

# Learning Places from Views: Variation in Scene Processing as a Function of Experience and Navigational Ability

Russell A. Epstein, J. Stephen Higgins, and Sharon L. Thompson-Schill

## Abstract

■ Humans and animals use information obtained from different viewpoints to form representations of the spatial structure of the world. We used functional magnetic resonance imaging (fMRI) adaptation to investigate the neural basis of this learning process and to show how the concomitant representations vary across individuals as a function of navigational ability. In particular, we examined the effect of repeating viewpoint and/or place information over both short (within-trial) and long (across-scan) intervals on the neural response in scene processing regions. Short-term fMRI adaptation effects in the parahippocampal cortex were initially highly viewpoint-specific but

became less so over time. Long-term fMRI repetition effects included a significant viewpoint-invariant component. When individual differences in navigational ability were considered, a significant correlation between the strength of these effects and self-reported navigational competence was observed. In particular, good navigators encoded representations that differed between new and old views and new and old places, whereas bad navigators did not. These results suggest that cortical scene representations evolve over time to become more viewpoint-invariant and that the quality of these representations directly impacts navigational ability. ■

## INTRODUCTION

Successful navigation involves the ability to encode and access representations of the spatial structure of the world. What is the nature of these representations in humans? Visual information is initially obtained from a particular point of view; however, navigation is facilitated by the use of world-centered representations that encode the intrinsic relationships between different locations (Tolman, 1948). Not surprisingly, behavioral and neuroscientific studies have indicated that humans and animals encode a variety of representations of local space (e.g., Mou & McNamara, 2002; Shelton & McNamara, 2001) including some that are tied to closely viewpoint (Chua & Chun, 2003; Nakatani, Pollatsek, & Johnson, 2002; Diwadkar & McNamara, 1997) and some that are viewpoint-independent (Burgess, 2002; King, Burgess, Hartley, Vargha-Khadem, & O'Keefe, 2002; Presson, DeLange, & Hazelrigg, 1989; O'Keefe & Nadel, 1978). These results could be unified by postulating a process by which viewpoint-invariant scene representations are learned from viewpoint-specific representations over time; however, to our knowledge, no direct evidence for such a process has yet been presented.

We address this issue by using functional magnetic resonance imaging (fMRI) adaptation to identify viewpoint-specific and viewpoint-invariant representations in cortical scene processing regions. fMRI adaptation is

defined as reduced response to repeated information compared to the response to novel information. This phenomenon can be used to determine the extent to which stimuli that are not physically identical are representationally similar with respect to processing within a particular cortical region (van Turennout, Bielsamowicz, & Martin, 2003; Grill-Spector & Malach, 2001; Kourtzi & Kanwisher, 2001; Henson, Shallice, & Dolan, 2000; James, Humphrey, Gati, Menon, & Goodale, 1999; Thompson-Schill, D'Esposito, & Kan, 1999; Buckner et al., 1998; Schacter & Buckner, 1998). In an earlier experiment, we used this technique to identify viewpoint-specific scene representations in the parahippocampal cortex (Epstein, Graham, & Downing, 2003). Here we extend these results by tracking the evolution of cortical scene representations over the course of an experimental session as subjects gained experience with different views of the same place. To anticipate, we report evidence that these representations may become less viewpoint-specific and more viewpoint-invariant over time.

fMRI adaptation effects were measured over both short (within-trial) and long (across-scan) repetition intervals. Previous studies that have used fMRI adaptation to identify object representations in the occipito-temporal cortex have given diverging results for short and long repetition intervals. For example, Grill-Spector, Kushnir, et al. (1999) found that short-term repetition effects in ventral occipital-temporal object-processing regions were viewpoint specific, whereas Vuilleumier,

---

University of Pennsylvania

Henson, Driver, and Dolan (2002) found that long-term repetition effects in the same regions in the left hemisphere were viewpoint invariant. We included both short and long repetition intervals within the same experiment to ensure that any effects found were not specific to the repetition technique used.

In addition to examining how scene representations vary over time, we also examined how they vary across individuals as a function of navigational competence by correlating fMRI adaptation effects with scores on the Santa Barbara Sense of Direction (SBSOD) scale (Hegarty, Richardson, Montello, Lovelace, & Ilavanil, 2002). The SBSOD is a 15-item survey in which subjects rate their agreement (on a scale of 1–7) with a number of statements about their navigational skill (e.g., “I very easily get lost in a new city.”). Previous work has demonstrated that the SBSOD has a high degree of test–retest reliability and that SBSOD scores predict performance on objective tests that require one to update one’s location and orientation in space but do not predict performance on tests of object rotation (Kozhevnikov & Hegarty, 2001). By measuring the correlation between fMRI adaptation effects and SBSOD scores, we aimed to establish a direct link between the cortical scene representations measured in this experiment and navigational ability. We specifically predicted that adaptation effects would be larger in good navigators than in bad navigators because good navigators should encode representations that more strongly differentiate between new and old places and new and old views.

## RESULTS

Each scan session was divided into three parts. In the first part (Scans 1–3), subjects viewed pairs of sequentially presented photographs of indoor and outdoor scenes (Figure 1) and reported whether the scenes were identical (no-change), different views of the same place (viewpoint-change), or photographs of two different places (place-change). In the second part (Scan 4), subjects made indoor/outdoor judgment on single-scene photographs, which were either identical to those used in the first part of the experiment (old views), previously unseen views of previously seen places (new views) or completely novel places (new places). Thus, pairwise comparisons of the three conditions in Scans 1–3 allowed us to examine the effect of repeating place or viewpoint over a short time scale of hundreds of milliseconds, whereas pairwise comparisons of the three conditions in Scan 4 allowed us to examine the effect of repeating place or viewpoint over a longer time scale of minutes. In the third part of the experiment (localizer scans), subjects viewed color photographs of scenes and objects in a block design in order to determine functional regions of interest as described previously (Epstein & Kanwisher, 1998).

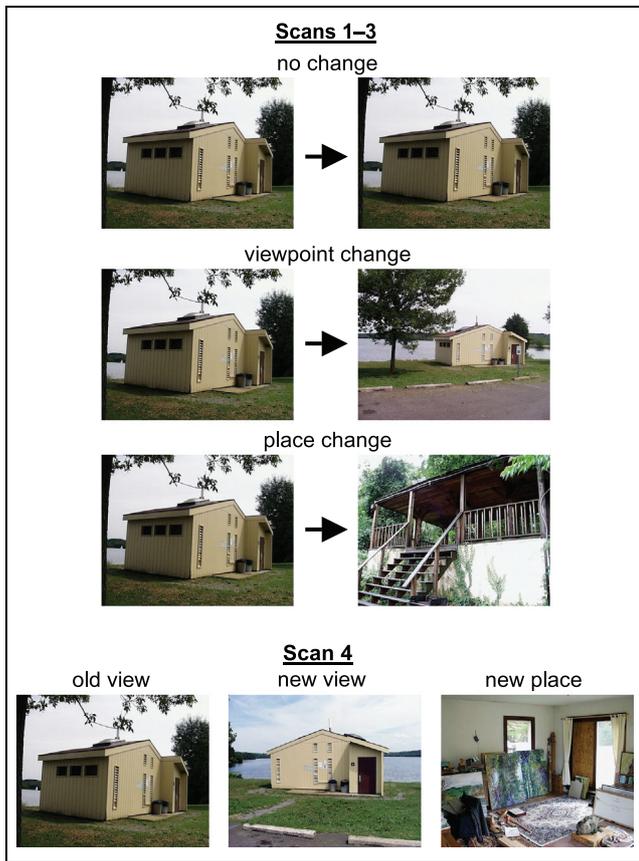
In Scans 1–3, the place-change condition is the baseline and the critical question is whether regional activity is reduced when the second photograph shows the same scene from a different viewpoint (viewpoint-change trials) or if it is only reduced when the second photograph shows the same scene from the same viewpoint (no-change trials). We interpret a reduced response to viewpoint-change trials relative to place-change trials as evidence for viewpoint-invariant representations and reduced response to no-change trials relative to viewpoint-change trials as evidence for viewpoint-specific representations. Similarly, in Scan 4, we interpret reduced response to new views relative to new places as evidence for viewpoint-invariant representations and reduced response to old views relative to new views as evidence for viewpoint-specific representations. Note that viewpoint-invariant and viewpoint-specific effects may coexist within the same region if the region supports scene representations that are neither completely viewpoint-invariant nor completely viewpoint-specific. In this article, we use the terms “fMRI adaptation” and “fMRI repetition reduction” interchangeably to refer to reductions of fMRI response due to repetition at both short (within-trial) and long (across-trial) time intervals.

## Behavioral Results

In Scans 1–3, subjects responded more quickly in the no-change condition ( $M = 1063$  msec) than in the viewpoint-change ( $M = 1183$  msec) or place-change conditions ( $M = 1159$  msec). Analysis of variance revealed that this difference was significant [ $F(2,22) = 11.1, p < .001$ ]; however, reaction times did not differ significantly between the viewpoint- and place-change conditions ( $p > .15$ ). There were no difference in response times across the three conditions in Scan 4 [old-view,  $M = 1003$ ; new-view,  $M = 1018$ ; new-place,  $M = 1032$ ;  $F(2,22) = 1.7, ns$ ]; however, the pairwise comparison between the old-view and new-place condition approached significance [ $t(11) = 1.9, p = .08$ ].

## Functional Regions of Interest

A group analysis of the data from the localizer scans indicated that three bilateral regions responded reliably more strongly across subjects to scenes than to objects: the parahippocampal place area (PPA) (Epstein, Harris, Stanley, & Kanwisher, 1999; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Aguirre, Zarahn, & D’Esposito, 1998; Epstein & Kanwisher, 1998), retrosplenial cortex (Maguire, 2001), and a region near the transverse occipital sulcus (TOS) (Grill-Spector, 2003; Hasson, Harel, Levy, & Malach, 2003; Nakamura et al., 2000) (Figure 2). Analysis of the data from individual subjects allowed us to functionally identify the left and right PPA and the left TOS region in all 12 subjects, the right TOS region in 10/12 subjects, the left retrosplenial region in



**Figure 1.** Examples of stimuli and schematic illustration of experimental design. In Scans 1–3, subjects viewed sequentially presented pairs of environmental scenes, which could be identical (no-change condition), different views of the same place (viewpoint-change condition), or different places (place-change condition). In Scan 4, the same subjects viewed single scenes, which could be identical to those viewed in Scans 1–3 (old views), previously unseen views of the places viewed in Scans 1–3 (new views), or previously unseen places (new places).

8/12 subjects, and the right retrosplenial region in 6/12 subjects. The PPA and TOS regions were marked in individual subjects for further analysis.

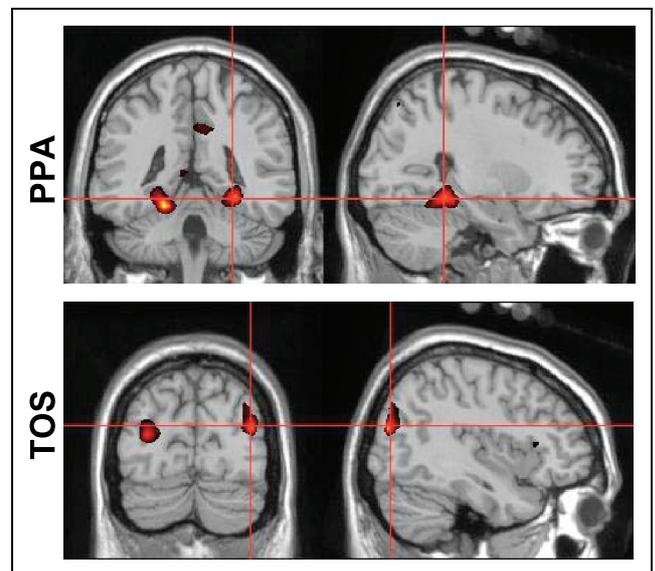
### fMRI Adaptation Effects

In Scans 1–3, fMRI adaptation was observed in the left and right PPA and the right TOS when place and viewpoint information were both repeated within a trial, as evidenced by reduced response in the no change condition relative to the place change condition [RPPA  $t(11) = 3.29, p < .01$ ; LPPA  $t(11) = 3.44, p < .01$ ; LTOS  $t(9) = 2.59, p < .05$ ]. A similar, although nonsignificant, trend was observed in the left TOS [ $t(11) = 1.88, p = .08$ ]. The portion of these effects attributable to repeating viewpoint is indexed by the difference between the viewpoint-change and no-change condition, whereas the portion of this effect attributable to repeating place information (while varying viewpoint) is indexed by the

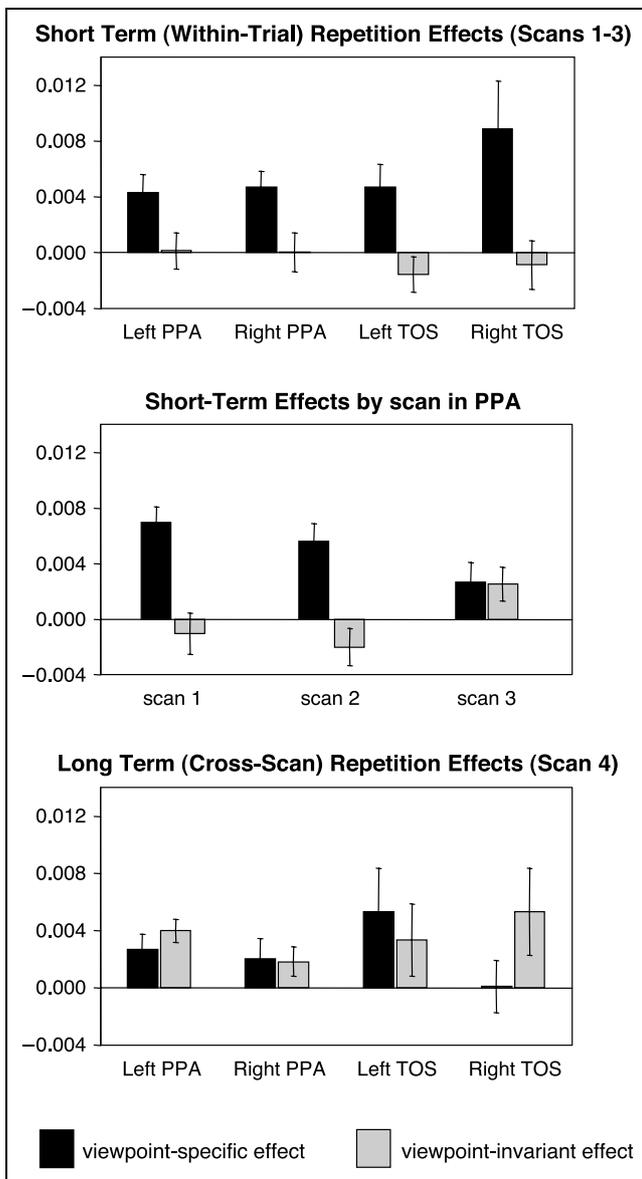
difference between the place- and viewpoint-change conditions. The viewpoint-specific effect was significant in the PPA and TOS [viewpoint-change > no-change: RPPA  $t(11) = 4.24, p < .002$ ; LPPA  $t(11) = 3.18, p < .01$ ; RTOS  $t(9) = 2.58, p < .05$ ; LTOS  $t(11) = 2.92, p < .02$ ], whereas the viewpoint-invariant effect was not (place change > viewpoint-change: all  $ts < 1$ ) (see Figure 3, top). Despite the smaller number of subjects for which a retrosplenial region of interest could be defined, the results in this area were similar (data not shown).

We also analyzed Scans 1 to 3 separately to determine whether the observed pattern of viewpoint specificity was stable over time or whether it changed as subjects gained experience with the scenes over the course of the scan session (see Figure 3, middle). As can be seen, the viewpoint-specific effect (viewpoint change > no change) became weaker over the course of Scans 1 to 3, whereas a viewpoint-invariant repetition effect (place change > viewpoint change) became evident by Scan 3. This Effect  $\times$  Scan interaction was significant [ $F(2,11) = 3.8, p < .05$ ]. No significant interaction with time was observed in the TOS region. Thus, although scenes are initially encoded in a viewpoint-specific manner in the PPA, we found evidence that these representations become more viewpoint-invariant over time.

This conclusion was buttressed by the data from Scan 4, which measured the effect of repeating place and/or viewpoint information over a longer time interval (see Figure 3, bottom). As before, significant MR signal



**Figure 2.** Scene processing regions. The PPA and the TOS responded more strongly to scenes than to objects in functional localizer scans. Here the results of a random-effects group analysis are shown. Voxels responding preferentially to scenes ( $p < .001$  uncorrected) are overlaid on top of a reference brain in standard space. Subject-specific regions of interest were defined in these areas for further analysis. Talairach coordinates of the crosshairs are (27 -40 -12) for the right PPA and (40 -78 22) for the right TOS.



**Figure 3.** fMRI adaptation effects. Top: Short-term (within-trial) fMRI repetition effects in the PPA and TOS for Scans 1–3. The viewpoint-specific effect is indexed by greater activity in the viewpoint-change than in the no-change condition, whereas the viewpoint-invariant effect is indexed by greater activity in the place-change than in the viewpoint-change condition. Error bars represent 1 SEM. Units on the y-axis are mean beta values. Middle: Interaction of the short-term fMRI repetition effects with time in the PPA (left and right combined). Over the course of Scans 1–3, the viewpoint-specific effect becomes weaker, whereas a viewpoint-invariant effect begins to develop. Bottom: Long-term (across-scan) fMRI repetition effect in the PPA and TOS for Scan 4. The viewpoint-specific effect is indexed by greater activity in the new-view than in the old-view condition, whereas the viewpoint-invariant effect is indexed by greater activity in the new-place than in the new-view condition.

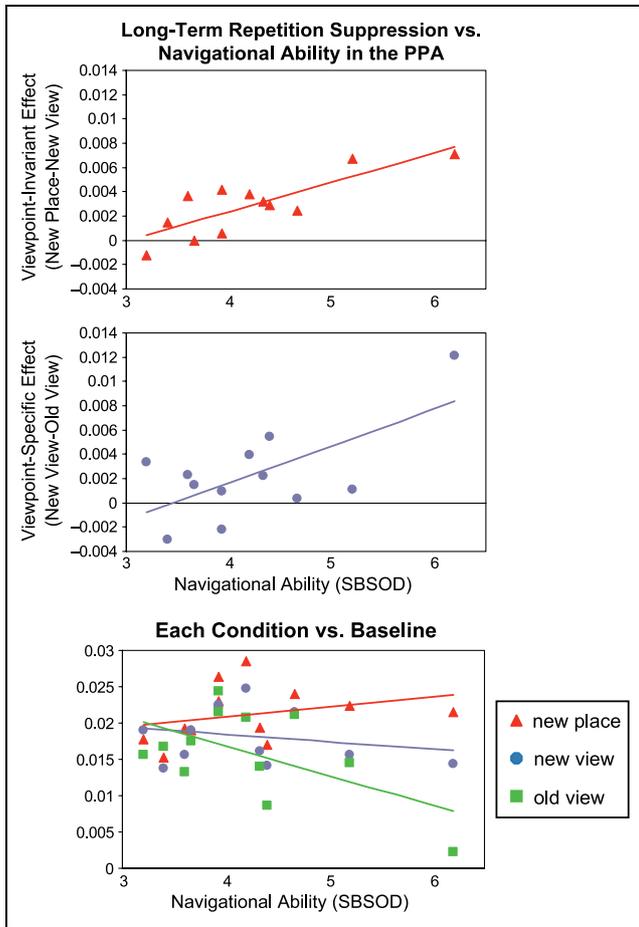
reduction was observed when viewpoint and place information were both repeated (now over a longer time interval) in the left PPA [ $t(11) = 4.54, p < .001$ ] and left TOS [ $t(11) = 4.17, p < .002$ ] with a similar, although

nonsignificant, trend in the right PPA [ $t(11) = 2.01, p = .07$ ]. In this part of the experiment, the repetition effect was neither completely viewpoint-invariant nor completely viewpoint-specific; rather, both kinds of effects were observed. These effects were significant in the left PPA [new place vs. new view:  $t(11) = 4.9, p < .001$ ; new view vs. old view:  $t(11) = 2.5, p < .05$ ]. A similar pattern was observed in the other regions, however, the effects in these regions did not reach significance. The fact that these effects were unreliable may be due to the small number of trials in Scan 4 (30 for each condition vs. 45 or more in Scans 1–3).

### Correlations with Navigational Ability

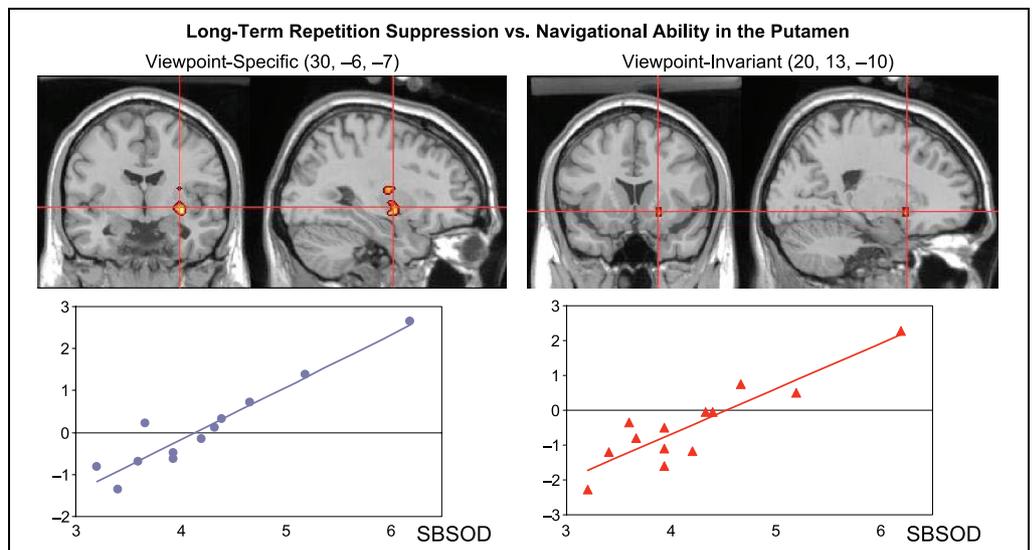
Are the repetition effects observed in this experiment indexing representations involved in navigation? To answer this question, we analyzed the relationship between the strength of these effects and individual differences in navigational ability, as assessed by the SBSOD scale (Hegarty et al., 2002). We observed a significant correlation between navigational ability and fMRI repetition effects in the PPA (Figure 4, top). The relationship between these two measures was most evident in Scan 4. SBSOD scores reliably predicted the strength of both the viewpoint-invariant (left:  $r = .58, p < .05$ ; right:  $r = .60, p < .05$ ; combined  $r = .81, p < .01$ ) and viewpoint-specific (left:  $r = .60, p < .05$ ; right:  $r = .66, p < .05$ ; combined  $r = 0.65, p < .05$ ) repetition effects in this region. When the subject with the highest SBSOD score was removed from the analysis, the correlation between the viewpoint-invariant repetition effect and navigational ability remained significant (left and right combined:  $r = .74, p < .01$ ) but the correlation with the viewpoint-specific effect was no longer reliable. Correlations were also found in Scans 1–3: SBSOD scores reliably predicted within-trial viewpoint-invariant adaptation (place change > view change) in the right PPA ( $r = .67, p < .02$ ). SBSOD scores did not correlate with short- or long-term repetition effects in the TOS region (all  $ps > .1$ ).

There are two possible scenarios that could account for the larger repetition effects in the good navigators than in the bad navigators. First, the good navigators might have developed more efficient representations of the familiar views and places, leading to reduced response to old views relative to new views and to new views of old places relative to new places. Alternatively, the response to all stimuli might be greater in the good navigators than in the bad navigators (perhaps because good navigators find topographical stimuli particularly interesting and thus attend to them more). If this were the case, then one might observe repetition effects that were larger in absolute terms in the good navigators but which were proportionately the same and thus *not* indicative of representations that are qualitatively different. To distinguish between these possibilities, we



**Figure 4.** Correlations with navigational ability in the PPA. Top: Correlation between navigational ability and viewpoint-specific (blue) and viewpoint-invariant (red) repetition effects in the PPA. Data are from Scan 4; left and right PPA have been combined. Navigational ability is indexed by scores on the SBSOD scale. Bottom: Response in all three conditions versus intertrial baseline as a function of navigational ability. The differences between conditions are larger in good navigators than in bad navigators even though the overall magnitude of the fMRI response does not increase.

**Figure 5.** Navigational ability predicted repetition effects in the putamen in Scan 4. Highlighted voxels are those in which a whole-brain analysis revealed a significant ( $p < .0001$  uncorrected) correlation between SBSOD score and viewpoint-specific (left) and viewpoint-invariant (right) repetition suppression. The relationship between SBSOD scores and the corresponding effects in these voxels is also plotted. Units on the  $y$ -axis are mean  $t$  values.



replotted the data from Scan 4 to show the response in all three conditions relative to the intertrial baseline (Figure 4, bottom). As can be seen, there is no evidence that the magnitude of the response to scenes is greater in good navigators than in bad navigators; indeed, if anything, the pattern is the opposite.

## Whole-Brain Correlational Analyses

We performed an exploratory whole-brain analysis to determine whether any areas outside of our regions of interest exhibited similar correlations between repetition reduction and navigational ability. Results for Scan 4 are shown in Figure 5. Although no regions showed differential response at the corrected level, when the threshold was reduced to  $p < .0001$  uncorrected [ $t(10) > 6.21$ ], a striking correlation between both viewpoint-specific and viewpoint-invariant long-term repetition reduction and navigational ability was observed in the putamen (viewpoint-specific:  $r = .96$ ,  $p < .00001$ ; viewpoint-invariant:  $r = .92$ ,  $p < .0001$ ). The loci for the correlations with the viewpoint-invariant effect were approximately 2 cm anterior to the locus for the correlation with the viewpoint-specific effect. No regions exhibited correlation between short-term viewpoint-specific repetition reduction and SBSOD score at this threshold in Scans 1–3; however, a correlation between SBSOD and short-term viewpoint-invariant repetition reduction was found in the cingulate near the corpus callosum ( $r = .93$ ,  $p < .00001$ ,  $x = 1$ ,  $y = 17$ ,  $z = 15$ ).

## DISCUSSION

This study had three main aims: (1) to measure the extent to which representations in scene processing regions are viewpoint-specific or viewpoint-invariant; (2) to track the development of these representations

over the course of an experimental session; (3) to show how these representations vary across individuals as a function of navigational ability. We found evidence that scene representations in the PPA are initially viewpoint-specific but become more viewpoint-invariant over time. In addition, we found evidence that both viewpoint-specific and viewpoint-invariant scene representations in the PPA are more reliable in good navigators than in bad navigators. We will address each of these results in turn.

Using a short-term repetition paradigm, we found evidence that scene representations in the PPA and TOS regions are largely viewpoint-specific, replicating results from an earlier study that used tabletop scenes as stimuli (Epstein, Graham, et al., 2003). In particular, we found that response to place-change and viewpoint-change trials was equivalent, whereas response to no-change trials was reduced, indicating that (at least initially) different views of the same place are as representationally distinct as views of different places. Thus, at first glance, our current data seem to support our earlier conclusion that scene representations in the PPA are viewpoint-specific and allow us to extend these conclusions to the TOS region. However, this interpretation must be qualified when the interaction with time is considered. By analyzing the data for Scans 1–3 separately, we were able to track the evolution of the short-term repetition effects as subjects gained experience with the individual images used. We observed an interaction in which the viewpoint-specific effect became weaker, whereas a viewpoint-invariant developed over the course of the scan session. We then examined the effect of repeating place and/or viewpoint over longer intervals by contrasting the response to novel views and places in Scan 4 with the response to views and places that had been seen several times previously. Here we found concurring evidence for the existence of a viewpoint-invariant effect. In sum, our data suggest that although scene representations in the PPA region are initially encoded in a viewpoint-specific manner, they evolve to become partially viewpoint-invariant.

These results have important consequences for our understanding of the functional role of scene-processing regions (Epstein, in press). The PPA responds strongly when scenes are viewed both in the context of a navigational task (Maguire, Burgess, et al., 1998; Aguirre, Detre, Alsop, & D'Esposito, 1996) and also when no such task is performed (Epstein, Harris, et al., 1999; Epstein & Kanwisher, 1998). Patients with damage to this region often suffer from severe navigational impairments (Mendez & Cherrier, 2003; Barrash, Damasio, Adolphs, & Tranel, 2000; Aguirre & D'Esposito, 1999; Habib & Sirigu, 1987). In a previous study of two of these patients, we found that they were particularly impaired at the encoding of scenes (but not objects) into memory (Epstein, DeYoe, Press, Rosen, & Kanwisher, 2001). In contrast, their on-line perception of scenes appeared to be relatively unimpaired. Taken as a whole with the

previous literature, the current data suggest that the parahippocampal cortex may be critical for navigation because it is the locus of a mechanism for the learning of places from individual views. This mechanism may work in concert with hippocampal learning mechanisms that represent places in a manner that is completely viewpoint-independent (Ekstrom et al., 2003; Burgess, 2002; King et al., 2002). Importantly, we cannot tell from the current data whether the formation of viewpoint-invariant representations in the PPA depends on the viewing of more than one view of the same place or whether a similar generalization process may occur even for scenes viewed from only a single vantage point (Sanocki, 2003; Intraub, Bender, & Mangels, 1992). Furthermore, it is currently unclear how the addition of self-movement information available during more realistic navigational episodes might affect these results. Previous studies suggest that PPA response is reduced when individual scenes are familiar but is less affected by familiarity with the larger environment in which those scenes are embedded (Epstein, Harris, et al., 1999); in contrast, hippocampal activity does depend on environmental familiarity (Maguire, Burgess, et al., 1998). Future experiments should address these issues.

These results also have the potential to inform a long-standing debate in the object recognition literature. Behavioral, neurophysiological, and neuroimaging studies have found evidence for both viewpoint-specific (Grill-Spector, Kushnir, et al., 1999; Tarr, Williams, Hayward, & Gauthier, 1998; Logothetis & Pauls, 1995) and viewpoint-invariant (Vuilleumier et al., 2002; Biederman & Gerhardstein, 1993) object representations, sometimes interspersed within the same cortical region (Booth & Rolls, 1998). The issue of which of these representations is most critical for recognition has been widely discussed. Although the present data do not resolve this debate, they do suggest that viewpoint-specific and viewpoint-invariant scene representations coexist in some cortical regions and that the viewpoint-invariant representations might be formed by reorganization of the viewpoint-specific representations. A similar process may occur for nonscene objects; indeed, previous neurophysiological studies have found evidence for reorganization of high-level object representations with experience (Erickson & Desimone, 1999; Sakai & Miyashita, 1991).

One unresolved issue is whether the adaptation effects observed in this experiment relate to repetition of the spatial aspects of the scene (i.e., viewpoint and place) or whether they relate to repetition of nonspatial visual information. For example, the reduced response in the no change condition relative to the viewpoint change condition might reflect the fact that the two images in the viewpoint-change condition are taken from different viewpoints, or it might simply reflect the fact that the two images in the viewpoint-change condition are visually dissimilar. In our previous study

(Epstein, Graham, et al., 2003), we demonstrated that the PPA was not sensitive to highly salient visual changes caused by manipulation of foreground objects; however, it is an open question whether the PPA would be sensitive to visual changes to the background elements of the scene that are not caused by a change in viewpoint. For example, one might predict that the PPA (and TOS) would initially encode two images of the same scene taken from the same viewpoint but under vastly different lighting conditions as being representationally distinct. In this view, the learning of “viewpoint-invariant” scene representations in this experiment may reflect the operation of a more general learning mechanism that ties together mental snapshots of the same place under different viewing conditions.

A related issue is the extent to which the adaptation effects reflect repetition of different aspects of the stimulus as opposed to repetition of task-related processes (Henson & Rugg, 2003). The fact that similar effects were observed both with a scene-comparison task (Scans 1–3) and with an indoor/outdoor judgment task (Scan 4) suggests that we are observing underlying aspects of cortical scene representations rather than task-related confounds. However, we cannot make the claim that these adaptation effects are entirely task-independent. In particular, some tasks (e.g., making an indoor/outdoor judgment) may require subjects to attend to aspects of the scene that are invariant across views, whereas other tasks may require them to attend to aspects of the scenes that are specific to individual views. The resulting adaptation effects may reflect these different attentional regimes.

The final goal of this study was to examine how cortical scene representations vary over individuals as a function of navigational ability. We found significant correlations between the magnitude of both viewpoint-specific and viewpoint-invariant fMRI repetition reductions and self-reported navigational ability in the PPA. These results suggest that representational differences between new and old places and new and old views are more salient in good navigators than in bad navigators. Insofar as these differences are evidenced by reduced response to the repeated stimuli in the good navigators, it may be the case that representations of familiar places and views are more efficient in good navigators than in bad navigators, and consequently, more useful for distinguishing between different places and different views (Wiggs & Martin, 1998). Navigational ability did not predict repetition effects in the TOS region, indicating that this region is less involved than the PPA in encoding representations that are used in navigational planning. These results further strengthen the argument that PPA scene representations are critical for spatial navigation.

An unexpected finding was that navigational ability correlated strongly with fMRI repetition effects in the basal ganglia. Although recent neuroimaging studies

have linked activity in the head of the caudate to the use of information about stimulus–response contingencies in navigation (Hartley, Maguire, Spiers, & Burgess, 2003; Iaria, Petrides, Dagher, Pike, & Bohbot, 2003), the activation observed here has a different locus—the putamen. Interestingly, neurophysiological studies have identified place and head direction cells in the basal ganglia of the rat (Ragozzino, Leutgeb, & Mizumori, 2001; Mizumori, Ragozzino, & Cooper, 2000). Furthermore, lesion studies have demonstrated a dissociation whereby damage to medial striatal regions lead to difficulties in making responses on the basis of stored spatial information, whereas damage to lateral striatal regions leads to difficulties making responses on the basis of visible cues (Devan & White, 1999). We hypothesize that when good navigators view a scene, they activate basal ganglia mechanisms involved in selecting an appropriate locomotor response (Wise, Murray, & Gerfen, 1996). Insofar as the appropriate response will differ for different viewpoints and different places, repetition reductions will be observed in this region for these subjects. In contrast, bad navigators do not prepare a response, so repetition reductions are not observed. The impressive strength of the correlation between cognitive ability and repetition effects in the basal ganglia suggests that the role of these structures in mediating our sense of direction deserves more attention.

These findings provide two clear examples of the value of individual differences analyses in the study of brain–behavior relations. First, only by examining the correlation between navigational ability and fMRI repetition effects were we able to identify a region of the basal ganglia that is potentially involved in scene processing. This area was not evident on group analyses because of the inconsistencies in the size of the effect across subjects. By accounting for individual differences, it might also be possible to detect small effects that would otherwise be swamped by explainable variance that is being treated as noise in group studies. Second, adopting an individual differences approach allows one to potentially dissociate effects that appear similar in a group analysis. In the current experiment, the group analyses indicated significant repetition reduction in both the PPA and TOS. However, correlations with navigational ability were only reliable in the former. This is potentially indicative of different roles for the PPA and TOS in spatial processing and navigation and suggests the need to contrast these areas in future studies. In general, correlations between physiological effects in a brain region and individual differences in a given process help to cement the relation between structure and function. Similar examples of the utility of such an approach can be seen in other neuroimaging studies of visual acuity (Duncan & Boynton, 2003), emotion regulation (Schaefer et al., 2002), and general intelligence (Gray, Chabris, & Braver, 2003).

In sum, the present results demonstrate that both viewpoint-specific and viewpoint-invariant representations are supported by scene processing regions and that they vary as a function of both experience and navigational ability. These regions appear to be critically involved in the transformation of information obtained from individual views into knowledge of the spatial structure of the world.

## METHODS

### Subjects

Fifteen healthy right-handed volunteers were recruited from the local community and gave informed consent according to procedures approved by the University of California, Berkeley. Data from 3 subjects were not analyzed because of data processing errors. Of the 12 remaining subjects, 8 were women, and the median age was 22.5 years.

Navigational ability was operationalized by scores on the SBSOD (Hegarty et al., 2002). Scores on this scale ranged from 3.2 to 6.2, with a mean of 4.2 and an *SD* of 0.84. These numbers were comparable to the mean of 4.7 and *SD* of 1.1 obtained by Hegarty et al. (2002) using a larger sample size of 107 undergraduates (Study 6). SBSOD scores did not differ significantly between men and women ( $t < 1$ ).

### fMRI Parameters

Scanning was carried out on a 4-Tesla Varian INOVA scanner at the Henry Wheeler Brain Imaging Center at UC Berkeley. T2\*-weighted images sensitive to blood oxygenation level-dependent contrasts were acquired using a two-shot gradient-echo echo-planar image sequence (field of view = 22.4 cm, matrix size = 64 × 64, repetition time = 1.1 sec per half of *k*-space, echo time = 28 msec, flip angle = 20°). Each functional volume consisted of twenty 5-mm-thick axial slices with 0.5-mm interslice gaps. High-resolution MP-Flash 3-D T1-weighted scans were acquired for anatomical localization and normalization.

### Procedure

Scans 1 to 3 were each 530.2 sec long and divided into seventy-five 6.6 sec trials (15 place-change, 15 viewpoint-change, 30 no-change, and 15 “null” trials) with 16-sec fixation periods at the beginning and end. Stimulus trials began with a 500-msec presentation of a dark gray “cue” image. After a 500-msec gap, subjects viewed two scenes for 500 msec each with a 500-msec interstimulus interval, followed by a 4100-msec poststimulus interval during which a fixation cross appeared on the screen and subjects used a button box to report the trial type (i.e.,

place change, viewpoint change, or no change). This task was intended to ensure that subjects explicitly attended to viewpoint and scene identity. During null trials, the fixation cross remained on the screen for 6.6 sec and subjects made no response. Stimuli were 30 digitized color photographs depicting two views of 15 unfamiliar indoor and outdoor locations. View pairs were chosen so that most of the same objects and topographical features were visible in both images; a separation of 30–40° between viewpoints was typical (see Figure 1 for an example). The same set of photographs was used to construct all three conditions in all three scans. Subjects saw each photograph 4 times within a scan (twice in the no-change condition, once in the viewpoint-change condition, and once in the place-change condition) and 12 times over the course of Scans 1–3. Thus, by measuring the response to each condition in each scan separately, it was possible to measure how these responses changed as the stimuli became more familiar to the subjects. Pairings between images in the place-change condition were randomized for each scan.

Scan 4 was 563.2 sec long and was divided into one hundred twenty 4.4-sec long events (30 new-place, 30 new-view, 30 old-view, and 30 “null” events) with 16-sec fixation periods at the beginning and end. During stimulus events, subjects viewed a single digitized color photograph for 500 msec and then used a button box to report whether it depicted an indoor or outdoor environment. This task required subjects to attend to the photographs without requiring them to make any explicit memory judgments. Accuracy rather than speed was stressed; however, subjects were instructed not to take any more time than necessary to make a decision. Stimuli in the old-view condition were the same set of 30 photographs used in the first part of the experiment. Stimuli in the new-view condition were photographs depicting two previously unseen views of the 15 places seen in the first part of the experiment. Stimuli in the new-place condition were photographs of 15 previously unseen unfamiliar places (2 views each). Order of event presentation was randomized in all scans subject to the constraint that events of every type were preceded equally often by events of every other type. The assignment of places and views to conditions was randomized across subjects to ensure that differences between conditions could not be attributed to stimulus differences.

In the two localizer scans, subjects viewed digitized color photographs of faces, common objects, landscapes, cityscapes, buildings, and scrambled objects. Each scan was 409.2 sec long and was divided into eighteen 17.6-sec picture epochs (3 epochs for each of the stimulus categories) interleaved with seven 13.2-sec epochs during which the screen was blank except for a fixation point. During each picture epoch, 20 photographs of the same type were presented for 400 msec each with a 480-msec interstimulus interval. Subjects

performed a one-back task in which they were required to press a button whenever two identical stimuli appeared in a row. There were two such repetitions in each epoch.

## Data Analysis

Functional images were preprocessed using SPM99 and analyzed using VoxBo ([www.voxbo.org](http://www.voxbo.org)). Images were corrected for differences in slice timing by resampling slices in time to match the first slice of each volume, and then realigned with respect to the first image of the scan using sinc interpolation. The mean realigned image was then normalized to the Montreal Neurological Institute (MNI) template using a set of affine and smoothly nonlinear transformations. This transformation was then separately applied to all functional images for the subject. Images were then resampled into 3-mm isotropic voxels and spatially smoothed with an 8-mm FWHM gaussian filter. Data were analyzed using the general linear model as implemented in VoxBo, including an empirically derived  $1/f$  noise model and filters that removed high and low temporal frequencies. The model included regressors to account for global signal and motion-specific effects.

Scene processing regions of interest were defined for each subject using data from the localizer scans. Each of the stimulus types was modeled as a boxcar function convolved with a canonical hemodynamic response function. Linear contrasts were used to identify clusters of contiguous voxels in the parahippocampal, retrosplenial, and TOS regions that responded more strongly ( $t > 2.5$  or  $3.5$ ) to landscapes and cityscapes than to objects. The time courses of activation during the mean experimental scans were then extracted and averaged over all voxels within the regions of interest. Three different analyses were then performed on these time courses. In the first, data from Scans 1 to 3 were fitted to a general linear model that included regressors for the three different event types (modeled as impulse functions convolved with a canonical hemodynamic response) and nuisance regressors to account for between-scan differences. The second analysis was similar except that data from Scans 1 to 3 were modeled separately. The third analysis modeled data from Scan 4. In all cases, beta values were calculated for each condition and used as the dependent variable in a random-effects analysis. Note that these analyses treat data from within a region of interest as if it were from a single (average) voxel so no correction for multiple comparisons across voxels is necessary. Whole-brain analyses were also performed to identify the strength of the contrasts of interest in each voxel of the brain and the cross-subject correlation between these values and SBSOD scores. For this analysis, a variance-normalized measure of contrast strength was used. The results for all analyses were qualitatively

the same irrespective of whether the measure of contrast strength was variance normalized or not.

## Acknowledgments

We are grateful for the invaluable assistance of members of the Henry H. Wheeler Jr. Brain Imaging Center at the University of California, Berkeley, including Charlotte Boettiger, Clayton Curtis, Jeffrey Cooney, Ben Inglis, John Jinks-Ollinger and Mark D'Esposito. We also thank Geoff Aguirre and Dan Kimberg for help with data analysis. This research was supported in part by grants from the Searle Scholars Program and NIH MH60414 (to S.T.-S.).

Reprint requests should be sent to Russell Epstein, 3720 Walnut Street, Philadelphia, PA, 19104-6241, or via e-mail: [epstein@psych.upenn.edu](mailto:epstein@psych.upenn.edu).

The data reported in this experiment have been deposited in the fMRI Data Center ([www.fmridc.org](http://www.fmridc.org)). The accession number is 2-2004-1166T.

## 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.
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to “building” stimuli: Evidence and implications. *Neuron*, *21*, 373–383.
- Barrash, J., Damasio, H., Adolphs, R., & Tranel, D. (2000). The neuroanatomical correlates of route learning impairment. *Neuropsychologia*, *38*, 820–836.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.
- Booth, M. C., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex*, *8*, 510–523.
- Buckner, R. L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., Rosen, B., & Dale, A. M. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron*, *20*, 285–296.
- Burgess, N. (2002). The hippocampus, space, and viewpoints in episodic memory. *Quarterly Journal of Experimental Psychology Section A*, *55*, 1057–1080.
- Chua, K. P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception and Psychophysics*, *65*, 72–80.
- Devan, B. D., & White, N. M. (1999). Parallel information processing in the dorsal striatum: Relation to hippocampal function. *Journal of Neuroscience*, *19*, 2789–2798.
- Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science*, *8*, 302–307.
- Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*, 659–671.

- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature*, *425*, 184–188.
- Epstein, R. A. (in press). The cortical basis of visual scene processing. *Visual Cognition*.
- Epstein, R., DeYoe, E. A., Press, D. Z., Rosen, A. C., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsychology*, *18*, 481–508.
- Epstein, R., Graham, K. S., & Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron*, *37*, 865–876.
- Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron*, *23*, 115–125.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Erickson, C. A., & Desimone, R. (1999). Responses of macaque perirhinal neurons during and after visual stimulus association learning. *Journal of Neuroscience*, *19*, 10404–10416.
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*, 159–166.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203.
- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321.
- Habib, M., & Sirigu, A. (1987). Pure topographical disorientation—A definition and anatomical basis. *Cortex*, *23*, 73–85.
- Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*, 877–888.
- Hasson, U., Harel, M., Levy, I., & Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron*, *37*, 1027–1041.
- Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Ilavani, S. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*, 425–447.
- Henson, R. N., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*, 263–270.
- Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*, 1269–1272.
- Iaria, G., Petrides, M., Dagher, A., Pike, B., & Bohbot, V. D. (2003). Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: Variability and change with practice. *Journal of Neuroscience*, *23*, 5945–5952.
- Intraub, H., Bender, R. S., & Mangels, J. A. (1992). Looking at pictures but remembering scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 180–191.
- Ishai, A., Ungerleider, L. G., Martin, A., Schouten, H. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences, U.S.A.*, *96*, 9379–9384.
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (1999). Repetition priming and the time course of object recognition: An fMRI study. *NeuroReport*, *10*, 1019–1023.
- King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). The human hippocampus and viewpoint dependence in spatial memory. *Hippocampus*, *12*, 811–820.
- Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *293*, 1506–1509.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory and Cognition*, *29*, 745–756.
- Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cerebral Cortex*, *5*, 270–288.
- 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.
- Mendez, M. F., & Cherrier, M. M. (2003). Agnosia for scenes in topographagnosia. *Neuropsychologia*, *41*, 1387–1395.
- Mizumori, S. J., Ragozzino, K. E., & Cooper, B. G. (2000). Location and head direction representation in the dorsal striatum of rats. *Psychobiology*, *28*, 441–462.
- Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 162–170.
- Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., Hatano, K., Ito, K., Fukuda, H., Schormann, T., & Zilles, K. (2000). Functional delineation of the human occipito-temporal areas related to face and scene processing—A PET study. *Brain*, *123*, 1903–1912.
- Nakatani, C., Pollatsek, A., & Johnson, S. H. (2002). Viewpoint-dependent recognition of scenes. *Quarterly Journal of Experimental Psychology A*, *55*, 115–139.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press and Oxford University Press.
- Presson, C. C., DeLange, N., & Hazelrigg, M. D. (1989). Orientation specificity in spatial memory: What makes a path different from a map of the path? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 887–897.
- Ragozzino, K. E., Leutgeb, S., & Mizumori, S. J. (2001). Dorsal striatal head direction and hippocampal place representations during spatial navigation. *Experimental Brain Research*, *139*, 372–376.
- Sakai, K., & Miyashita, Y. (1991). Neural organization for the long-term-memory of paired associates. *Nature*, *354*, 152–155.
- Sanocki, T. (2003). Representation and perception of scenic layout. *Cognitive Psychology*, *47*, 43–86.
- Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, *20*, 185–195.
- Schaefer, S. M., Jackson, D. C., Davidson, R. J., Aguirre, G. K., Kimberg, D. Y., & Thompson-Schill, S. L. (2002). Modulation of amygdalar activity by the conscious

- regulation of negative emotion. *Journal of Cognitive Neuroscience*, *14*, 913–921.
- Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, *43*, 274–310.
- Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, *1*, 275–277.
- Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, *23*, 513–522.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208.
- van Turennout, M., Bielowicz, L., & Martin, A. (2003). Modulation of neural activity during object naming: Effects of time and practice. *Cerebral Cortex*, *13*, 381–391.
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, *5*, 491–499.
- Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, *8*, 227–233.
- Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex–basal ganglia system in primates. *Critical Reviews in Neurobiology*, *10*, 317–356.