

## The cortical basis of visual scene processing

Russell A. Epstein

*Department of Psychology, University of Pennsylvania, Philadelphia, USA*

Several lines of evidence suggest that the human brain contains special-purpose machinery for processing information about visual scenes. In particular, a region in medial occipitotemporal cortex—the “parahippocampal place area”, or PPA—represents the geometric structure of scenes as defined primarily by their background elements. Neuroimaging studies have demonstrated that the PPA responds preferentially to scenes but not to the objects within them, while neuropsychological studies have shown that damage to this region leads to an impaired ability to learn new scenes. More recent evidence suggests that the PPA encodes novel scenes in a viewpoint-specific manner and that these codes are more reliable in good navigators than bad navigators. The PPA may be part of a larger network of regions involved in processing navigationally relevant spatial information. The role of this region in place recognition and gist comprehension is also discussed.

How are visual scenes processed in the human brain? To answer this question, we first need to define the term “scene” in a way that distinguishes scenes from other kinds of visual stimuli. Henderson and Hollingworth (1999) provide the following definition: “[A] semantically coherent (and often nameable) view of a real-world environment comprising background elements and multiple discrete objects arranged in a spatially licensed manner”. Implicit in this definition is a contrast between scenes and nonscene objects such as cars, faces, and coffee cups, which do not have background elements and are not “environments”. A simple heuristic is that objects are spatially compact entities that one acts upon, while scenes are spatially distributed entities that one acts within. Thus, to answer the question of how scenes are processed in the cortex, we first need to determine whether there is special-purpose cortical machinery for processing scenes that can be distinguished from the machinery involved in processing nonscene objects.

---

Please address all correspondence to: Russell Epstein, Department of Psychology, University of Pennsylvania, 3720 Walnut St., Philadelphia, PA 19104-6241, USA.

Email: [epstein@psych.upenn.edu](mailto:epstein@psych.upenn.edu)

I thank Nancy Kanwisher and Sharon Thompson-Schill for comments on this manuscript.

---

© 2005 Psychology Press Ltd

<http://www.tandf.co.uk/journals/pp/13506285.html>

DOI:10.1080/13506280444000607

However, the project does not end there. According to Henderson and Hollingworth's (1999) definition, a scene has at least four primary aspects: (1) it is semantically coherent (i.e., it has a "gist"), (2) it depicts a real-world environment, (3) it contains background elements such as walls, and (4) it contains objects that are spatially organized in a meaningful way relative to each other and relative to the background elements. It is quite possible that there may be separate mechanisms for identifying a scene as a particular kind of environment or situation ("desert", "picnic", "party"), identifying the real-world location of the depicted environment ("North side of campus, next to the library"), processing the overall spatial structure of the scene as defined by the background elements, and processing individual objects and their locations. Thus, once we have identified regions of cortex involved in scene processing, we then need to find out which aspect of the scene the region processes.

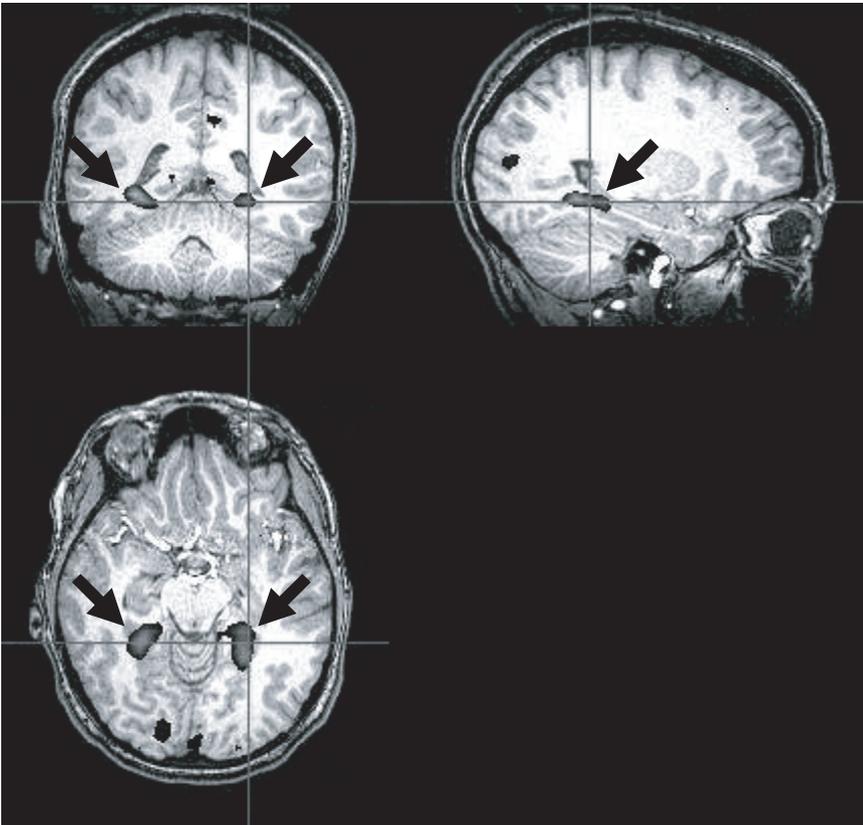
Here I review a series of neuroimaging and neuropsychological experiments that suggest that a particular region of the brain—parahippocampal cortex—is critical for at least one aspect of scene processing: Representing the spatial structure of the scene as defined primarily by its background elements. I will argue that this region is part of a larger cortical network involved in processing navigationally relevant spatial information. Indeed, one of the lessons of this work is that scenes are special because they depict environments, and thus convey critical information about where one can move.<sup>1</sup>

### THE PARAHIPPOCAMPAL PLACE AREA RESPONDS SELECTIVELY TO VISUAL SCENES

As a first step towards uncovering the cortical basis of scene processing, Kanwisher and I used functional magnetic resonance imaging (fMRI) to identify a region of the brain that responded selectively to visual scenes (Epstein & Kanwisher, 1998). In our first experiment, nine subjects were scanned while viewing greyscale photographs of faces, objects, and scenes presented in 16 s long "blocks". In all nine subjects, significantly greater activation was observed when subjects viewed scenes than when they viewed faces or objects in a bilateral region of posterior parahippocampal cortex (Figure 1). The locus of this activation—straddling the collateral sulcus near the parahippocampal/lingual boundary—was remarkably consistent across subjects. We named this region the "parahippocampal place area" (PPA) because it appeared to respond strongly to

---

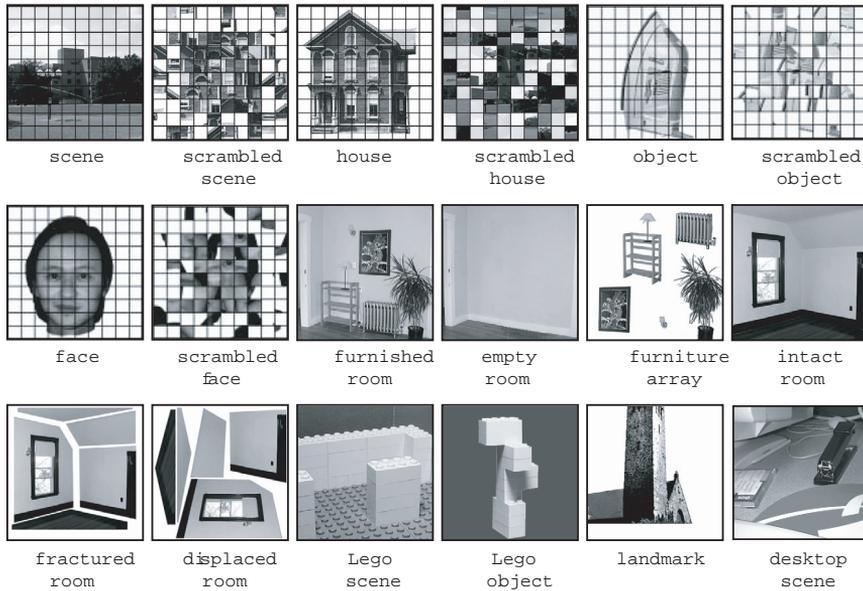
<sup>1</sup> A word about terminology. I will use the term "scene" in two ways: First, to refer to a *view* of a section of the world such as a room or a tabletop (as in Henderson and Hollingworth's, 1999, definition), and second, to refer to the section of the world itself. The meaning should be clear from the context; however, to forestall confusion, I will occasionally use the term "layout" to refer to the real-world entity. The term "place" also refers to the real-world entity, but with a different emphasis: Places are locations in the world, and thus have a spatial relationship to each other that cannot always be ascertained from a single view.



**Figure 1.** The parahippocampal place area in one subject. Voxels responding significantly more strongly ( $t > 3.5$ ) to scenes than to objects during blocked viewing are highlighted. Right hemisphere is on the right. The PPA is generally found straddling the collateral sulcus near the parahippocampal/lingual boundary in both hemispheres.

photographs of places (i.e., indoor and outdoor scenes). Importantly, the region of activation did not include the hippocampus proper.

We then used a region of interest analysis to examine the response in the PPA more closely. In this analysis, the PPA is defined for each subject using functional data from a set of independent “localizer” scans. All contiguous voxels in the parahippocampal region that exceed a predefined statistical criterion for responding significantly more strongly to scenes than to faces and objects in these localizer scans are included. We then measured the timecourse of MR activity in these voxels while subjects viewed scrambled and intact photographs of scenes, faces, houses, and objects in a second set of scans acquired within the same scan session (see Figure 2 for examples of stimuli for this and subsequent experiments). Subjects either viewed the photographs passively (half of the

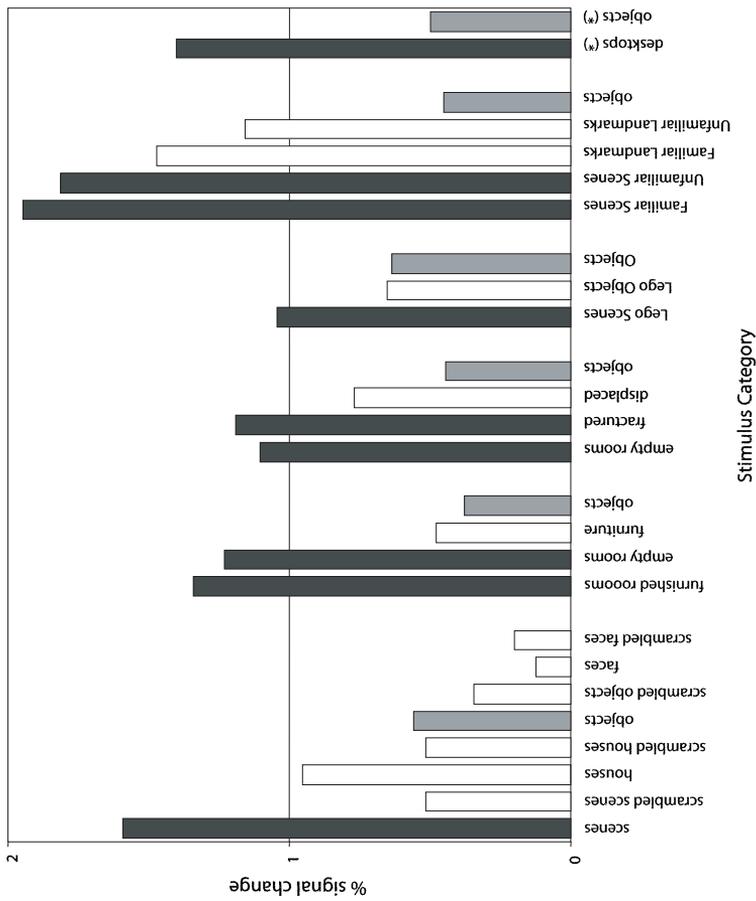


**Figure 2.** Examples of stimuli used in the experiments. (Grid lines were only overlaid on stimuli in the first experiment.)

scans) or performed a “1-back” repetition detection task in which they pressed a button whenever they saw two identical photographs in a row.

The average percentage signal change (from a fixation baseline) for each of the eight stimulus conditions in this experiment is shown in Figure 3. (Results for subsequent experiments are also plotted in this figure.) The PPA responded more strongly to scenes than to faces, objects, or houses. This preferential response to scenes held even after the response to the scrambled versions of the photographs was subtracted out (i.e., the scene vs. house  $\times$  intact vs. scrambled interaction was significant). Thus, the greater response to scenes than to other stimuli is unlikely to be attributable to low-level differences between the stimulus classes (e.g., texture density, average lightness), as many of these differences should be present in the scrambled images as well. The pattern of the PPA response did not differ between passive viewing and 1-back runs (data not shown). Although the mean PPA response to intact buildings (0.95%) was larger than the mean response to intact objects (0.59%), the scene vs. house  $\times$  intact vs. scrambled interaction was not significant.

This experiment demonstrated two novel and important facts. First, there is a swath of territory in ventromedial occipitotemporal cortex (the PPA) that responds preferentially to visual scenes. A number of other research groups have subsequently confirmed this basic result (e.g., Kohler, Crane, & Milner, 2002;



**Figure 3.** Response of the PPA to different stimulus classes in several experiments. Values are percentage signal change relative to a fixation baseline. Stimulus classes appearing within the same experiment are grouped together. Across experiments, the PPA response to scene-like stimuli (black bars) was significantly higher than the response to object-like stimuli (grey bars). Buildings gave an intermediate response. (\*) indicates stimuli presented in colour.

Maguire, Frith, & Cipolotti, 2001). Second, this region responds automatically to scenes, even when subjects are not actively involved in navigation. Earlier experiments had demonstrated increased parahippocampal response when navigation tasks were compared to less-navigationally demanding control tasks (Aguirre, Detre, Alsup, & D'Esposito, 1996; Ghaem et al., 1997; Maguire, Frackowiak, & Frith, 1997). For example, Aguirre et al. (1996) found that parahippocampal activity was greater when subjects actively navigated through a virtual reality maze than when they simply moved through an endlessly looping virtual reality corridor. Neuropsychological evidence (Aguirre & D'Esposito, 1999, see discussion below; Habib & Sirigu, 1987) also implicated parahippocampal cortex in navigation-related processing. The current results allowed us to go beyond these earlier results by assigning a particular role to parahippocampal cortex in navigation: Processing information about the local scene.

### THE PPA PROCESSES INFORMATION ABOUT THE SPATIAL STRUCTURE OF THE LOCAL ENVIRONMENT

What aspect of scenes drives the strong PPA response to them? There are many differences between scenes, objects, and faces that could potentially explain this pattern. For example, the PPA might be responding to the greater visual and/or semantic complexity of the scenes, or it might be computing the relative locations of the various objects in the scenes. Alternatively, the PPA might be responding to the background elements of the scene, which define the geometry of the surrounding environment. In this section, I will present evidence that this last explanation is correct: The PPA processes information about the shape of local space. I will refer to this as the *spatial layout hypothesis*.

PPA response is strongly affected by the presence of background elements in a scene, but not by the presence or absence of discrete objects (Epstein & Kanwisher, 1998, Exp. 2). We demonstrated this by scanning subjects while they viewed grayscale photographs of (1) furnished rooms, (2) empty rooms, which were the same rooms from the same viewpoint but empty of all furniture, and (3) furniture arrays, which consisted of the furniture and objects from the rooms on a white background (Figure 2). Except for the use of different stimuli, the procedure for this experiment was identical to the one described above. The task in this experiment (and in all of the experiments described in this section and the next) was passive viewing in half of the scans and 1-back repetition detection in the other half; data from both tasks were averaged together. As before, the PPA was defined for each subject as the set of contiguous voxels in the parahippocampal region that responded significantly more strongly to scenes than to faces and objects in a separate set of localizer scans.

Strikingly, the PPA responded much more strongly to the furnished and empty rooms than to the object arrays (see Figure 3). Furthermore, the response to the empty rooms was not significantly different from the response to the furnished rooms. Thus, the response of the PPA depended critically on the presence of background structures such as walls and floors that define the geometry of the three-dimensional space, but not on the presence or absence of individual objects within that space. This result provides strong support for the hypothesis that the PPA processes information about the spatial structure of the local environment.

In addition to furnished and empty rooms, subjects also viewed photographs of a college campus and natural landscapes in this experiment. The PPA response to these outdoor scenes was comparable to the response to the rooms (Epstein & Kanwisher, 1998). Thus, the PPA responds generally to all kinds of scenes, irrespective of whether they depict an indoor or outdoor location, or a natural or manmade environment. Subsequent experiments have shown that the PPA also responds strongly to desktop scenes (Epstein, Graham, & Downing, 2003; see Figures 2 and 3), scenes made out of Lego blocks (Epstein et al., 1999; see Figures 2 and 3), and even fairly minimal scenes created by rendering software (Downing, Epstein, & Kanwisher, unpublished data). This pattern of results is consistent with the spatial layout hypothesis: Despite their visual dissimilarity, all of these types of scenes convey spatial information. Importantly, this pattern is difficult to explain in terms of Haxby's object form topography model (Haxby, Gobbini, Furey, Ishai, Schouten, & Pietrini, 2001), which posits that regions of ventral temporal cortex respond preferentially to stimuli that share the same visual features. It is unclear what features are shared by the widely varying stimuli that activate the PPA, except for the "feature" of conveying information about the geometry of local space (see Gauthier, 2000 for discussion).

A further experiment provided additional support for the spatial layout hypothesis. Five subjects viewed images of empty rooms that were "fractured" into their component surfaces (Epstein & Kanwisher, 1998, Exp. 3). In one condition (fractured rooms), the relative positions of the resulting surfaces were preserved. In another condition (fractured + rearranged), the surfaces were rearranged so that they no longer defined a coherent space. Subjects also viewed the original unfractured images of the rooms (intact rooms). Percentage signal change in the PPA was significantly lower in the fractured + rearranged condition (0.8%) than in the fractured condition (1.2%). Insofar as the stimuli in both conditions contained the exact same image components, the difference in response must be attributed to the fact that the surfaces compose a coherent three-dimensional space in the fractured condition but not in the fractured + rearranged condition.

The idea that there might be a special-purpose mechanism for processing the geometric structure of scenes as defined by their background elements is

consistent with at least one line of behavioural results. Cheng (1986) and Margules and Gallistel (1988) explored the cues that rats used to reorient themselves after disorientation. They trained rats to find food in one corner of a rectangular box and then removed them from the box and slowly spun them around. When returned to the box, the rats searched for food with equal probability in the correct corner and in the geometrically equivalent diagonally opposite corner, indicating that the rats knew where the food was relative to the shape of the room. From these findings, Cheng and Margules and Gallistel hypothesized that an informationally encapsulated “geometric module” in the rat’s brain uses information about the overall shape of the environment to orient the animal. Evidence for the primacy of geometric cues has also been reported for rhesus monkeys (Gouteux, Thinus-Blanc, & Vauclair, 2001) and human infants (Hermer & Spelke, 1994, 1996), although the efficacy of nongeometric cues during reorientation remains controversial in the human case (Learmonth, Newcombe, & Huttenlocher, 2001). Although it would be too simplistic to suggest that the PPA is the anatomical locus of the geometric module, it would be surprising if the representation of spatial layout in the PPA did not play an important role in this geometry-based reorientation process.

### THE PPA RESPONSE DOES NOT DEPEND ON THE “SEMANTIC” QUALITIES OF THE SCENE

The above experiments demonstrate that the PPA responds strongly to visual scenes and that the critical factor driving this response is the presence in the stimulus of information about the spatial structure of the surrounding environment. However, the scene stimuli used in these experiments were mostly photographs of real places in the world. Thus, the PPA response to scenes could reflect either (1) the operation of mechanisms directly involved in analysing the spatial structure of the scene (irrespective of its “semantic” qualities, such as whether or not it depicts a familiar place), or (2) the operation of higher level navigational or recognition mechanisms that relate this spatial structure (which might be computed elsewhere in the brain) to stored representations of different places in the world. Note that these two possibilities are by no means mutually exclusive. However, in experiments designed to test whether the PPA was sensitive to the “semantic” (i.e., nonperceptual) qualities of the stimulus, we failed to observe any such sensitivity.

First, we demonstrated that the level of PPA response is not affected by real-world familiarity with the environment depicted in the scene (Epstein, Harris, Stanley, & Kanwisher, 1999). Four MIT students and four Tufts University students were scanned while viewing scenes and buildings from the MIT and Tufts campus. The MIT students had never visited Tufts, and the Tufts students had never visited MIT. All subjects reported that they recognized most or all of the scenes and buildings from their own campus and none from the other

campus. We predicted that the PPA should respond more strongly to the familiar than to the unfamiliar places if it were involved in retrieval of information about either (1) the real-world location of the place depicted in the photograph, or (2) the meaning, identity, or significance of the place. Strikingly, no such pattern was observed. The PPA responded equally strongly to photographs of familiar in unfamiliar places, even though semantic and topographical information about the depicted place could only be retrieved in the familiar condition (Epstein et al., 1999, Exp. 1). This result suggests that the PPA supports processes that operate solely on the spatial information that is locally available within the scene. A further experiment demonstrated that the PPA responded strongly to scenes even when they depict layouts made out of Lego blocks, which clearly did not depict places in the world and thus were unlikely to engage place recognition mechanisms (Epstein et al., 1999, Exp. 2). Note that these results do not speak to the issue of whether scenes from familiar and unfamiliar environments are represented in the same manner at the *neural* level; they simply demonstrate that both kinds of scenes activate the PPA to the same extent.

In contrast to this lack of a familiarity effect for scenes, the PPA *did* respond more strongly to photographs of familiar buildings than to photographs of unfamiliar buildings when the images were abstracted from their surroundings and placed on a white background. This pattern may have to do with the ambiguous nature of buildings, which are neither completely “objects” nor completely “scenes”. Several experiments have shown that the response to buildings is intermediate between the response to scenes and the response to other objects. Indeed, the PPA is often described (not quite accurately) as a “building” area (Aguirre, Zarahn, & D’Esposito, 1998; Haxby et al., 2001; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999). Importantly, this preference for buildings over household objects such as tools, shoes, and blenders is not due to the larger real-world size of the buildings: The PPA responds no more strongly to large movable objects such as ships and airplanes than to smaller objects (Epstein & Kanwisher, unpublished data). The key factor driving the strong PPA response to buildings may be the fact that buildings are stable objects that usually play an important role in defining the spatial structure of the environment. In essence, a photograph of a building is simultaneously a photograph of a discrete object and a photograph of a partial scene. Familiarity with a building might bias the brain to process it more as a partial scene (i.e., as a stimulus that tells you something about the geometry of surrounding space) than as a potentially movable object. This hypothesis is consistent with behavioural literature from rats indicating that these animals only use objects as landmarks if they have been shown by experience to be stable over time (Biegler & Morris, 1993).

Not only is the overall level of PPA response unaffected by familiarity with the place depicted in the scene, it is also unaffected by the navigational relevance of the scene. When subjects viewed colour photographs of outdoor scenes

and tabletop layouts, the response to the tabletop scenes was just as strong as the response to the outdoor scenes (Epstein et al., 2003; see Figure 3). This was despite the fact that the photographs of tabletop scenes were all taken from quite close up so that they did not provide any information about the surrounding room. This result is consistent with the idea that the PPA supports an analysis of the spatial structure of visual scenes as defined primarily by background elements rather than an analysis of the semantic qualities of the scene. Note that these results do not contradict the general notion that the PPA is involved in navigation: The PPA may support a spatial processing mechanism whose primary purpose is representation of scenes for navigation but which engages automatically to all scenes, including desktop scenes.

Interestingly, while the magnitude of the PPA response was not modulated by extraexperimental experience with the *place* depicted in the scene, it was modulated by within-experimental experience with particular scene photographs. In particular, response to repeated scene photographs was reduced compared to response to novel scene photographs (Epstein et al., 1999, Exp. 4). This result is particularly illuminating when considering the functional relationship between the PPA and the hippocampus. In the rat hippocampus, cells are found that represent the animal's position relative to other locations in the world. Although these cells are partially controlled by information about the local visual scene (see below), they do not directly represent the scene. In contrast, the PPA appears to represent the spatial structure of the currently visible scene, but does not represent where that scene is in the world. Consequently, the magnitude of the PPA response is affected by familiarity with particular scene images, but not by familiarity with the environment in which those scenes are embedded. Clearly, these two different kinds of spatial representations are complementary, and both are necessary for successful navigation.

### THE PPA PLAYS A CRITICAL ROLE IN THE LEARNING OF NEW SCENES

Damage to parahippocampal cortex is associated with deficits in spatial navigation and scene processing. In particular, patients with damage to this region can have trouble recognizing scenes, landmarks, and places (Habib & Sirigu, 1987; Landis, Cummings, Benson, & Palmer, 1986; Whiteley & Warrington, 1978), and occasionally show a more general impairment in the learning of new topographical information (Aguirre & D'Esposito, 1999). They are also impaired on an analogue to the Morris water maze task in which they have to remember the location of a hidden platform within an unfamiliar room (Bohbot, Kalina, Stepankova, Spackova, Petrides, & Nadel, 1998) as well as on a spatial memory task in which they have to remember the location of a target on the screen (Ploner et al., 2000). In a study inspired by the fMRI results described above, my colleagues and I tested two such patients and found that they were

selectively impaired at encoding new scenes but not new objects (Epstein, DeYoe, Press, Rosen, & Kanwisher, 2001).

The two patients suffered strokes that damaged the parahippocampal territory where the PPA is usually found (one bilaterally, one in the right hemisphere). Both patients reported an almost complete inability to learn new topographical information but only a mild general memory impairment. Interestingly, both patients were able to recall topographical information learned prior to their injury; for example, they were able to draw detailed maps of prior residences that contrasted favourably to their less detailed and less accurate maps of more recently encountered environments. Based on the neuroimaging data, we predicted that both patients should have a scene-processing deficit. A key question was whether such a deficit would manifest itself in perceptual tasks or memory tasks. Comparison of the results from two very similar recognition memory tasks suggested that mnemonic processing was particularly impacted:

In the *different-views matching task*, subjects viewed two photographs of layouts made out of Lego blocks sequentially presented with a 10 s interstimulus interval and reported whether they depicted the same layout from different viewing angles or different layouts. Previous fMRI studies demonstrated that the PPA responds strongly to these stimuli but not to “objects” made out of the same materials (Epstein et al., 1999; see Figure 2). In order to perform this task successfully, subjects must form a representation of the Lego layout, hold it in short-term memory for 10 s, and match it to a second Lego layout that is not visually identical. The patients were unimpaired at this task compared to age-matched normal controls. Furthermore, they did no better on the task when stimuli were photographs of Lego objects than when they were photographs of Lego layouts.

In the *continuous visual memory task*, subjects viewed a continuous sequence of photographs and reported whether or not they had seen each one before in the experiment. When the photographs repeated themselves, they did so either immediately (1-back), with two intervening items (3-back), or with four intervening items (5-back). In different versions of the test, stimuli were either photographs of Lego layouts or Lego objects. The patients were relatively impaired for the Lego layouts but not for the Lego objects, but only at the longest interval (i.e., in the 5-back but not the 1-back case).

Taken together, these experiments demonstrated that: (1) PPA damage results in a specific deficit for processing scenes but not objects, as predicted from the fMRI data; and (2) this deficit is found in recognition tasks only when the scene must be encoded into memory and retrieved after the presentation of a number of intervening items. Thus, these experiments suggest that the PPA may play a critical role in encoding information about the geometry of local space into memory.

A recent report by Mendez and Cherrier (2003) has confirmed and extended these results. They examined a patient who suffered damage to the right lingual/

parahippocampal region after a posterior cerebral artery stroke. This patient could describe routes without difficulty and could recognize landmarks such as buildings. However, he was severely impaired at recognizing scenes that did not contain distinguishing landmarks. This dissociation of recognition performance between landmarks and scenes was observed both in familiar environments (such as his home neighbourhood) and in novel environments. This report provides convincing evidence that one can be impaired specifically at the encoding and recognition of scenes without being impaired at the encoding and recognition of nonscenic sources of topographical information.

### THE PPA ENCODES SCENES IN A VIEWPOINT-SPECIFIC MANNER

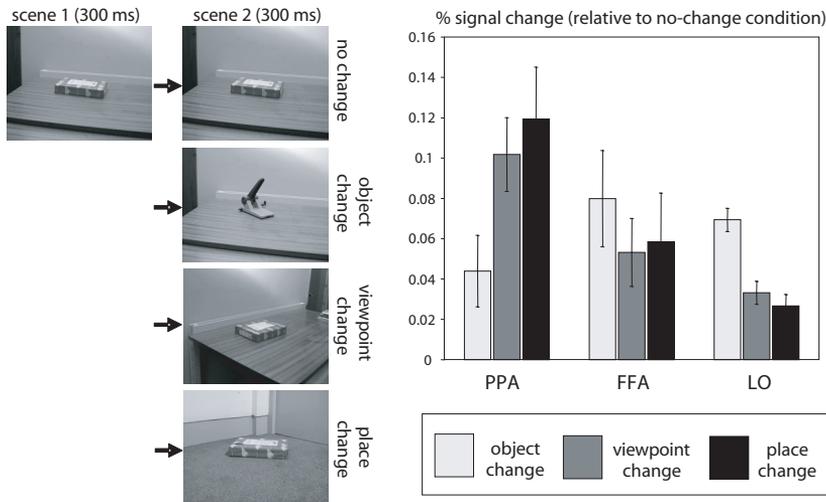
The above experiments strongly suggest that the PPA mediates a representation of the spatial structure of the current visual scene, but they do not reveal the nature of this representation. For example, they do not indicate whether scene representations in the PPA are viewpoint specific or viewpoint invariant. In the viewpoint-specific case, the PPA might represent the spatial relationship between the observer and the surfaces in the scene—a relationship that is different for different viewpoints. In the viewpoint-invariant case, the PPA might represent the intrinsic geometric arrangement of the surfaces without reference to the location or orientation of the observer. Either account could explain the strong PPA response to scenes in the previous experiments. For example, when subjects view many different photographs of empty rooms within a stimulus block, the PPA might be calculating the shape of each room, or it might be calculating the implied location of the observer relative to the walls of each room.

In order to distinguish between these two possibilities, we must have some way of determining whether the PPA considers two views of the same scene to be representationally identical (i.e., the viewpoint-invariant case) or representationally distinct (i.e., the viewpoint-specific case). In the viewpoint-invariant case, different views of the same place would activate the same subset of neurons in the PPA, while in the viewpoint-specific case they would activate different neural subsets. Although the spatial resolution of fMRI is far too coarse for direct examination of neural firing patterns, information about these patterns can be obtained indirectly by looking at neural adaptation effects (Grill-Spector & Malach, 2001). When two identical stimuli are presented sequentially, neurophysiological experiments have demonstrated that response to the second item is reduced compared to the response to the first item (Baylis & Rolls, 1987; Brown, Wilson, & Riches, 1987; Brown & Xiang, 1998; Li, Miller, & Desimone, 1993; Miller, Li, & Desimone, 1991; Muller, Metha, Krauskopf, & Lennie, 1999; Ringo, 1996). This reduction in neural response can be observed using fMRI (Buckner et al., 1998; Grill-Spector, Kushnir, Edelman, Avidan,

Itzhak, & Malach, 1999; Grill-Spector, Kushnir, Edelman, Itzhak, & Malach, 1998; Henson, Shallice, & Dolan, 2000; Kourtzi & Kanwisher, 2001; Schacter & Buckner, 1998; Stern et al., 1996; van Turennout, Bielarowicz, & Martin, 2003; Vuilleumier, Henson, Driver, & Dolan, 2002). For example, if two scenes are presented sequentially within a 1–2 s long “event”, the overall fMRI response to the event will be smaller if the scenes are representationally identical than if they are representationally distinct, because the two stimuli will engage the same set of neurons in the “same” case (leading to neural adaptation) but a nonoverlapping set of neurons in the “different” case. Importantly, this effect should also be observed when the stimuli in the “same” case are not physically identical but are “the same” in terms of the kind of information represented by a cortical region (Grill-Spector & Malach, 2001; Kourtzi & Kanwisher, 2001).

My colleagues and I (Epstein et al., 2003) used this logic to investigate the viewpoint-specific vs. viewpoint-invariant question in the PPA. We scanned subjects while they viewed events consisting of two successively presented photographs of tabletop layouts presented for 300 ms each with a 100 ms interstimulus interval. The two photographs could have one of four possible relationships to each other: (1) They could be completely identical (no-change condition); (2) they could depict the same layout from the same viewing angle, but differ in the identity of a prominent central object (object-change condition); (3) they could depict the same layout with the same central object, but shown from different viewpoints (viewpoint-change condition); or (4) they could show the same central object placed in different spatial contexts (place-change condition) (Figure 4). We predicted that the PPA would respond more strongly to a change in scene layout (i.e., to the place-change condition) than to a change in viewpoint (i.e., to the viewpoint-change condition) if it represented scenes in a viewpoint-invariant manner, but would respond equally to both kinds of changes if it represented scenes in a viewpoint-specific manner.

The results indicated viewpoint-specific scene representations: PPA response to viewpoint- and place-changes was almost exactly the same. In other words, as far as the PPA is concerned, two photographs of the same scene taken from different viewpoints were as representationally distinct as two photographs of different scenes. In contrast, neural adaptation led to significant reduction of response when viewpoint was repeated in the object-change and no-change conditions—even though the visual change was quite large in the object-change case. The finding that scene representations in the PPA are encoded in a viewpoint-specific manner is consistent with a number of behavioural studies showing viewpoint-specific memory and priming effects for scenes (Christou & Bulthoff, 1999; Chua & Chun, 2003; Diwadkar & McNamara, 1997; Mou & McNamara, 2002; Nakatani, Pollatsek, & Johnson, 2002; Shelton & McNamara, 2001; but see Sanocki & Epstein, 1997). Interestingly, subsequent experiments suggested that the complete viewpoint specificity observed in this experiment might relate to the unfamiliarity of the stimuli and that some viewpoint



**Figure 4.** Scene representations in the PPA are encoded in a viewpoint-specific manner. Left: Examples of stimuli. PPA activity was measured in response to “events” in which object, viewpoint, or place information could change. Right: PPA response to viewpoint and place changes was equivalent, and greater than response to object changes. In contrast, the lateral occipital complex (LO) responded more strongly to object changes, and the fusiform face area did not distinguish between the conditions.

invariance can develop as scenes become more familiar (Epstein, Higgins, & Thompson-Schill, 2005—see next section for discussion).

In contrast to the greater PPA response to spatial changes, the “lateral occipital complex” (LO; Malach et al. 1995) responded significantly more strongly to object changes than to viewpoint or place changes, consistent with previous work implicating this region in the perception of object shape (Grill-Spector, Kourtzi, & Kanwisher, 2001; Kanwisher, Woods, Iacoboni, & Mazziotta, 1997b; Kourtzi & Kanwisher, 2001; Malach et al., 1995). Thus, we observed a double dissociation between the PPA and LO: The former responds more strongly to spatial changes, while the latter responds more strongly to object changes. These results suggest a scheme by which ventral occipito-temporal cortex subserves at least two anatomically and functionally distinct processing mechanisms: One that processes information about the spatial relationship between the observer and the background elements of the scene, and one that processes information about the objects within the scene. Consistent with this claim, Levy, Hasson, Avidan, Hendler, and Malach (2001) have noted the PPA is well-positioned to receive inputs from early visual areas that process information about the periphery of the visual field. In contrast, object processing areas such as the fusiform face area (FFA; Kanwisher, McDermott, & Chun,

1997a) are well-positioned to receive inputs from the centre of the visual field. Levy et al. (2001; see also Epstein et al., 2003) also provided evidence that the PPA responds to some extent even to nonscenic information in the periphery, although this effect is secondary to the main “categorical” effect.

## THE QUALITY OF PPA SCENE REPRESENTATIONS CORRELATES WITH NAVIGATIONAL ABILITY

As noted above, a number of neuroimaging studies have obtained parahippocampal activation when subjects perform navigational tasks (Aguirre et al., 1996; Ghaem et al., 1997; Maguire et al., 1997). The claim that the PPA is involved in processing information about the spatial structure of the local scene is consistent with these earlier results. In a recent study, my colleagues and I (Epstein et al., 2005) used fMRI adaptation to establish a more direct link between PPA scene representations and navigation. Specifically, we showed that the strength of fMRI adaptation effects in the PPA correlates with self-reported navigational ability, as assessed by Hegarty’s Santa Barbara Sense of Direction scale (SBSOD; Hegarty, Richardson, Montello, Lovelace, & Ilavani, 2002). These results suggest that one component of successful navigation is the encoding of accurate representations of the local scene.

The experiment consisted of two parts, conducted within the same scan session. The first part was similar to the previous experiment: Subjects viewed pairs of successively presented photographs of indoor and outdoor scenes, which could either be identical (no-change condition), different views of the same place (viewpoint-change condition), or photographs of two different places (place-change). In the second part, subjects viewed single scene photographs and made indoor/outdoor judgements. These photographs could be 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 the first part of the experiment allowed us to examine short-term fMRI adaptation effects (operating over the course of hundreds of milliseconds), while pairwise comparison of the three conditions in the second part of the experiment allowed us to examine long-term fMRI adaptation effects (operating over the course of minutes).

The first part of the experiment replicated our earlier results implicating the PPA in viewpoint-specific scene processing: The PPA responded equivalently to viewpoint and place changes but response in the no-change condition was reduced. The second part of the experiment revealed both viewpoint-specific (new view > old view) and viewpoint-invariant (new place > new view) long-term adaptation effects in the PPA. Interestingly, these effects were significantly correlated with navigational ability as asses-

sed by the SBSOD score. For example, SBSOD scores accounted for approximately 44% of the variance in the magnitude of the viewpoint-specific effect in the right PPA. Both viewpoint-specific and viewpoint-invariant adaptation effects were larger in good navigators than in bad navigators (despite the fact that the overall level of response to scenes did not vary as a function of navigational ability). These results suggest that navigational competence relates directly to the quality of scene representations in the PPA. Specifically, good navigators encode representations that distinguish between new and old views of a previously seen place, while bad navigators do not.

Beyond the correlation with navigational ability, the finding that long-term adaptation in the PPA was partially viewpoint invariant is intriguing because it suggests that scene representations may become less viewpoint specific as one gains experience with different views of a place. A similar learning process has been postulated for nonscene objects, but to my knowledge no evidence for the formation of viewpoint-invariant object representations over the course of an experimental session has been observed. The above results suggest that the PPA might be the locus of a mechanism for learning places from individual views; however, this hypothesis must be tested through further experiments. For example, it would be interesting to see if a greater degree of viewpoint invariance is found for photographs of familiar environments than was found for the photographs of unfamiliar environments used in the present experiment.

### THE PPA MAY BE PART OF A LARGER NETWORK OF SPATIAL-PROCESSING REGIONS

The PPA is not the only region of the brain that processes navigationally relevant spatial information. *Place cells* in the rat hippocampus fire when the animal is in a particular location but are quiescent otherwise (O'Keefe & Dostrovsky, 1971). Furthermore, these cells often exhibit considerably different place fields in different environments, indicating that the rat hippocampus not only represents the location of the animal within the local environment but also distinguishes between different local environments (Best, White, & Minai, 2001). Results such as these led O'Keefe and Nadel (1978) to postulate that the hippocampus supports a cognitive map (Tolman, 1948) that represents the spatial structure of the environment in allocentric (i.e., world-centred) coordinates. Supporting this contention is the fact that rats with hippocampal lesions are severely impaired on tasks that require recall of the relationship between different world locations, such as swimming from a random starting position to a hidden fixed platform in a water-filled pool (Morris, Garrud, Rawlins, & O'Keefe, 1982).

Some recent studies have identified place cells in the primate hippocampus (Matsumura, Nishijo, Tamura, Eifuku, Endo, & Ono, 1999), while others have

identified *spatial view cells* in the same region that fire when the animal looks at particular locations in the world (Rolls, Robertson, & Georges-Francois, 1997). Place cells have also been identified in the human hippocampus using intracranial recordings (Ekstrom et al., 2003). Spatial view cells and place cells are similar in two critical respects. First, they both respond to locations in an allocentric manner: Their firing does not depend on the bodily orientation, eye position, or head position of the animal (Georges-Francois, Rolls, & Robertson, 1999). Second, they respond to locations *per se* rather than particular visual cues experienced in those locations. For example, place cells fire in the dark and retain their firing pattern with some degree of accuracy as the animal moves through an unlit environment, while spatial view firing fields are unchanged when the walls of the experimental chamber are obscured by curtains (Robertson, Rolls, & Georges-Francois, 1998). Thus, both the rat and monkey hippocampus appear to be involved in supporting an amodal (i.e., nonvisual) representation of the locations of “places” in a world-centred coordinate frame. Interestingly, a separate set of neurons in the postsubiculum, anterior thalamus, mammillary bodies, and retrosplenial cortex appears to be involved in representing the orientation of the animal in the world (Taube, 1998). These *head direction (HD) cells* are similar to place and spatial view cells in that they represent orientation in allocentric coordinates (i.e., they represent orientation relative to a world-centred coordinate frame, rather than bearings to specific landmarks) and can be updated using both visual and nonvisual cues (Stackman & Taube, 1997).

The parietal lobes are also critically involved in processing spatial information. Neurophysiological data indicate that parietal regions represent the locations of behaviourally relevant targets in a number of different coordinate frames tied to effector and receptor surfaces such as the eye, hand, or mouth (Andersen, Snyder, Bradley, & Xing, 1997; Colby & Goldberg, 1999). For example, neurons in the medial intraparietal region (MIP) respond when the animal reaches for a target and show selectivity both for the retinotopic location of the target and for the reaching hand, while neurons in the lateral intraparietal region (LIP) and area 7a encode the locations of eye-movement targets in retinotopic coordinates (Colby, Duhamel, & Goldberg, 1996; Gottlieb, Kusunoki, & Goldberg, 1998). Although parietal involvement in mediating movements of the whole body through space has been comparatively less studied in monkeys, human neuropsychological and neuroimaging data implicate several parietal subregions in active navigation. For example, medial parietal regions are active during scene viewing and both medial and inferior parietal regions are active during navigational tasks (Burgess, Maguire, Spiers, & O’Keefe, 2001; Maguire, Burgess, Donnett, Frackowiak, Frith, & O’Keefe, 1998). Patients with damage to parietal regions can show a severe deficit in navigating through the immediate environment, often bumping into walls and other obstacles (Stark, Coslett, & Saffran, 1996). These data suggest that parietal regions play a critical

*online* role in guiding locomotion—for example, by tracking the locations of navigationally relevant surfaces as one moves. Importantly, this kind of information is constantly changing, and thus parietal representations may be too transient to support navigational planning (Milner & Goodale, 1995). However, they may be important as inputs to more durable spatial representations supported by other regions.

How is the PPA connected to these other spatial processing regions? In the monkey, there are strong anatomical projections from parietal regions LIP and 7a to parahippocampal cortex region TF (Suzuki & Amaral, 1994) and indirect projections from parietal cortex to parahippocampal regions TF and TH via retrosplenial cortex (Morris, Paxinos, & Petrides, 2000; Morris, Petrides, & Pandya, 1999; Suzuki, 1996). Parahippocampal cortex also receives input from ventral stream visual areas TE and TEO, putting it in an ideal position to combine parietal information about the locations of different surfaces with ventral stream information about the overall “shape” of the surrounding environment. Parahippocampal cortex projects to entorhinal cortex, which is the primary source of input to the hippocampus. Thus, the anatomical data suggest that a pathway from parietal cortex to parahippocampal cortex to the hippocampus may be critical for processing navigationally relevant spatial information. Supporting this contention, Burgess et al. (2001) observed increased activation in all of these regions when recall of the spatial aspects of an event (“Where did this event take place?”) was contrasted with recall of the non-spatial aspects (e.g., “Who was involved?”).

Indeed, functional studies suggest that these regions form a processing hierarchy in humans. Maguire et al. (1997) found both hippocampal and parahippocampal activation when subjects retrieved routes through complex environments, but only parahippocampal activation when they recalled individual landmarks or scenes. Furthermore, King, Burgess, Hartley, Vargha-Khadem, and O’Keefe (2002) found that a patient with bilateral hippocampal atrophy was strongly impaired at recalling the locations of objects within a virtual-reality courtyard when he was tested using a viewpoint different from the one used at encoding, but not when he was tested using the same viewpoint used at encoding. Taken together, these results indicate that the PPA supports viewpoint-specific scene representations while the hippocampus may be critical for encoding viewpoint-independent scene representations and understanding the routes between different places. This formulation fits both the cognitive map theory, and also the more general idea that the hippocampus is critical for encoding sequences of locations or events (Eichenbaum, Dudchenko, Wood, Shapiro, & Tanila, 1999; Fortin, Agster, & Eichenbaum, 2002; Levy, 1996).

Neurophysiological data from animals also support this scheme. Neither place cells nor HD cells represent visual cues (as evidenced by their firing in darkness), but both place fields and HD fields are strongly controlled by certain kinds of visual cues when they are present. In particular, the relationship

between the animal and the overall configuration of distant visible landmarks (Knierim, Kudrimoti, & McNaughton, 1995; Muller & Kubie, 1987; O'Keefe & Speakman, 1987) or barriers (O'Keefe & Burgess, 1996) can be a critical determinant of place cell response. Similar results have been obtained for HD cells—indeed, place and HD fields usually move in a tightly coupled manner when landmarks locations change (Knierim, Kudrimoti, & McNaughton, 1998). These results suggest that both place and HD cells receive input from a common representation of the geometry of local scene (Golob & Taube, 1997). Although the locus of such a representation has not been identified in animals, the human data suggest that parahippocampal cortex may be a source of such geometric information.

### UNRESOLVED ISSUES AND CONCLUSION

The evidence discussed above strongly suggests that parahippocampal cortex plays a critical role in processing information about the spatial structure of scenes as defined primarily by their background elements. Thus, we have identified the neural basis for processing information about one major aspect of visual scenes. What about the other aspects? Can we say anything, for example, about the cortical regions involved in place recognition, or in processing gist? Note that both of these operations involve linking visual to nonvisual information. In place recognition, the aim is to determine the *identity* of the place depicted in the scene, both in conceptual terms (“What is the name of this particular place?”) and spatial terms (“Where is this place relative to other places in the world?”). In gist recognition, the aim is to determine the *category* of the scene (“What *kind* of scene is this?”).

Place recognition appears to involve multiple regions of the brain. Nakamura et al. (2000) used PET to measure the neural response to photographs of faces and places that were either personally familiar or unfamiliar to the subjects. In occipital-temporal regions, the well-established face/place division was observed: Both familiar and unfamiliar faces activated fusiform gyrus, while both familiar and unfamiliar places activated parahippocampal cortex and the parietal/occipital junction. In contrast, the temporal pole responded to familiar but not unfamiliar stimuli, irrespective of stimulus type. (Similar results were obtained by Gorno-Tempini and Price, 2001.) These results fit well with McCarthy, Evans, and Hodges' (1996) report of a patient with anterior temporal damage who is impaired not only at place recognition but also in his ability to recall semantic information about different places. Thus, place recognition (in the sense of being able to identify different places in the world) may require the linking of parahippocampal representations of scene geometry with long-term semantic memory stores in anterior temporal cortex. Interestingly, it is unclear whether long-term information about the actual *locations* of places is stored in anterior temporal cortex: McCarthy et al.'s patient could draw an accurate map

of his home town (even though he couldn't recognize locations in that town from photographs), while Teng and Squire (1999) reported that a patient with extensive anterior and medial temporal damage could retrieve information about the relative locations of landmarks in his former hometown. (In both patients, this location information was acquired before their injury.) Thus, at least for highly familiar environments, retrieving the relative locations of different places might involve different brain regions than those involved in retrieving the names and meanings of the places (see also Aguirre & D'Esposito, 1997).

As for gist processing, comparatively little is known about how this is implemented in the brain. Behavioural studies have demonstrated that we can interpret scenes ("picnic", "beach", "airplane flying over clouds") quite quickly (Potter, 1975)—almost as quickly as we can identify individual objects (Thorpe, Fize, & Marlot, 1996). The mechanisms behind this impressive ability are unclear. In many cases, "gist" might be identifiable from a few key objects or materials such as a cactus or camel in the desert. The visual system might operate in parallel on several different objects within a scene in order to classify the gist of the scene in approximately the same amount of time that it takes to classify the individual objects. In other cases, gist recognition—at least insofar as it involves classifying a scene into a preexisting category such as "beach" or "desert"—might rely on mechanisms that analyse the overall shape, texture, or colour distribution within a scene (Oliva & Schyns, 1997, 2000; Oliva & Torralba, 2001) and are independent of the mechanisms for object form analysis (Steeves, Humphrey, Culham, Menon, & Goodale, 2002).

Given the strong response of the PPA to scenes, one might expect that it plays an important role in processing scene gist. For example, the PPA might combine information about the overall spatial layout of the scene with a summary of either object or feature information to come up with a unified interpretation of the scene's meaning. To date, this hypothesis has not been tested. However, recent results from Bar and Aminoff (2003) are consistent with the idea that the PPA may be involved in rapidly summarizing object information within a scene. These researchers found that parahippocampal cortex responded more strongly to objects that were strongly associated with particular contexts than to objects that had no such association. Interestingly, the "contexts" in this experiment could either be particular kinds of places ("farm", "kitchen") or more abstract situations or ideas ("music", "birthday"). Although preliminary, these results point to a role for the PPA in processing scene gist that bears further investigation.

Finally, the fourth aspect of visual scenes is that they contain objects. The mechanisms involved in identifying objects and specifying their locations have been extensively studied. Indeed, as it is generally formulated, the standard division of the visual system into what vs. where pathways *assumes* that object recognition and localization are the two main goals of vision (Mishkin, Ungerleider, & Macko, 1983). However, just as actors cannot act without a

stage, objects cannot appear except within the context of a scene. Thus, one of salutary aspect of studying scene perception is that it expands our conception of what vision is for. Vision scientists have spent many years studying the actors; now it is time to direct some attention to the stage.

## REFERENCES

- Aguirre, G. K., & D'Esposito, M. (1997). Environmental knowledge is subserved by separable dorsal/ventral neural areas. *Journal of Neuroscience*, *17*, 2512–2518.
- 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.
- Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, *20*, 303–330.
- Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, *38*, 347–358.
- Baylis, G. C., & Rolls, E. T. (1987). Response of neurons in inferotemporal cortex in short-term and serial recognition memory tasks. *Experimental Brain Research*, *65*, 614–622.
- 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.
- Biegler, R., & Morris, R. G. M. (1993). Landmark stability is a prerequisite for spatial but not discrimination-learning. *Nature*, *361*, 631–633.
- Bohbot, V. D., Kalina, M., Stepankova, K., Spackova, N., Petrides, M., & Nadel, L. (1998). Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex. *Neuropsychologia*, *36*, 1217–1238.
- Brown, M. W., Wilson, F. A. W., & Riches, I. P. (1987). Neuronal evidence that inferomedial temporal cortex is more important than hippocampus in certain processes underlying recognition memory. *Brain Research*, *65*, 158–162.
- Brown, M. W., & Xiang, J. Z. (1998). Recognition memory: Neuronal substrates of the judgement of prior occurrence. *Progress in Neurobiology*, *55*, 149–189.
- Buckner, R. L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., et al. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron*, *20*, 285–296.
- Burgess, N., Maguire, E. A., Spiers, H. J., & O'Keefe, J. (2001). A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage*, *14*, 439–453.
- Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*, 149–178.
- Christou, C. G., & Bulthoff, H. H. (1999). View dependence in scene recognition after active learning. *Memory and Cognition*, *27*, 996–1007.
- Chua, K. P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception and Psychophysics*, *65*, 72–80.
- Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1996). Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *Journal of Neurophysiology*, *76*, 2841–2852.
- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, *22*, 319–349.

- Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science*, *8*, 302–307.
- Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., & Tanila, H. (1999). The hippocampus, memory, and place cells: Is it spatial memory or a memory space? *Neuron*, *23*, 209–226.
- Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., et al. (2003). Cellular networks underlying human spatial navigation. *Nature*, *425*, 184–188.
- 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. A., Higgins, J. S., & Thompson-Schill, S. L. (2005). Learning places from views: Variation in scene processing as a function of experience and navigational ability. *Journal of Cognitive Neuroscience*, *17*, 73–83.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, *5*, 458–462.
- Gauthier, I. (2000). What constrains the organization of the ventral temporal cortex? *Trends in Cognitive Sciences*, *4*, 1–2.
- Georges-Francois, P., Rolls, E. T., & Robertson, R. G. (1999). Spatial view cells in the primate hippocampus: Allocentric view not head direction or eye position or place. *Cerebral Cortex*, *9*, 197–212.
- Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., et al. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport*, *8*, 739–744.
- Golob, E. J., & Taube, J. S. (1997). Head direction cells and episodic spatial information in rats without a hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 7645–7650.
- Gorno-Tempini, M. L., & Price, C. J. (2001). Identification of famous faces and buildings: A functional neuroimaging study of semantically unique items. *Brain*, *124*, 2087–2097.
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*, 481–484.
- Gouteux, S., Thinus-Blanc, C., & Vauclair, J. (2001). Rhesus monkeys use geometric and nongeometric information during a reorientation task. *Journal of Experimental Psychology: General*, *130*, 505–519.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- 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., Kushnir, T., Edelman, S., Itzhak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, *21*, 191–202.
- 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.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.

- 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.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243–271.
- Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science, 287*, 1269–1272.
- Hermer, L., & Spelke, E. (1996). Modularity and development: The case of spatial reorientation. *Cognition, 61*, 195–232.
- Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370*, 57–59.
- 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 of the United States of America, 96*, 9379–9384.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997a). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.
- Kanwisher, N., Woods, R. P., Iacoboni, M., & Mazziotta, J. C. (1997b). A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience, 9*, 133–142.
- 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.
- Knierim, J. J., Kudrimoti, H. S., & McNaughton, B. L. (1995). Place cells, head direction cells, and the learning of landmark stability. *Journal of Neuroscience, 15*, 1648–1659.
- Knierim, J. J., Kudrimoti, H. S., & McNaughton, B. L. (1998). Interactions between idiothetic cues and external landmarks in the control of place cells and head direction cells. *Journal of Neurophysiology, 80*, 425–446.
- Kohler, S., Crane, J., & Milner, B. (2002). Differential contributions of the parahippocampal place area and the anterior hippocampus to human memory for scenes. *Hippocampus, 12*, 718–723.
- Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293*, 1506–1509.
- Landis, T., Cummings, J. L., Benson, D. F., & Palmer, E. P. (1986). Loss of topographic familiarity: An environmental agnosia. *Archives of Neurology, 43*, 132–136.
- Learmonth, A. E., Newcombe, N. S., & Huttenlocher, J. (2001). Toddlers' use of metric information and landmarks to reorient. *Journal of Experimental Child Psychology, 80*, 225–244.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4*, 533–539.
- Levy, W. B. (1996). A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus, 6*, 579–590.
- Li, L., Miller, E. K., & Desimone, R. (1993). The representation of stimulus-familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69*, 1918–1929.
- 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.
- Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around London: Activation of the right hippocampus in taxi drivers. *Journal of Neuroscience, 17*, 7103–7110.
- Maguire, E. A., Frith, C. D., & Cipolotti, L. (2001). Distinct neural systems for the encoding and recognition of topography and faces. *Neuroimage, 13*, 743–750.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic-resonance-imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 8135–8139.
- Margules, J., & Gallistel, C. R. (1988). Heading in the rat—Determination by environmental shape. *Animal Learning and Behavior, 16*, 404–410.
- Matsumura, N., Nishijo, H., Tamura, R., Eifuku, S., Endo, S., & Ono, T. (1999). Spatial- and task-dependent neuronal responses during real and virtual translocation in the monkey hippocampal formation. *Journal of Neuroscience, 19*, 2382–2393.

- McCarthy, R. A., Evans, J. J., & Hodges, J. R. (1996). Topographic amnesia: Spatial memory disorder, perceptual dysfunction, or category specific semantic memory impairment? *Journal of Neurology, Neurosurgery, and Psychiatry*, *60*, 318–325.
- Mendez, M. F., & Chierri, M. M. (2003). Agnosia for scenes in topographagnosia. *Neuropsychologia*, *41*, 1387–1395.
- Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, *254*, 1377–1379.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford, UK/New York: Oxford University Press.
- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision—2 cortical pathways. *Trends in Neurosciences*, *6*, 414–417.
- Morris, R., Paxinos, G., & Petrides, M. (2000). Architectonic analysis of the human retrosplenial cortex. *Journal of Comparative Neurology*, *421*, 14–28.
- Morris, R., Petrides, M., & Pandya, D. N. (1999). Architecture and connections of retrosplenial area 30 in the rhesus monkey (*Macaca mulatta*). *European Journal of Neuroscience*, *11*, 2506–2518.
- Morris, R. G., Garrud, P., Rawlins, J. N., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*, 681–683.
- Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 162–170.
- Muller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, *285*, 1405–1408.
- Muller, R. U., & Kubie, J. L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience*, *7*, 1951–1968.
- Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., et al. (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., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature*, *381*, 425–428.
- O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*, 171–175.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, UK: Clarendon Press/Oxford University Press.
- O'Keefe, J., & Speakman, A. (1987). Single unit activity in the rat hippocampus during a spatial memory task. *Experimental Brain Research*, *68*, 1–27.
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.
- Oliva, A., & Schyns, P. G. (2000). Diagnostic colours mediate scene recognition. *Cognitive Psychology*, *41*, 176–210.
- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*, 145–175.
- Ploner, C. J., Gaymard, B. M., Rivaud-Pechoux, S., Baulac, M., Clemenceau, S., Samson, S., et al. (2000). Lesions affecting the parahippocampal cortex yield spatial memory deficits in humans. *Cerebral Cortex*, *10*, 1211–1216.
- Potter, M. C. (1975). Meaning in visual search. *Science*, *187*, 965–966.
- Ringo, J. L. (1996). Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behavioural Brain Research*, *76*, 191–197.
- Robertson, R. G., Rolls, E. T., & Georges-Francois, P. (1998). Spatial view cells in the primate hippocampus: Effects of removal of view details. *Journal of Neurophysiology*, *79*, 1145–1156.
- Rolls, E. T., Robertson, R. G., & Georges-Francois, P. (1997). Spatial view cells in the primate hippocampus. *European Journal of Neuroscience*, *9*, 1789–1794.

- Sanocki, T., & Epstein, W. (1997). Priming spatial layout of scenes. *Psychological Science*, 8, 374–378.
- Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, 20, 185–195.
- Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, 43, 274–310.
- Stackman, R. W., & Taube, J. S. (1997). Firing properties of head direction cells in the rat anterior thalamic nucleus: Dependence on vestibular input. *Journal of Neuroscience*, 17, 4349–4358.
- Stark, M., Coslett, H. B., & Saffran, E. M. (1996). Impairment of an egocentric map of locations: Implications for perception and action. *Cognitive Neuropsychology*, 13, 481–523.
- Steeves, J. K. E., Humphrey, G. K., Culham, J. C., Menon, R. S., & Goodale, M. A. (2002). *Scene classification and parahippocampal place area activation in an individual with visual form agnosia*. Paper presented at the Vision Sciences Society, Sarasota, FL.
- Stern, C. E., Corkin, S., Gonzalez, R. G., Guimaraes, A. R., Baker, J. R., Jennings, P. J., et al. (1996). The hippocampal formation participates in novel picture encoding: Evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 8660–8665.
- Suzuki, W. A. (1996). Neuroanatomy of the monkey entorhinal, perirhinal and parahippocampal cortices: Organization of cortical inputs and interconnections with amygdala and striatum. *Seminars in the Neurosciences*, 8, 3–12.
- Suzuki, W. A., & Amaral, D. G. (1994). Perirhinal and parahippocampal cortices of the macaque monkey—Cortical afferents. *Journal of Comparative Neurology*, 350, 497–533.
- Taube, J. S. (1998). Head direction cells and the neurophysiological basis for a sense of direction. *Progress in Neurobiology*, 55, 225–256.
- Teng, E., & Squire, L. R. (1999). Memory for places learned long ago is intact after hippocampal damage. *Nature*, 400, 675–677.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208.
- Van Turennout, M., Bielarowicz, 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.
- Whiteley, A. M., & Warrington, E. K. (1978). Selective impairment of topographical memory: A single case study. *Journal of Neurology, Neurosurgery, and Psychiatry*, 41, 575–578.