



HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2016 June 08.

Published in final edited form as:

Nat Neurosci. ; 14(10): 1224–1226. doi:10.1038/nn.2936.

Is that a bathtub in your kitchen?

Marius V. Peelen and

Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

Sabine Kastner

Department of Psychology and at the Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, USA

Marius V. Peelen: marius.peelen@unitn.it; Sabine Kastner: skastner@princeton.edu

Abstract

We can efficiently and rapidly recognize daily-life visual settings. A study finds that scene recognition involves the posterior object-selective visual cortex, where multiple within-scene objects are represented in parallel.

At a brief glance, we perceive grazing animals and trees or we see cars and traffic lights. This object information helps to distinguish a pasture from a city street. Recognizing such daily-life visual scenes is a common task for the visual system and is one at which it is highly adept. Yet despite the apparent ease with which we categorize scenes, the neural mechanisms of real-world scene recognition are complex and largely unknown. In this issue, MacEvoy and Epstein¹ provide compelling neural evidence that the parallel analysis of multiple diagnostic objects, objects typically found in the context of a specific scene, may constitute a mechanism by which real-world scenes are efficiently recognized.

Although the neural basis of visual object perception has been studied for decades, much of this research has evaluated brain responses evoked by objects shown in isolation. As a result, little is known about how the brain processes objects that are present in meaningful, but typically cluttered, real-world scenes. In daily-life scenes, the visual appearance of an object (for example, a chair) may change radically depending on variations typical of the real world. For example, an object may be partly occluded (a chair behind a desk). We have a remarkable ability to quickly recognize objects across such naturalistic variations². Recent functional magnetic resonance imaging (fMRI) studies^{3,4} have provided evidence that object-selective regions in higher-level visual cortex may support this ability. These regions, defined by their strong response to intact as compared with scrambled objects, have been found to represent within-scene objects independently of real-world variations in their appearance and location^{3,4}.

However, for rapid scene recognition, the brain would need to represent multiple objects in parallel, as one object may not provide sufficient information about a given scene type. For example, a chair may be found both in the kitchen and in the living room. The study by

Competing Financial Interests: The authors declare no competing financial interests.

MacEvoy and Epstein¹ takes a substantial step toward a better understanding of real-world scene perception by investigating the neural processing of multiple simultaneously presented objects in scenes. They provide evidence that multiple within-scene objects are indeed processed in parallel and that this occurs in the posterior part of object-selective cortex. Such information about the presence of multiple objects in a scene can be used to support reliable inferences about scene category; for example, to distinguish a kitchen from a living room.

Using fMRI, MacEvoy and Epstein¹ recorded brain activity of participants while they viewed pictures of four different scene types (bathroom, kitchen, intersection and playground), each of which contained two diagnostic objects (for example, all bathrooms contained a bathtub and a toilet). Participants also viewed different exemplars of these object types individually and without scene background. By measuring the multi-voxel patterns of brain responses evoked by these isolated objects, the authors established a 'neural signature' of each object type. Previous research has shown that such multi-voxel activity patterns allow the discrimination of object categories⁵. Here the similarity between the isolated object-evoked responses and the scene-evoked responses was used as a measure of the degree to which within-scene objects were processed³.

Their results revealed that in one brain area in particular, the object-selective lateral occipital cortex, there was a close correspondence between response patterns evoked by within-scene objects and response patterns evoked by these objects when shown in isolation. For example, response patterns evoked by pictures of bathrooms were similar to response patterns evoked by isolated pictures of bathtubs, relative to other objects not present in the bathroom scene. These results indicate that lateral occipital cortex carries information about the objects in the scenes and that it processes within-scene objects similarly to how it processes these objects without scene context. Other brain regions did not carry such information. In particular, activity patterns in the parahippocampal place area (PPA), a region implicated in scene perception⁶, did not generalize from the isolated objects to the within-scene objects. This is consistent with the proposal that PPA processes global scene characteristics rather than within-scene objects⁷⁻¹⁰.

Next, the authors tested whether averaging the patterns evoked by two isolated objects (such as a bathtub and toilet) would improve the prediction of lateral occipital cortex activity patterns evoked by the full scene (such as a bathroom). Notably, a significantly better prediction of the scene-evoked response patterns consisted of the mean of the response patterns evoked by the two isolated objects as compared with response patterns evoked by one of the two objects separately. This provides evidence that response patterns in lateral occipital cortex carry information about multiple objects present in the scene. Thus, responses to multiple objects seem to be combined linearly in lateral occipital cortex, similarly to the way responses to multiple objects without scene background are computed in this region¹¹.

In the study described above, participants viewed the scenes for 1 s, which is sufficient to make several eye movements and to process the objects one by one. Because of the poor temporal resolution of fMRI, responses to two serially processed objects would not be discriminable from responses to two objects processed in parallel. To provide evidence for

parallel processing of the two objects in the scenes, the authors conducted a second study and presented the scenes for just 150 ms, followed by a masking pattern. Such a short presentation time is too brief to make eye movements and helps to reduce (although not eliminate) serial deployment of attention. Despite the much shorter presentation time, the results were remarkably similar to the results of the first study, suggesting that a brief glance at a scene is sufficient for lateral occipital cortex to process multiple within-scene objects.

The authors conclude that lateral occipital cortex, through its ability to represent multiple within-scene objects, embodies an object-based channel for scene recognition. Such a mechanism is likely to be most useful for the recognition of scenes that are closely related to the identity of one or multiple foreground objects, such as a kitchen or a bathroom: the information that a bathtub and a toilet are present in the scene would be sufficient to infer a bathroom scene. In other cases, however, scenes may share the same set of objects and may only be discriminable on the basis of spatial relations between objects (for example, parking lot versus highway) or on the basis of global scene characteristics (for example, mountain versus desert). MacEvoy and Epstein¹ propose that an additional pathway, which includes the PPA, may support recognition of scenes based on global scene properties. This dual-pathway hypothesis is consistent with recent studies that found that the PPA contains information about the spatial layout of scenes^{8,9}, whereas object-selective cortex contains information about within-scene objects^{7,9}. At a behavioral level, coherent scene context facilitates the perception of within-scene objects^{12,13}, and objects facilitate scene recognition when these are mutually coherent¹³. A critical question then is where and how the two proposed pathways interact to support scene recognition. In particular, does lateral occipital cortex represent objects independently of scene context or does it integrate object representations with contextual information (Fig. 1)?

Owing to capacity limitations, the brain cannot possibly represent all of the information that is typically present in our cluttered natural environment. In daily life, we overcome such limitations by using cognitive ‘top-down’ mechanisms. Such goal-driven mechanisms may enhance the processing of relevant information and filter out irrelevant information. For example, we actively search for objects by forming mental templates of the things that we are looking for (such as a friend in a crowd of people), and we make predictions about the presence or absence of objects on the basis of typical scene characteristics (for example, we anticipate a bathtub when we enter a bathroom). Accordingly, previous research has shown that the processing of real-world scenes is biased toward relevant objects³ and that this selection is mediated by anticipatory processes⁴. In contrast, MacEvoy and Epstein’s results¹ suggest that the averaging of neural responses evoked by multiple objects in lateral occipital cortex is a highly automatic and ‘bottom-up’ process, given that participants were not instructed to attend to any particular object in the rapidly presented scenes. Thus, the object-based channel for scene recognition might operate independently of top-down sources that provide feedback to the object recognition system. Alternatively, it is possible that the repeated presentation of a relatively small number of scene types and diagnostic objects may have prompted participants to anticipate these objects while viewing the scenes. An important question for future studies is to address how many within-scene objects can be processed in parallel without prior knowledge of the scenes and objects that will be presented¹⁴.

In sum, MacEvoy and Epstein¹ showed that the response in lateral occipital cortex to multiple within-scene objects can be approximated by a linear combination of responses to these objects in isolation. This raises the intriguing possibility that lateral occipital cortex constitutes a pathway involved in scene recognition based on a scene's component objects.

References

1. MacEvoy SP, Epstein RA. *Nat Neurosci.* 2011; 14:1323–1329. [PubMed: 21892156]
2. Thorpe S, Fize D, Marlot C. *Nature.* 1996; 381:520–522. [PubMed: 8632824]
3. Peelen MV, Fei-Fei L, Kastner S. *Nature.* 2009; 460:94–97. [PubMed: 19506558]
4. Peelen MV, Kastner S. *Proc Natl Acad Sci USA.* 2011; 108:12125–12130. [PubMed: 21730192]
5. Haxby JV, et al. *Science.* 2001; 293:2425–2430. [PubMed: 11577229]
6. Epstein R, Kanwisher N. *Nature.* 1998; 392:598–601. [PubMed: 9560155]
7. Goh JO, et al. *J Neurosci.* 2004; 24:10223–10228. [PubMed: 15537894]
8. Kravitz DJ, Peng CS, Baker CI. *J Neurosci.* 2011; 31:7322–7333. [PubMed: 21593316]
9. Park S, Brady TF, Greene MR, Oliva A. *J Neurosci.* 2011; 31:1333–1340. [PubMed: 21273418]
10. Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L. *Proc Natl Acad Sci USA.* 2011; 108:9661–9666. [PubMed: 21593417]
11. MacEvoy SP, Epstein RA. *Curr Biol.* 2009; 19:943–947. [PubMed: 19446454]
12. Biederman I, Mezzanotte RJ, Rabinowitz JC. *Cognit Psychol.* 1982; 14:143–177. [PubMed: 7083801]
13. Davenport JL, Potter MC. *Psychol Sci.* 2004; 15:559–564. [PubMed: 15271002]
14. Evans KK, Horowitz TS, Wolfe JM. *Psychol Sci.* 2011; 22:739–746. [PubMed: 21555522]

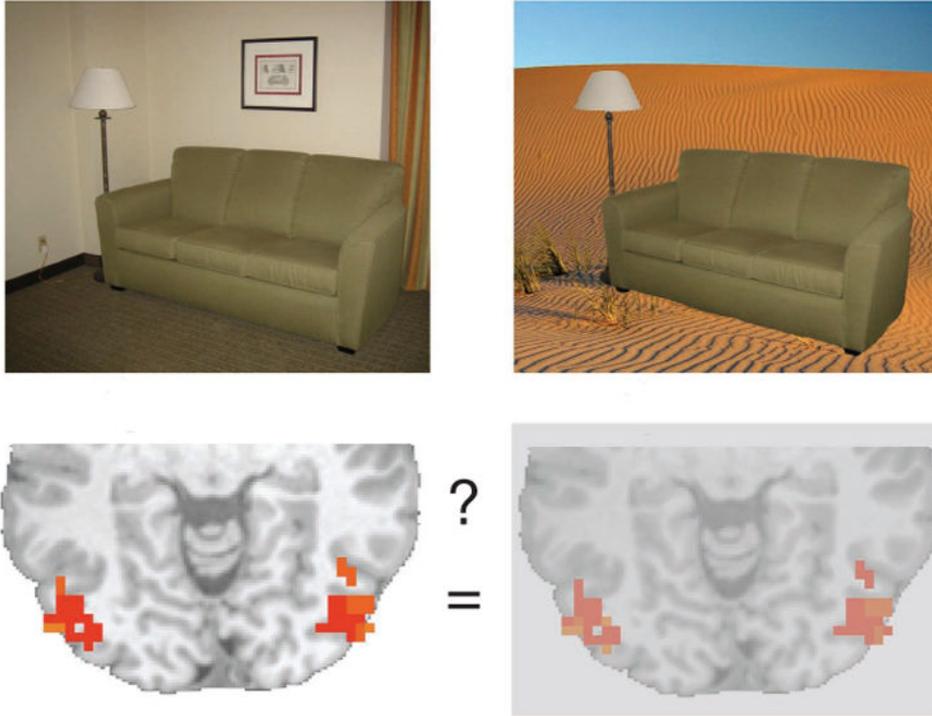


Figure 1. Does lateral occipital cortex represent scene category or within-scene objects? One possibility is that lateral occipital cortex responds to within-scene objects independently of scene category: lateral occipital response patterns would be similar on presentation of the same set of objects in a congruent context, in an incongruent context and in isolation. In this scheme, lateral occipital cortex may contribute to scene recognition by ‘listing’ the objects present in a scene, but would not represent scene category (for example, desert or living room) directly. Alternatively, lateral occipital cortex may integrate object with contextual information (computed in other regions), responding differently to the same object in congruent versus incongruent context. In this scenario, lateral occipital cortex would distinguish scene categories even when they largely share the same set of objects, suggesting a more direct role in scene representation. The results of MacEvoy and Epstein¹, together with those of other recent studies⁷⁻¹⁰, are more consistent with the first scenario.