

Temporal Components in the Parahippocampal Place Area Revealed by Human Intracerebral Recordings

Julien Bastin,^{1,2} Juan R. Vidal,³ Seth Bouvier,⁵ Marcela Perrone-Bertolotti,³ Damien Bénis,^{1,2} Philippe Kahane,^{2,4} Olivier David,^{1,2} Jean-Philippe Lachaux,³ and Russell A. Epstein⁵

¹Fonctions Cérébrales et Neuromodulation, Université Joseph Fourier, 38042 Grenoble, France, ²Institut National de la Santé et de la Recherche Médicale, U836, Grenoble Institut des Neurosciences, 38042 Grenoble, France, ³Institut National de la Santé et de la Recherche Médicale U1028, CNRS UMR 5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, France and Université Claude Bernard, Lyon 1, 69100 Lyon, France, ⁴Laboratoire de Physiopathologie de l'Épilepsie, Centre Hospitalier Universitaire, 38700 Grenoble, France, and ⁵Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Many high-level visual regions exhibit complex patterns of stimulus selectivity that make their responses difficult to explain in terms of a single cognitive mechanism. For example, the parahippocampal place area (PPA) responds maximally to environmental scenes during fMRI studies but also responds strongly to nonscene landmark objects, such as buildings, which have a quite different geometric structure. We hypothesized that PPA responses to scenes and buildings might be driven by different underlying mechanisms with different temporal profiles. To test this, we examined broadband γ (50–150 Hz) responses from human intracerebral electroencephalography recordings, a measure that is closely related to population spiking activity. We found that the PPA distinguished scene from nonscene stimuli in ~ 80 ms, suggesting the operation of a bottom-up process that encodes scene-specific visual or geometric features. In contrast, the differential PPA response to buildings versus nonbuildings occurred later (~ 170 ms) and may reflect a delayed processing of spatial or semantic features definable for both scenes and objects, perhaps incorporating signals from other cortical regions. Although the response preferences of high-level visual regions are usually interpreted in terms of the operation of a single cognitive mechanism, these results suggest that a more complex picture emerges when the dynamics of recognition are considered.

Introduction

Functional magnetic resonance imaging (fMRI) studies have identified occipitotemporal brain regions that respond preferentially to specific visual categories, including faces, bodies, and scenes. The interpretation of these results is a matter of considerable debate. Some argue that these regions are dedicated to recognition of optimal stimulus categories (Kanwisher, 2010). Others argue that they support more general recognition mechanisms that can be engaged by both optimal and suboptimal stimuli (Gauthier et al., 1999; Bar, 2004). In both cases, however, the assumption is that the response of each region can be explained by a single underlying cognitive mechanism.

Here we use local field potential (LFP) recordings to explore an alternative view. Rather than postulating a single mechanism that explains all selectivities within a brain region, we suggest that

multiple recognition mechanisms might process different kinds of information at different time points. To test this idea, we examined LFP responses in the parahippocampal place area (PPA) during viewing of scenes and objects. The PPA is a canonical category-selective region defined by maximal response to visual scenes (e.g., landscapes, street scenes, rooms) (Epstein and Kanwisher, 1998). Notably, the PPA also distinguishes among nonscene categories. For example, it responds more strongly to buildings than to common everyday objects (Epstein, 2005) and more strongly to objects that are larger, more distant, more space-defining, or more suitable as landmarks than to objects that are smaller, closer, less space-defining, and less suitable as landmarks (Committeri et al., 2004; Janzen and van Turennout, 2004; Cate et al., 2011; Mullally and Maguire, 2011; Amit et al., 2012; Konkle and Oliva, 2012; Bastin et al., 2013). Attempts to explain this complex set of scene- and object-related selectivities in terms of a single mechanism face a fundamental challenge because scenes and objects differ substantially in their visual and geometric properties. Thus, it is unclear how a mechanism specialized for distinguishing between scenes could also be efficient for distinguishing between objects.

This conundrum could potentially be resolved if scene- and object-related processing in the PPA were driven by two distinct mechanisms. We hypothesized that scene selectivity in the PPA might be driven in part by bottom-up processing of scene-specific features and thus would be exhibited early, consistent with behavioral reports that scene interpretation occurs very rap-

Received Sept. 25, 2012; revised April 22, 2013; accepted May 12, 2013.

Author contributions: J.B., S.B., J.-P.L., and R.A.E. designed research; J.B., J.R.V., M.P.-B., D.B., and P.K. performed research; J.B., J.R.V., M.P.-B., and D.B. analyzed data; J.B., O.D., J.-P.L., and R.A.E. wrote the paper.

This work was supported by Fondation pour la Recherche Médicale and by Agence Nationale de la Recherche (CONTINT "OpenVibe2"; ANR blanc "MLA"). We thank all patients for their participation; the staff of the Grenoble Neurological Hospital epilepsy unit; and Dominique Hoffmann, Patricia Boschetti, Carole Chatelard, and Véronique Dorlin for their support.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Julien Bastin, Institut des Neurosciences de Grenoble, Bâtiment Edmond J. Safra des Neurosciences, Chemin Fortuné Ferrini, Université Joseph Fourier, Site Santé La Tronche, BP 170, 38042 Grenoble, Cedex 9, France. E-mail: julien.bastin@ujf-grenoble.fr.

DOI:10.1523/JNEUROSCI.4646-12.2013

Copyright © 2013 the authors 0270-6474/13/3310123-09\$15.00/0

idly (Biederman et al., 1974; Potter, 1975; Greene and Oliva, 2009), whereas object selectivity might reflect the incorporation of top-down inputs about the size, meaning, orientational value, or spatial qualities of the object and thus would be exhibited late. To quantify the time course of scene- and object-related selectivity, we determined the amount of time it takes the LFP signal to distinguish scenes from objects and the amount of time it takes to distinguish between two object categories (buildings and non-buildings). We focus in particular on broadband γ activity (BGA, 50–150 Hz), as recent studies combining multiunit recordings with LFP in humans (Manning et al., 2009) and macaque monkeys (Ray and Maunsell, 2011) have demonstrated strong positive correlations between BGA and co-occurring spike rates at the population level.

Materials and Methods

Participants. Intracranial recordings were obtained from nine neurosurgical patients with intractable epilepsy (five women; mean age, 27 ± 10 years) at the Epilepsy Department of the Grenoble University Hospital. To localize epileptic foci that could not be identified through noninvasive methods, neural activity was monitored in lateral, intermediate, and medial wall structures in these patients using stereotactically implanted multilead electrodes (stereotactic electroencephalography [sEEG]). Electrode implantation was performed according to routine clinical procedures, and all target structures for the presurgical evaluation were selected strictly according to clinical considerations with no reference to the current study. Patients selected for the study were those whose electrodes primarily sampled visual areas (see Fig. 3). All participants had normal or corrected to normal vision and provided written informed consent. Experimental procedures were approved by the Institutional Review Board and by the National French Science Ethical Committee.

Electrode implantation. Eleven to 15 semirigid, multilead electrodes were stereotactically implanted in each patient. Electrodes had a diameter of 0.8 mm and, depending on the target structure, contained 10–15 contact leads 2-mm-wide and 1.5-mm-apart (DIXI Medical Instruments). All electrode contacts were identified on each patient's individual postimplantation MRI. Each subject's individual preimplantation MRI was coregistered with the postimplantation MRI (Carmichael et al., 2008) to determine the anatomical location of each contact and to compute all contact coordinates in the Montreal Neurological Institute (MNI) space using standard Statistical Parametric Mapping algorithms. Visual inspection of the contact locations was also used to check whether each sEEG contact was located in gray or white matter.

Intracranial EEG recordings. sEEG recordings of 128 contacts in each patient were conducted using a commercial video-sEEG monitoring system (System Plus, Micromed). The data were bandpass-filtered online from 0.1 to 200 Hz and sampled at 512 Hz, using a reference electrode located in white matter. Each electrode trace was subsequently rereferenced with respect to its direct neighbor (bipolar derivations with a spatial resolution of 3.5 mm) to achieve high local specificity by cancelling out effects of distant sources that spread equally to both adjacent sites through volume conduction. All electrodes exhibiting epileptiform signals were discarded from the present study. This was achieved in collaboration with the medical staff and was based on visual inspection of the recordings and by systematically excluding data from any electrode site that was a posteriori found to be located within the seizure onset zone. Overall, 863 recording sites were obtained over the nine patients included in this study.

Experimental procedure. Each patient participated in two experiments. In Experiment 1, they viewed images from six stimulus categories (scenes, houses, faces, animals, objects, and scrambled images; see Fig. 1, left). Comparison of responses across these six categories allowed us to identify electrode contacts that responded in a scene-selective manner equivalent to that typically observed in the fMRI-defined PPA. We refer to contacts that express a scene-selective response pattern as the sEEG-defined PPA. In Experiment 2, patients viewed images of building or nonbuilding objects, presented either in a scenic context or on a gray

background (see Fig. 1, right). Comparison of response between scenes (objects with scenic background) and isolated objects (objects without scenic background) allowed us to determine the time course of scene-related selectivity, whereas comparison of response between buildings and nonbuildings allowed us to determine the time course of object-related selectivity. (Note that under our formulation, buildings are considered "objects" because they are compact entities with a bounding contour.)

Pictures in Experiment 1 were grayscale pictures that had the same average luminance and were presented within an oval aperture (400×300 pixels). Pictures were sequentially presented for 200 ms each followed by a 800–1000 ms interstimulus interval. Patients viewed blocks of 5 randomly chosen pictures, and each block was followed by 3 s pause periods during which patients could freely blink. Fifty exemplars were shown for each of the six categories. Participants were instructed to press a button each time a picture of a fruit appeared on screen (visual oddball paradigm, 19 oddballs of 319 pictures total). Pictures in Experiment 2 were full-color scenic images (400×400 pixels) collected from internet searches. Pictures were sequentially presented for 200 ms each followed by a 800–1000 ms interstimulus interval in blocks of six randomly chosen pictures interleaved by 3 s pause periods. A total of 79 images were used for each four categories (building with scenic background, building with gray background, object with scenic background, and object with gray background) for a total of 316 images in the complete stimulus set. Participants performed a 1-back repetition detection task, which required them to press a button whenever they saw two identical pictures in a row (6% of trials).

Visual stimuli were delivered on a 19 inch TFT monitor with a refresh rate of 60 Hz, controlled by a PC (Pentium 133, Dos) with Presentation 14.1 (Neurobehavioral Systems). The monitor was placed 60 cm away from the subject's eyes. To control the timing of stimulus delivery, a transistor-transistor logic pulse was sent by the stimulation PC to the EEG acquisition PC. Patients responded to the tasks through right-hand button presses.

Data analysis. The first step of our analysis was to determine the typical response to visual stimulation (Bastin et al., 2012). To do this, we used Elan-pack (developed at the laboratory Institut National de la Santé et de la Recherche Médicale U1028) to compute standard time-frequency (TF) wavelet decomposition, which allowed us to identify major TF components of interest across all trials at each recording site (Lachaux et al., 2003). The frequency range was 1–200 Hz, and the time interval included a 250 ms prestimulus baseline and lasted until 800 ms after stimulus onset. Wilcoxon nonparametric tests were used to compare, across the trials, the total energy in a given TF tile ($100 \text{ ms} \times 3 \text{ Hz}$) with that of a tile of similar frequency extent covering the prestimulus baseline period from -600 to -100 ms.

The TF decomposition analyses indicated a strong response in the γ range (50–150 Hz) during visual stimulation (e.g., see Fig. 2A), so we focused on BGA in the subsequent analyses. BGA was extracted with the Hilbert transform of sEEG signals (MATLAB, MathWorks) using the following procedure. sEEG signals were first bandpass filtered in 10 successive 10-Hz-wide frequency bands (e.g., 10 bands, beginning with 50–60 Hz up to 140–150 Hz). For each bandpass filtered signal, we computed the envelope using standard Hilbert transform. The obtained envelope had a time resolution of 15.625 ms. Again, for each band, this envelope signal (i.e., time-varying amplitude) was divided by its mean across the entire recording session and multiplied by 100 for normalization purposes. Finally, the envelope signals computed for each consecutive frequency bands (e.g., 10 bands of 10 Hz intervals between 50 and 150 Hz) were averaged together, to provide one single time-series (the BGA) across the entire session, expressed as percentage of the mean. This time series was then divided into trialwise responses, which were analyzed in three steps as detailed below.

First, we used the BGA responses from Experiment 1 to identify scene-selective contacts for further analysis. For each sEEG contact, we calculated average BGA during the poststimulus interval (0–800 ms) for each of the 50 scene trials and 50 object trials. We then compared these values using a *t* test ($df = 98$) and corrected the *p* values for multiple comparisons over the $n = 96 \pm 6$ contacts for each subject using the false discov-

ery rate (Genovese et al., 2002). Contacts that showed significantly greater response to scenes than to objects were deemed to be PPA contacts.

Next, we examined the response magnitudes across conditions for the PPA contacts in Experiments 1 and 2. To do this, we calculated the average trialwise BGA for each PPA contact during the poststimulus interval (0–800 ms). We then averaged these trialwise values across all trials for each condition to get a single response value for each condition at each contact. We then performed ANOVA and *post hoc* tests on these values.

Finally, we examined the timing of scene- and object-related response in Experiment 2. For these analyses, the scene effect was defined as the difference in response between background-present and background-absent stimuli (regardless of the identity of the focal object), whereas the object effect was defined as the difference between buildings and non-building objects (regardless of whether a background was present or not). The latencies of these effects were compared using two complementary methods. In the first set of analyses, we determined the onset and peak latency for each effect at each contact and then compared these two quantities using a two-tailed *t* test to test the hypothesis that the scene effect should onset and peak earlier than the object effect. To calculate onset and peak latencies for each contact for the first set of analyses, we calculated the strength of the scene and object effect at each 15.625 ms time bin in the (–250 to 1000 ms) time interval, using single trial responses at each time bin as observations (*t* test, *df* = 198). The onset latency for each effect was taken to be the first time bin at which a significant *p* value (false discovery rate corrected for 81 multiple comparisons in the time domain) was observed and subsequently maintained for at least 125 ms (8 bins), whereas the peak latency for each effect was taken to be the time point at which the difference between conditions reached its maximum. In the second set of analyses, we compared the relative strength of the two effects after first normalizing each effect to its peak magnitude. We compared their normalized magnitudes at each time point in the (0–500 ms) time interval, using the average responses across trial at each time bin as observations (*t* test, *df* = 13). This analysis tests the idea that the scene effect should be larger than the object effect at early time points, but not at later time points when both effects are fully developed.

Mapping intracranial EEG data to standard MNI brain. The anatomical representation of all significant BGA modulations was obtained by pooling data from all subjects and mapping it into the standard MNI brain atlas using normalized contact coordinates.

Results

Experiment 1: BGA in the PPA is stronger during scene viewing

In the first experiment, we identified the PPA using BGA as a neural marker of scene selectivity. Previous fMRI (Epstein and Kanwisher, 1998; Epstein et al., 1999; Epstein et al., 2003) and sEEG (Bastin et al., 2013) studies have demonstrated that a region in the collateral sulcus, including posterior parahippocampal cortex and adjoining territory in the lingual and fusiform gyrii, responds significantly more strongly to scenes than to objects, and we wished to replicate these results. sEEG data were collected from a total of 863 intracranial bipolar recording sites across nine patients while they viewed pictures drawn from six categories (houses, faces, animals, scenes, objects, and scrambled images; Fig. 1). Higher spectral power was found in a broadband γ frequency range (50–150 Hz) during picture viewing than during fixation (Fig. 2A) at many occipitotemporal sites. BGA was stronger during scene than during object viewing in 14 sEEG contacts (Fig. 2B, C; *t* test on single trial BGA measured during the 0–800 ms after stimulus; *df* = 98; *p* < 0.05 corrected for multiple comparisons across all contacts for each subject). These contacts, found in 6 of 9 subjects (Table 1; Fig. 3), were clustered around the collateral sulcus ~1 cm behind the parahippocampal/lingual boundary (Fig. 2B; center MNI coordinates: ± 27 , -53 , -9 ; see

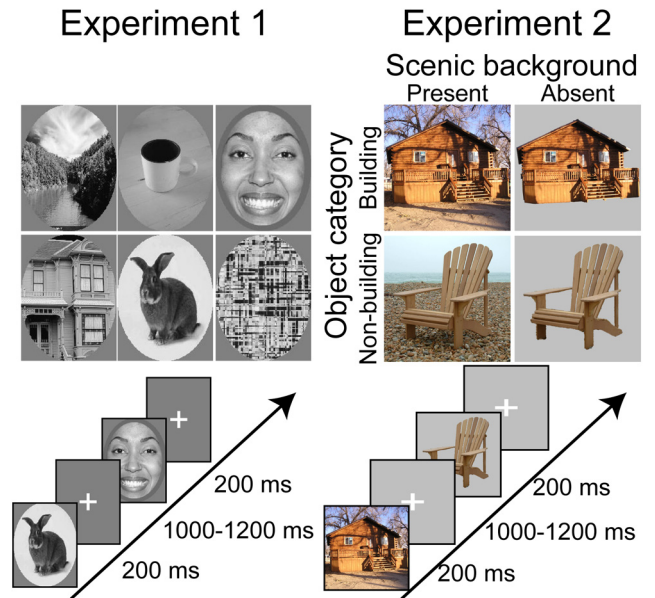


Figure 1. Experimental design: examples of stimuli and presentation rates used in Experiments 1 and 2. Stimulus duration was set to 200 ms during both experiments.

also Table 1; Fig. 4) and will be referred to the PPA in the following.

Next, we examined BGA in the 0–800 ms time interval in the PPA for all stimulus categories (Fig. 2D). Repeated-measures ANOVA revealed that stimulus category had a significant effect on BGA ($F_{(5,65)} = 54.5$; *p* < 0.001, using averaged BGA observed for each PPA contact as observations). *Post hoc* analyses confirmed the PPA selectivity to scene stimuli (scene > all other categories, Newman–Keuls test, *p* < 0.01) and also revealed a strong PPA response to buildings (house > [object, face, and scrambled stimuli], Newman–Keuls test, *p* < 0.01). The apparent trend toward greater PPA response to animals compared with objects was not significant. Notably, the relative level of response across scenes, buildings, objects, and faces was almost identical to that observed in previous fMRI studies (Epstein and Kanwisher, 1998).

To exclude the possibility that the relative differences between scenes and objects could have been overestimated because the same dataset was used to identify PPA and to estimate PPA selectivity profile, we ran a control analysis that used half of the trials (25 randomly selected trials) to identify PPA contacts and the other half to plot the selectivity profile of the PPA. This analysis showed very similar results: we identified the same 14 PPA contacts, and the selectivity profile was identical ($F_{(5,65)} = 49$; *p* < 0.001 and Newman–Keuls test, *p* < 0.01).

Experiment 2: BGA onsets dissociate scenic background from object category encoding

Although the PPA responds strongly to both scene and building stimuli, we reasoned that these responses might be driven by different underlying mechanisms; that is, in information processing terms, they may reflect fundamentally different computations. Scenes and objects are quite different in their visual and geometric properties, so it is feasible that greater response to scenes than to objects in the PPA could reflect bottom-up processing of properties unique to scenes; for example, information about spatial layout provided by large fixed scene elements, such as walls and ground planes. In contrast, distinguishing buildings

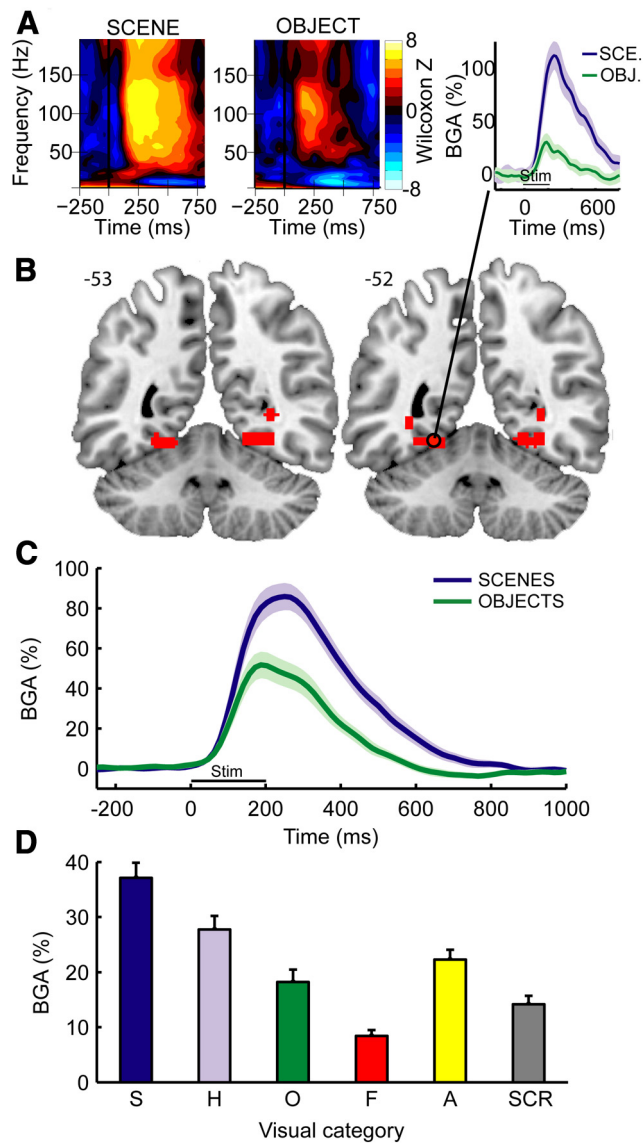


Figure 2. TF response and broadband BGA in the PPA. **A**, Example of statistical TF maps (Wilcoxon z scores) and BGA response time course (right) during scene and object viewing from a single recording site (Patient 2; MNI coordinates: $-23, -54, -12$) located in the posterior part of the collateral sulcus. Shaded areas represent ± 1 SEM across trials. **B**, Anatomical location of the contacts that respond preferentially to scenes. Two adjacent slices from MNI brain are shown. Numbers on the right of each image indicate MNI coordinates in the anterior–posterior (y) axis. **C**, Average BGA responses aligned to stimulus onset during scene and object encoding in scene-selective contacts ($n = 14$ contacts recorded from 6 patients). Shaded areas represent ± 1 SEM across contacts. **D**, Average BGA responses during a 0–800 ms time interval to six different stimulus types in these contacts. S, Scenes; H, houses; O, objects; F, faces; A, animals; SCR, scrambled. Error bars represent SEM across contacts.

from nonbuildings is a more challenging problem, which might require analysis of higher-level object properties, such as size, fixedness, and meaning. The stimuli in Experiment 1 were not optimal for testing this hypothesis because many of the object stimuli were taken under naturalistic conditions and included some scenic background; thus, the scene versus object difference in this case would reflect a conflation of the two predicted effects. To test this idea, therefore, we ran a second experiment, using stimuli that allowed us to directly compare the temporal profiles of the scene- and building-related response in the PPA.

The same nine patients that participated in Experiment 1 viewed objects (building and nonbuilding) presented either in a

Table 1. Anatomical locations of the sEEG-defined PPA contact-pairs

Electrode name (contact-pairs)	MNI coordinates			Patient
	x	y	z	
f2-1	-28	-51	-11	P1
f3-2	-31	-51	-11	P1
f3-2	-19	-54	-12	P2
f4-3	-23	-54	-12	P2
f5-4	-26	-54	-11	P2
f4-3	22	-53	-9	P2
f6-5	30	-53	-9	P2
f2-1	20	-54	-10	P5
f3-2	24	-54	-10	P5
f2-1	-35	-51	-2	P7
x4-3	31	-54	2	P8
f2-1	24	-52	-11	P9
f3-1	28	-52	-11	P9
f4-3	31	-52	-11	P9

scenic context or on a gray background (Fig. 1right). ANOVA revealed that response was higher for background compared with no-background stimuli (scene effect, $F_{(1,13)} = 81.1, p < 0.01$) and also higher for buildings than for nonbuilding objects (object effect, $F_{(1,13)} = 20.3, p < 0.01$). There was a significant interaction between these two effects ($F_{(1,13)} = 34.8, p < 0.01$), reflecting the fact that the object effect was significant for no-background stimuli (Newman–Keuls, $p < 0.01$) but not for background stimuli (Newman–Keuls, $p = 0.69$, not significant). The pattern (Fig. 5A) is consistent with previous observations that the PPA responds more strongly to scenes (with-background stimuli) than to non-scenes (no-background stimuli) regardless of object content, and also responds more strongly to buildings than to nonbuildings when these stimuli are presented in isolation (Epstein and Kanwisher, 1998; Troiani et al., 2012).

To compare the time courses of the scene and object effects (Fig. 5B,C), we calculated the onset latency and the timing of the peak response for each recording site (Fig. 5D). The onset of the scene effect was identified for 13 of 14 PPA contacts; the onset of the object effect was identified in 10 of 14 contacts (in 5 of 6 patients); the peak response was identified for both effects in all 14 contacts. The average onset latency of the scene effect (background present vs background absent) was 82 ± 10 ms, whereas the average onset latency of the object effect (building vs nonbuilding) was 170 ± 34 ms (mean \pm SEM). Crucially, when the onsets of the two effects were compared, the onset of the scene effect ($n = 13$) was found to precede the onset of the object effect ($n = 10$) by a significant amount (unpaired t test: $t_{(21)} = 2.6, p = 0.014$). Echoing this result, the peak of the scene effect (223 ± 10 ms, $n = 14$) also preceded the peak of the object effect (340 ± 38 ms, $n = 14$) by a significant amount (unpaired t test: $t_{(26)} = 3; p = 0.006$). Substantially identical results were obtained when the object effect was defined based on response to no-background trials only: in this case, the onset of the object category effect was 176 ± 107 ms ($n = 11$, unpaired t test vs scene onset: $t_{(22)} = 2.8, p = 0.009$), whereas the peak was 288 ± 26 ms ($n = 14$, unpaired t test vs scene onset: $t_{(26)} = 2.3; p = 0.026$).

As a second test of the relative timing of the scene and object effects, we compared their magnitudes at different time bins, after first normalizing each effect to its peak magnitude for each contact. We found that the normalized scene effect was significantly greater than the normalized object effect from 203 to 266 ms after stimulus onset (paired t tests; $df = 13; p < 0.05$ after false discovery rate correction). This result reflects the fact that the scene effect is more fully actualized than the object effect at earlier time

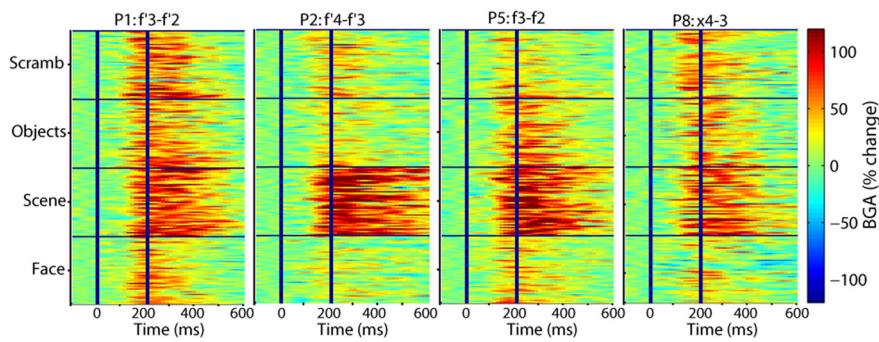


Figure 3. Reproducibility across trials and across patients of BGA responses recorded from the PPA during Experiment 1 (2 left PPA and 2 right PPA contacts recorded from four different patients are shown).

points (in the sense that it has reached a higher percentage of its peak). This analysis provides additional support for the idea that the scene effect develops more rapidly than the object effect.

To further examine reliability of results across subjects, we reanalyzed our data after first averaging across contacts within each patient's hemisphere ($n = 7$ PPA hemispheres). The selectivity profile of the PPA (Experiment 1) was unaffected by this analysis, and the time course of the scene and building effects were qualitatively identical (i.e., the scene effect always preceded the building effect); however, the statistical power was too low to reach significance. Finally, to exclude the possibility that low-level differences between stimuli might explain results from Experiment 2, we conducted another control analysis. We first identified sEEG contacts that were more sensitive to objects than to scenes in Experiment 1, restricting this analysis to object-sensitive electrodes that were found to be located more laterally than the PPA contacts (objects > scene contacts, $n = 9$ sEEG contacts). We found that these object contacts were neither reliably sensitive nor inversely sensitive to the manipulations that affected the PPA contacts in Experiment 2, making it impossible to define the timing of the scene and object effects in this case.

Effects of spatial frequency

Previous studies indicate that the PPA responds significantly more strongly to stimuli that are dominated by high spatial frequencies (HSF) compared with stimuli that contain less relative power at HSF (Rajimehr et al., 2011). To explore whether it was possible to explain the scene and/or object effects in Experiment 2 in terms of differences of spatial frequency content in the stimulus set, we examined the spatial frequency content of the images as a function of background-presence and object category. First, we extracted the high (>5 cycles/°) spatial frequency (HSF) power for each image using a previously described procedure (Rajimehr et al., 2011). Then we examined these HSF power values using a 2×2 ANOVA, with background-presence and object category as factors. Background-present stimuli had significantly more HSF power than background-absent stimuli ($F_{(1,268)} = 10.3$; $p < 0.05$). Building stimuli, on the other hand, did not differ in spatial frequency from nonbuilding stimuli ($F_{(1,268)} = 0.07$, $p = 0.78$, not significant), and the interaction between background presence and object category was not significant ($F_{(1,268)} = 0.37$, $p = 0.54$, not significant).

To test the possibility that these SF differences could explain the PPA preference for stimuli with scenic background, we compared BGA response to pictures with more HSF power ($n = 100$ highest HSF power pictures) with BGA response to pictures with less HSF power ($n = 100$ low HSF). Because we wished to exam-

ine the effects of SF and stimulus category independently, 50 of the pictures for each SF condition were background-present stimuli (specifically, the 25 highest/lowest SF buildings with background and the 25 highest/lowest SF nonbuildings with background) and 50 were background-absent stimuli (also distributed equally between buildings and nonbuildings). To verify the efficacy of this manipulation, we performed a 2×2 ANOVA on HSF power for the 200 selected images. This confirmed the expected difference between high and low SF pictures ($F_{(1,195)} = 161.7$; $p < 0.01$) and showed that, for this reduced stimulus set, there was no significant difference in SF

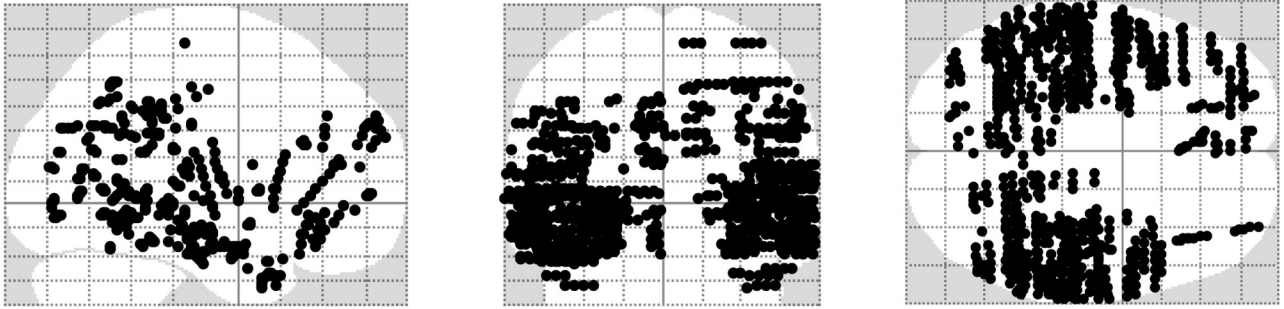
between background-present and background-absent stimuli ($F_{(1,195)} = 2.84$; $p = 0.09$, not significant) and no significant interaction ($F_{(1,195)} < 1$, not significant).

We then performed the same analysis on BGA power in PPA contacts in the (0–800 ms) time interval. This analysis confirmed the significant effect of background on BGA ($F_{(1,13)} = 93$; $p < 0.01$; the average percentage change of BGA was $26 \pm 2.7\%$). The analysis also revealed a significant effect of high spatial frequency ($F_{(1,13)} = 6.81$; $p = 0.02$; the average percentage change of BGA was $2.8 \pm 1.1\%$). These results argue against the idea that the background effect can be explained by SF differences, as the background effect is observed here even for a reduced stimulus set for which SF is controlled, and the magnitude of the background effect is 10 times bigger than the magnitude of the SF effect.

Discussion

The main finding of this study is that broadband γ responses at scene-selective sEEG recording sites (the “sEEG-defined PPA”) distinguish between scenes and nonscenes more quickly than they distinguish between buildings and nonbuildings. Furthermore, the anatomical location of the scene-selective recording sites, and the ordering of responses across stimulus categories, was similar to that observed previously with fMRI (Epstein, 2005). Our results suggest the existence of two information processing stages in the human PPA: an early stage that distinguishes scenes from nonscenes, followed by a later stage that distinguishes navigationally relevant objects (i.e., buildings) from other objects.

Although the speed at which the human visual system distinguishes between faces and objects have been studied in previous sEEG experiments (Privman et al., 2007; Tsuchiya et al., 2008; Fisch et al., 2009; Liu et al., 2009), this is the first precise measurement, to our knowledge, of the speed at which it distinguishes between scenes, buildings, and objects. Previous neurophysiological work in monkeys has shown that selectivity for scene stimuli is exhibited in some occipitotemporal neurons at a latency of ~ 125 ms after stimulus onset (Bell et al., 2011). In humans, the neuronal latency of the scene selectivity within the PPA is unclear: some parahippocampal neurons increased their activity compared with fixation ~ 270 ms after visual stimulus onset (Mormann et al., 2008), but these neurons did not specifically respond to scene stimuli and were therefore probably located outside of the PPA, in the main body of the human parahippocampal cortex. We previously reported preliminary results from two patients indicating that BGA was selectively modulated in the collateral sulcus during scenes compared with objects (Bastin et al., 2013). Here we replicate this finding in a larger cohort of

A • sEEG contacts
B

- sEEG PPA
- Objects contacts
- Visual (non selective) contacts
- No task-related BGA

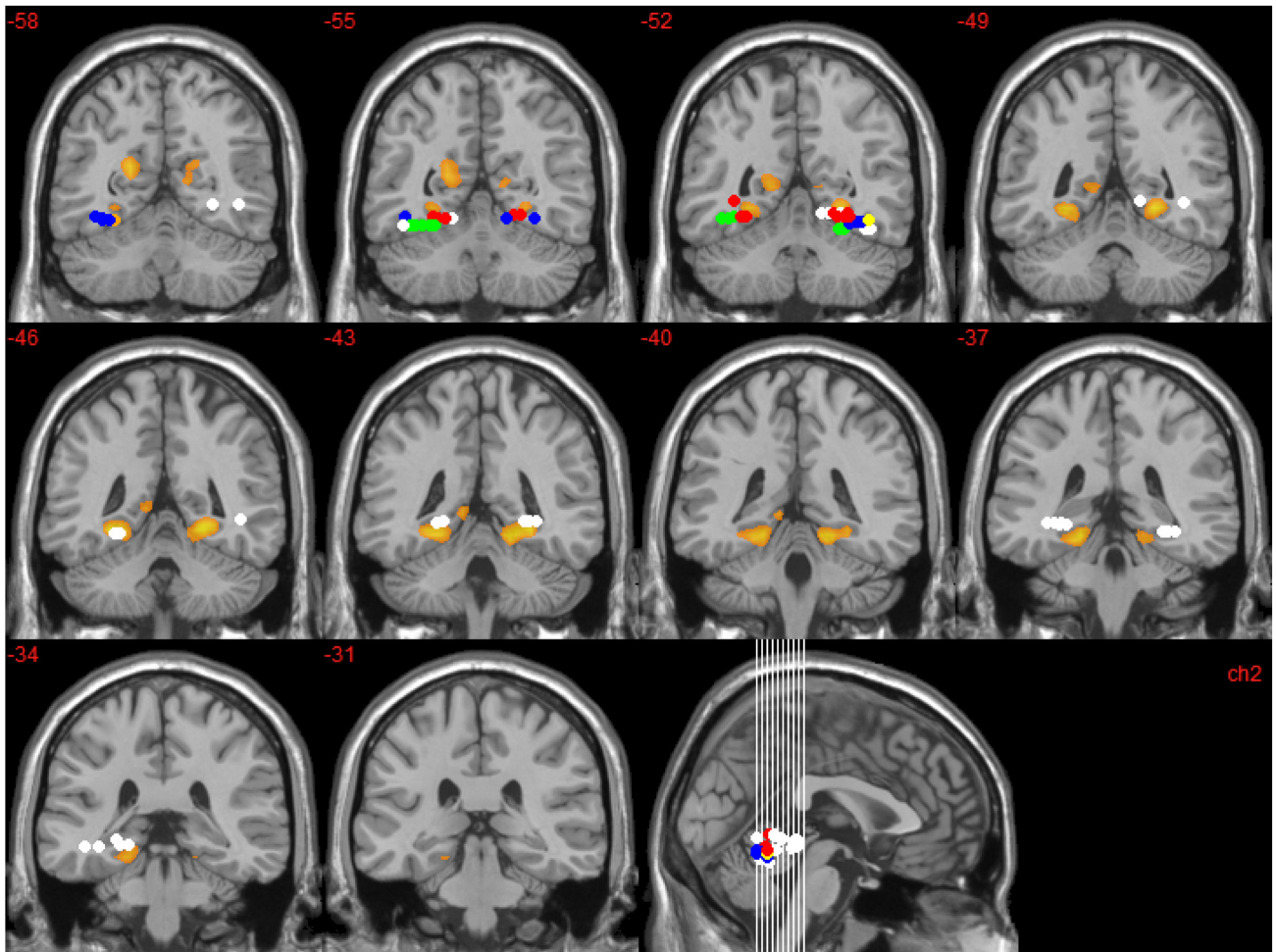


Figure 4. Anatomical locations of all intracerebral contact-pairs (**A**) and relationship between fMRI-defined PPA and contact-pairs that were within an ROI centered on the fMRI-defined PPA (MNI coordinates: $abs(x) < 45$; $-35 < y < -60$; $-18 < z < 0$) (**B**). Voxels that respond more to scenes than to objects are shown in orange (fMRI data from R.A.E. laboratory). sEEG contacts that were located within the ROI were categorized into four functionally distinct response types: red dots represent sEEG-defined PPA contacts (i.e., BGA for these contacts was selective to scene stimuli); green dots, contacts that were selective to objects; blue dots, contacts that showed a BGA increase that was similar across categories (visual nonselective); white dots, sEEG contacts that did not respond to visual stimulation (nonresponsive). The sEEG-defined PPA fits within the posterior tail of the fMRI-defined PPA. The center of fMRI-defined PPA was not sampled by intracerebral contacts. The more lateral location of intracerebral contacts selective to objects is consistent with the neuroimaging literature.

patients and show that scenes and objects can be distinguished at ~ 80 ms (comparable to latencies observed in monkey visual cortices), whereas buildings and nonbuildings are distinguished later, at 170 ms.

We hypothesize that the early scene-selective component might reflect bottom-up processing of visual or geometric el-

ements unique to scenes, such as spatial layout or visual summary statistics. Previous fMRI studies have indicated that these global scene properties are extracted by the PPA (Epstein and Kanwisher, 1998; Kravitz et al., 2011a; Park et al., 2011; Cant and Xu, 2012) and might underlie the PPA's critical involvement in scene categorization and scene recognition

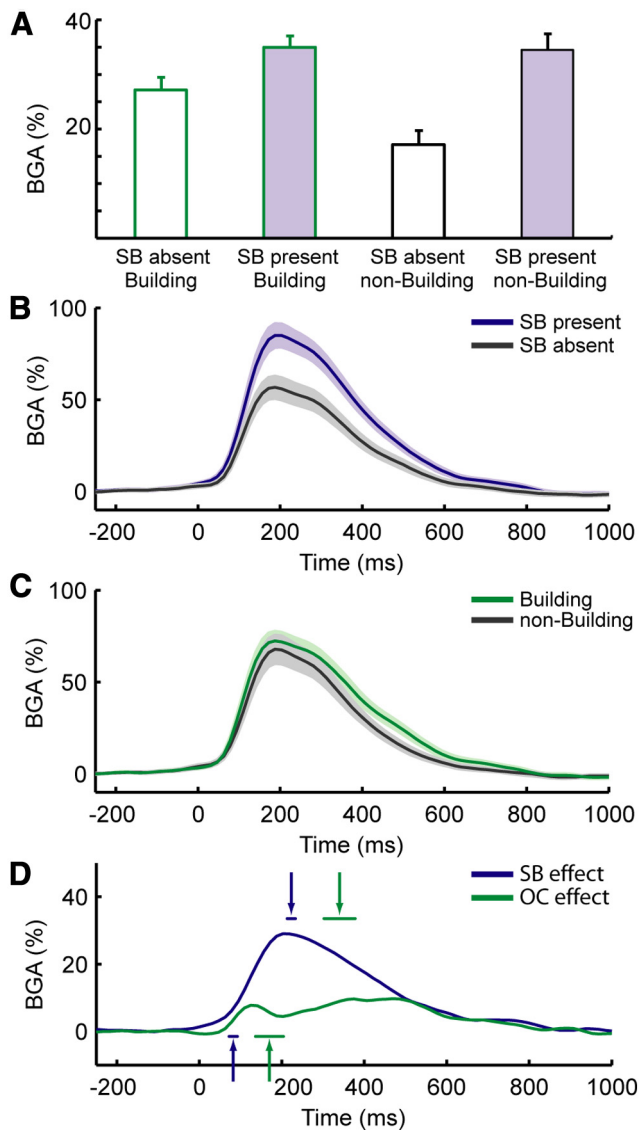


Figure 5. BGA at PPA contacts in Experiment 2. **A**, Grand-average BGA activity during the four conditions of Experiment 2 (average BGA during the 0–800 ms after stimulus interval): building with no background, building with background, nonbuilding object without background, and nonbuilding object with background. **B**, Grand-average BGA responses during viewing of stimuli that either had a scenic background (SB present) or not (SB absent). **C**, Grand-average BGA responses during viewing of stimuli that either contained buildings or nonbuilding objects. **D**, Time course of the scene (SB present — SB absent stimuli) versus object (building — nonbuilding stimuli) effects. Arrows indicate the average latencies of the onsets (arrows pointing upward) and peak times (arrows pointing downward) of experimental effects. Shaded areas and horizontal lines below or above arrows represent ± 1 SEM. SB effect, Scenic background effect; OC effect, object category effect.

(Habib and Sirigu, 1987; Steeves et al., 2004; Walther et al., 2009). Dovetailing with these results, behavioral studies indicate that human observers have a remarkable ability to interpret complex visual scenes very rapidly, perhaps in part through analysis of such global scene properties (Biederman et al., 1974; Potter, 1975; Antes et al., 1981; Li et al., 2002; Greene and Oliva, 2009). Our current results provide an important data point that links together these previous findings by showing that scene-specific responses within the PPA are exhibited at short latencies, exactly as one would expect if the PPA extracts global scene properties for discrimination between scenes using a feedforward mechanism.

The object category effect, on the other hand, might reflect a later stage of processing in which information about some (but not necessarily all) objects is incorporated into the PPA response. Recent studies have convincingly demonstrated PPA sensitivity to object qualities. For example, the PPA has been shown to respond more strongly to the following: large objects compared with small objects (Cate et al., 2011; Konkle and Oliva, 2012), space-defining objects compared with nonspace-defining objects (Mullally and Maguire, 2011), objects in navigationally relevant locations compared with objects in less navigationally relevant locations (Janzen and van Turenout, 2004), nearby objects compared with distant objects (Amit et al., 2012), and objects strongly associated with a place or context compared with objects only weakly associated with a place or context (Bar and Aminoff, 2003; but see Epstein and Ward, 2010). Attempts to explain these object-related responses in terms of scene-recognition mechanisms face a challenge because scenes and objects are quite different in their visual and geometric properties; for example, objects are compact, convex, and have a well-defined border, whereas scenes are distributed, concave, and have no clear border (Epstein and MacEvoy, 2011). As such, it seems unlikely that the same bottom-up mechanisms used for scene recognition would be useful for extracting information about objects. This is especially true given that the PPA appears to be sensitive to higher-level object properties, such as size, fixedness, and spatial orientation value, which are not immediately ascertainable on the basis of appearance. Our results suggest a possible solution to this problem: the PPA might incorporate this object-based information at a later stage, relying in part on top-down or lateral signals from other areas of the brain.

Although it is unclear what kind of information about the objects is encoded at this later stage, we can suggest two (nonexclusive) possibilities. The first is that the later response component reflects coding of abstract or semantic aspects of the objects, such as information about their size (Konkle and Oliva, 2012), contextual associations (Bar, 2004), or their value as a navigational landmark (Janzen and van Turenout, 2004). The second possibility is that the later component reflects spatial analyses that occur for some objects but not others (Mullally and Maguire, 2011). We have previously proposed that the PPA represents information about the layout of local space. If signals from other brain regions indicate to the PPA that an object is large enough and stable enough to define the permanent geometry of local space, then the PPA might engage more strongly in an effort to extract the spatial coordinate frame anchored to the object, in effect turning the “object” into a “scene” (Committeri et al., 2004; Troiani et al., 2012; Bastin et al., 2013). Under this account, the PPA might indeed have two functions: (1) scene identification performed by the early component; and (2) the extraction of navigationally useful spatial information from a stimulus, regardless of whether the stimulus is a scene or an object, performed by the later component. Consistent with this idea, we previously found a late response (~ 600 – 800 ms after stimulus onset) in the PPA when subjects were asked to judge the location of a target object relative to an allocentric coordinate frame anchored on a building (Bastin et al., 2013), suggesting that late components in the PPA reflect spatial processes that go beyond the perceptual analysis of an item.

Under the first proposal (abstract/semantic coding), the later stage might involve the incorporation of inputs from object-processing areas, such as the lateral occipital complex (Harel et

al., 2013) or frontal lobe regions involved in object interpretation (Bar et al., 2006). Under the second proposal (spatial processing), the later stage might involve the incorporation of spatial inputs from the retrosplenial and medial parietal regions, which ultimately originate from the dorsal stream (Suzuki and Amaral, 1994; Stefanacci et al., 1996; Kravitz et al., 2011b; Bastin et al., 2013). Although the latter account may seem to contradict previous observations that neurons in the dorsal visual stream have faster latencies than neurons in the ventral stream (Schmolsky et al., 1998), we note it is unclear whether these latencies apply to the medial parietal regions that provide input to the PPA. Furthermore, because our contact sites are posterior to parahippocampal cortex proper, which is the locus of parietal afferents, it is possible that spatial information would take additional time to propagate backwards to these locations.

In conclusion, these results demonstrate the existence of two temporal components in the PPA response. Whereas scenes and nonscenes are distinguished relatively early, navigationally relevant objects (buildings) and non-navigationally relevant objects (nonbuildings) are distinguished relatively late. These findings suggest the possibility that the PPA represents different kinds of information at different points in time. We suspect that similar temporal dissociations might exist in other high-level visual regions (Sugase et al., 1999), such as the fusiform face area (Kanwisher, 2000; Bukach et al., 2006), and might explain otherwise conflicting accounts of the functions of these regions.

References

- Amit E, Mehoudar E, Trope Y, Yovel G (2012) Do object-category selective regions in the ventral visual stream represent perceived distance information? *Brain Cogn* 80:201–213. [CrossRef Medline](#)
- Antes JR, Penland JG, Metzger RL (1981) Processing global information in briefly presented pictures. *Psychol Res* 43:277–292. [CrossRef Medline](#)
- Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5:617–629. [CrossRef Medline](#)
- Bar M, Aminoff E (2003) Cortical analysis of visual context. *Neuron* 38:347–358. [CrossRef Medline](#)
- Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hamalainen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E (2006) Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A* 103:449–454. [CrossRef Medline](#)
- Bastin J, Lebranchu P, Jerbi K, Kahane P, Orban G, Lachaux JP, Berthoz A (2012) Direct recordings in human cortex reveal the dynamics of gamma-band [50–150 Hz] activity during pursuit eye movement control. *Neuroimage* 63:339–347. [CrossRef Medline](#)
- Bastin J, Committeri G, Kahane P, Galati G, Minotti L, Lachaux JP, Berthoz A (2013) Timing of posterior parahippocampal gyrus activity reveals multiple scene processing stages. *Hum Brain Mapp* 34:1357–1370. [CrossRef Medline](#)
- Bell AH, Malecek NJ, Morin EL, Hadj-Bouziane F, Tootell RB, Ungerleider LG (2011) Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J Neurosci* 31:12229–12240. [CrossRef Medline](#)
- Biederman I, Rabinowitz JC, Glass AL, Stacy EW Jr (1974) On the information extracted from a glance at a scene. *J Exp Psychol* 103:597–600. [CrossRef Medline](#)
- Bukach CM, Gauthier I, Tarr MJ (2006) Beyond faces and modularity: the power of an expertise framework. *Trends Cogn Sci* 10:159–166. [CrossRef Medline](#)
- Cant JS, Xu Y (2012) Object ensemble processing in human anterior-medial ventral visual cortex. *J Neurosci* 32:7685–7700. [CrossRef Medline](#)
- Carmichael DW, Thornton JS, Rodionov R, Thornton R, McEvoy A, Allen PJ, Lemieux L (2008) Safety of localizing epilepsy monitoring intracranial electroencephalograph electrodes using MRI: radiofrequency-induced heating. *J Magn Reson Imaging* 28:1233–1244. [CrossRef Medline](#)
- Cate AD, Goodale MA, Köhler S (2011) The role of apparent size in building- and object-specific regions of ventral visual cortex. *Brain Res* 1388:109–122. [CrossRef Medline](#)
- Committeri G, Galati G, Paradis AL, Pizzamiglio L, Berthoz A, LeBihan D (2004) Reference frames for spatial cognition: different brain areas are involved in viewer-, object-, and landmark-centered judgments about object location. *J Cogn Neurosci* 16:1517–1535. [CrossRef Medline](#)
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601. [CrossRef Medline](#)
- Epstein R, Harris A, Stanley D, Kanwisher N (1999) The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23:115–125. [CrossRef Medline](#)
- Epstein R, Graham KS, Downing PE (2003) Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37:865–876. [CrossRef Medline](#)
- Epstein RA (2005) The cortical basis of visual scene processing. *Visual Cognition* 12:954–978. [CrossRef](#)
- Epstein RA, MacEvoy SP (2011) Making a scene in the brain. In: *Vision in 3d environments* (Harris L, Jenkin M, eds). Cambridge: Cambridge UP.
- Epstein RA, Ward EJ (2010) How reliable are visual context effects in the parahippocampal place area? *Cereb Cortex* 20:294–303. [CrossRef Medline](#)
- Fisch L, Privman E, Ramot M, Harel M, Nir Y, Kipervasser S, Andelman F, Neufeld MY, Kramer U, Fried I, Malach R (2009) Neural “ignition”: enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron* 64:562–574. [CrossRef Medline](#)
- Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (1999) Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nat Neurosci* 2:568–573. [CrossRef Medline](#)
- Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15:870–878. [CrossRef Medline](#)
- Greene MR, Oliva A (2009) Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cogn Psychol* 58:137–176. [CrossRef Medline](#)
- Habib M, Sirigu A (1987) Pure topographical disorientation: a definition and anatomical basis. *Cortex* 23:73–85. [Medline](#)
- Harel A, Kravitz DJ, Baker CI (2013) Deconstructing visual scenes in cortex: gradients of object and spatial layout information. *Cereb Cortex* 23:947–957. [CrossRef Medline](#)
- Janzen G, van Turenout M (2004) Selective neural representation of objects relevant for navigation. *Nat Neurosci* 7:673–677. [CrossRef Medline](#)
- Kanwisher N (2000) Domain specificity in face perception. *Nat Neurosci* 3:759–763. [CrossRef Medline](#)
- Kanwisher N (2010) Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc Natl Acad Sci U S A* 107:11163–11170. [CrossRef Medline](#)
- Konkle T, Oliva A (2012) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74:1114–1124. [CrossRef Medline](#)
- Kravitz DJ, Peng CS, Baker CI (2011a) Real-world scene representations in high-level visual cortex: it’s the spaces more than the places. *J Neurosci* 31:7322–7333. [CrossRef Medline](#)
- Kravitz DJ, Saleem KS, Baker CI, Mishkin M (2011b) A new neural framework for visuospatial processing. *Nat Rev Neurosci* 12:217–230. [CrossRef Medline](#)
- Lachaux JP, Rudrauf D, Kahane P (2003) Intracranial EEG and human brain mapping. *J Physiol Paris* 97:613–628. [CrossRef Medline](#)
- Li FF, VanRullen R, Koch C, Perona P (2002) Rapid natural scene categorization in the near absence of attention. *Proc Natl Acad Sci U S A* 99:9596–9601. [CrossRef Medline](#)
- Liu H, Agam Y, Madsen JR, Kreiman G (2009) Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62:281–290. [CrossRef Medline](#)
- Manning JR, Jacobs J, Fried I, Kahana MJ (2009) Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J Neurosci* 29:13613–13620. [CrossRef Medline](#)
- Mormann F, Kornblith S, Quiroga RQ, Kraskov A, Cerf M, Fried I, Koch C (2008) Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J Neurosci* 28:8865–8872. [CrossRef Medline](#)
- Mullally SL, Maguire EA (2011) A new role for the parahippocampal cortex in representing space. *J Neurosci* 31:7441–7449. [CrossRef Medline](#)
- Park S, Brady TF, Greene MR, Oliva A (2011) Disentangling scene content

- from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *J Neurosci* 31:1333–1340. [CrossRef Medline](#)
- Potter MC (1975) Meaning in visual search. *Science* 187:965–966. [CrossRef Medline](#)
- Privman E, Nir Y, Kramer U, Kipervasser S, Andelman F, Neufeld MY, Mukamel R, Yeshurun Y, Fried I, Malach R (2007) Enhanced category tuning revealed by intracranial electroencephalograms in high-order human visual areas. *J Neurosci* 27:6234–6242. [CrossRef Medline](#)
- Rajimehr R, Devaney KJ, Bilenko NY, Young JC, Tootell RB (2011) The “parahippocampal place area” responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biol* 9:e1000608. [CrossRef Medline](#)
- Ray S, Maunsell JH (2011) Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol* 9:e1000610. [CrossRef Medline](#)
- Schmolsky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG (1998) Signal timing across the macaque visual system. *J Neurophysiol* 79:3272–3278. [Medline](#)
- Steeves JK, Humphrey GK, Culham JC, Menon RS, Milner AD, Goodale MA (2004) Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *J Cogn Neurosci* 16:955–965. [CrossRef Medline](#)
- Stefanacci L, Suzuki WA, Amaral DG (1996) Organization of connections between the amygdaloid complex and the perirhinal and parahippocampal cortices in macaque monkeys. *J Comp Neurol* 375:552–582. [CrossRef Medline](#)
- Sugase Y, Yamane S, Ueno S, Kawano K (1999) Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400:869–873. [CrossRef Medline](#)
- Suzuki WA, Amaral DG (1994) Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J Comp Neurol* 350:497–533. [CrossRef Medline](#)
- Troiani V, Stigliani A, Smith ME, Epstein RA (2012) Multiple object properties drive scene-selective regions. *Cereb Cortex*. Advanced online publication Dec. 12, 2012. doi: 10.1093/cercor/bhs364.
- Tsuchiya N, Kawasaki H, Oya H, Howard MA 3rd, Adolphs R (2008) Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. *PLoS One* 3:e3892. [CrossRef Medline](#)
- Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009) Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci* 29:10573–10581. [CrossRef Medline](#)