Scene-selective coding by single neurons in the human parahippocampal cortex

Florian Mormann[a,b,c,1,2], Simon Kornblith[b,d,1], Moran Cerf[b,c,e], Matias J. Ison[f,g], Alexander Kraskov[b,h], Michelle Tran[c], Simeon Knieling[a], Rodrigo Quian Quiroga[b,f], Christof Koch[b,i], and Itzhak Fried[c,j,k]

[a]Department of Epileptology, University of Bonn, 53105 Bonn, Germany; [b]Division of Biology, California Institute of Technology, Pasadena, CA 91125; [c]Department of Neurosurgery and Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, CA 90095; [d]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; [e]Kellogg School of Management, Northwestern University, Evanston, IL 60208; [f]Centre for Systems Neuroscience and Department of Engineering, University of Leicester, Leicester LE1 7QR, United Kingdom; [g]School of Psychology, University of Nottingham, Nottingham NG7 2RD, United Kingdom; [h]Institute of Neurology, University College London, London WC1N 3BG, United Kingdom; [i]Allen Institute for Brain Science, Seattle, WA 98109; [j]Functional Neurosurgery Unit, Tel-Aviv Sourasky Medical Center, Tel Aviv 64239, Israel; and [k]Functional Neurosurgery Unit, Sackler Faculty of Medicine, Tel Aviv 64239, Israel

Edited by Larry R. Squire, Veterans Affairs San Diego Healthcare System, San Diego, CA, and approved November 4, 2016 (received for review May 21, 2016)

Imaging, electrophysiological, and lesion studies have shown a relationship between the parahippocampal cortex (PHC) and the processing of spatial scenes. Our present knowledge of PHC, however, is restricted to the macroscopic properties and dynamics of bulk tissue; the behavior and selectivity of single parahippocampal neurons remains largely unknown. In this study, we analyzed responses from 630 parahippocampal neurons in 24 neurosurgical patients during visual stimulus presentation. We found a spatially clustered subpopulation of scene-selective units with an associated event-related field potential. These units form a population code that is more distributed for scenes than for other stimulus categories, and less sparse than elsewhere in the medial temporal lobe. Our electrophysiological findings provide insight into how individual units give rise to the population response observed with functional imaging in the parahippocampal place area.

electrophysiology | single units | scene selectivity | population code

The involvement of posterior parahippocampal cortex (PHC) in perceiving landmarks and scenes is well established. Studies using fMRI and intracranial electroencephalography (iEEG) have demonstrated that a region in posterior PHC exhibits significantly greater activation to passively viewed scenes and landscapes than to single objects or faces (1, 2). Moreover, damage to posterior PHC produces anterograde disorientation, a deficit in the ability to navigate in novel environments (3, 4), and electrical stimulation in this area produces complex topographic visual hallucinations (5). Beyond gross scene-selectivity, other studies suggested that the parahippocampal place area (PPA) responds more strongly to outdoor scenes (1), to images of objects with a spatial background (6), to objects that are larger in the real world regardless of retinotopic size (7–9), and to images with greater perceived depth (10, 11). However, as a voxel in a typical fMRI study corresponds to several cubic millimeters of cortex, and as iEEG contacts record the activity of large numbers of neurons, our present knowledge of PHC is restricted to the properties and dynamics of bulk tissue properties (12). The selectivity of single parahippocampal neurons thus remains largely unknown.

In this work we set out to investigate the single neuron responses underlying these results. At least three different, although not necessarily mutually exclusive, types of single neuron selectivity profiles could potentially produce the scene-selective population response observed with fMRI and iEEG recordings. First, units could exhibit sparse responses, each of them tuned to one or relatively few individual scenes, similar to the semantically invariant neurons observed in the human medial temporal lobe that fire selectively to specific familiar individuals (13). In this case, the scene-selective responses observed with fMRI and iEEG would be given by the spatial average of neurons with different responses. Second, each unit could be scene-selective, but respond to many scenes, thus representing a distributed code, as found in macaque face and scene patches (14, 15). Third, units might represent a low-level feature or conjunction of features present in both scene and nonscene stimuli, but more prevalent in the former, such that population activity to scenes exceeds that to nonscenes. In this scenario, strong scene selectivity would be present at the population level, but single neurons would be only weakly scene-selective. Neurocomputational models of PHC function are scarce (16, 17) and do not make specific predictions about the sparseness of neuronal scene responses. By analyzing the responses of single neurons in PHC to visual stimuli in subjects with pharmacologically intractable epilepsy, we sought to determine how viewing pictures of scenes modulated spiking responses of individual parahippocampal neurons.

Results

We recorded a total of 1,998 units (668 single and 1,330 multiunits) (Table S1) from the hippocampus (829 single and multiunits), entorhinal cortex (EC; 539 units), and PHC (630 units) of 24 neurosurgical patients undergoing epilepsy monitoring while they viewed images on an LCD monitor. Stimulus sets contained images of persons, animals, and landscapes (with and without buildings; *Materials and Methods*).

PHC Neurons Respond to Landscapes and Scenes. Although neurons in EC and hippocampus showed little consistent preference for any particular stimulus category, neurons in PHC responded strongly to landscapes (Figs. 1 and 2 and Fig. S1). To statistically compare neuronal selectivity across regions and categories, we

Significance

Neurons in the human parahippocampal cortex explicitly code for scenes, rather than people, animals, or objects. More specifically, they respond to outdoor pictures, rather than to indoor pictures, and to stimuli with rather than without spatial layout. These scene-selective neurons are spatially clustered and receive spatially clustered inputs reflected by an event-related local field potential (LFP). Furthermore, these neurons form a distributed population code that is less sparse than codes found elsewhere in the human medial temporal lobe. Our findings thus provide insight into the electrophysiological (single unit and LFP) substrates underlying the parahippocampal place area, a structure well-known from neuroimaging.

Author contributions: F.M., S. Kornblith, C.K., and I.F. designed research; F.M., M.C., M.J.I., A.K., R.Q.Q., and I.F. performed research; F.M., M.T., and S. Knieling assessed electrode localizations; I.F. performed all neurosurgical procedures; F.M. and S. Kornblith analyzed data; and F.M., S. Kornblith, C.K., and I.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]F.M. and S. Kornblith contributed equally to this work.

[2]To whom correspondence should be addressed. Email: florian.mormann@ukb.uni-bonn.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1608159113/-/DCSupplemental.

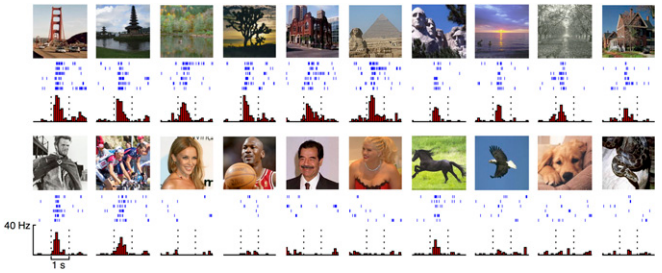

Fig. 1. Typical response of a scene-selective single neuron in PHC to a variety of landscape stimuli (*Upper*) and nonlandscape stimuli (*Lower*). Landscape stimuli and nonlandscape stimuli with indications of spatial layout (scenes) elicited robust responses, whereas stimuli without indications of spatial layout (nonscenes) elicited no response. Note that because of insurmountable copyright problems, all original celebrity pictures were replaced by very similar ones (same person, similar background, etc.) from the public domain. Images courtesy of Wikimedia Commons/ThiloK, flickr/Renan Katayama, Basketballphoto.com/Steve Lipofsky, flickr/doggiesrule04.

calculated the mean baseline-normalized response magnitude of every neuron to each stimulus category. Comparison of the mean response to different stimulus categories in the three MTL regions showed a highly significant category selectivity in the PHC ($P < 10^{-12}$, repeated-measures one-way ANOVA). Landscapes evoked a significantly stronger response than persons ($P < 10^{-9}$, paired t test) or animals ($P = 0.0003$; Fig. 3*A* and Fig. S2*A*).

Furthermore, outdoor photographs evoked a significantly stronger response than indoor photographs (Fig. S2*B*), even after excluding landscape stimuli, which consisted exclusively of outdoor photographs (Fig. 3*B*; $P < 10^{-9}$). In addition, all stimuli were divided into groups with and without cues of spatial layout (Fig. 3*C*, Fig. S2*C*, and *Materials and Methods*), subsequently referred to as scenes and nonscenes, respectively. Within both outdoor and indoor categories (excluding landscapes), the PHC neurons responded more strongly to scenes than to nonscenes (outdoor: $P < 10^{-5}$; indoor: $P < 10^{-6}$, paired t test).

We ran an additional series of analyses to investigate relationships between response magnitude and stimulus content. We divided images into three categories: no background present, background present but unrecognizable, and background clearly visible. The responses of parahippocampal neurons to images of persons and animals with a recognizable background were significantly stronger than the responses to images with no background or an unrecognizable background (vs. no background: $P < 10^{-10}$; vs. unrecognizable background: $P < 10^{-6}$, paired t test; Fig. S3 *A* and *B*), but there was only a minor difference between images with no background and images with an unrecognizable background ($P = 0.02$).

Among images with a recognizable background, images with greater perceived depth (i.e., with more spatial information) evoked a stronger response. We obtained a rank ordering of the real-world distance between the closest and farthest points of images with a recognizable background from 21 nonpatient subjects, using a merge sort procedure (*Materials and Methods*). For each PHC neuron, we computed the Spearman correlation between the average rankings of the images and the corresponding firing rates and compared the mean of the Fisher-z transformed correlation coefficients against zero with a t test. The relationship between depth and firing rate was highly significant [$P < 10^{-5}$; $\bar{\rho} = 0.060$ (95% confidence interval [CI], 0.035–0.084)] and remained significant when including only persons and animals [$P = 0.001$; $\bar{\rho} = 0.038$ (95% CI 0.018–0.057)].

PHC Responses Are Not Explained by Low-Level Features. The above analyses indicate that parahippocampal neurons showed strong selectivity for images with greater indications of spatial layout. It is, however, possible that such selectivity could be driven by selectivity for low-level features that were more commonly present in images with spatial layout. To rule out this possibility, we trained a linear support vector machine (SVM) classifier to discriminate images with and without spatial layout based on low-level visual features, using the hierarchical model and X complex 1 (HMAX C1) layer, which is intended to model neural representation at the level of V1 with additional scale invariance (18). For each stimulus, we obtained a specific label for the stimulus, using a classifier trained on the remaining stimuli. This procedure was 89% accurate at reproducing our manual labels, correctly identifying 93% of nonscenes and 71% of scenes. We then computed the mean responses for each of the 630 recorded single and multiunits in PHC for stimuli within four categories: stimuli manually labeled as nonscenes that the classifier also classified as nonscenes (true negatives), stimuli manually labeled as nonscenes that the classifier classified as scenes (false positives), stimuli manually labeled as scenes that the classifier also labeled as scenes (true positives), and stimuli manually labeled as scenes that the classifier labeled as nonscenes (false negatives).

If neurons were more strongly tuned to low-level features than to the presence or absence of spatial layout, we would expect that nonscenes that the classifier incorrectly classified as scenes (false positives) should elicit high response magnitude, whereas scenes that the classifier incorrectly classified as nonscenes (false negatives) should elicit low response magnitude. This was not the case. Instead, false negatives elicited significantly stronger responses than false positives ($P < 10^{-5}$, paired t test; Fig. S3*C*). Nonscenes elicited similar responses regardless of the classifier output ($P = 0.14$), although scenes classified as scenes elicited a slightly

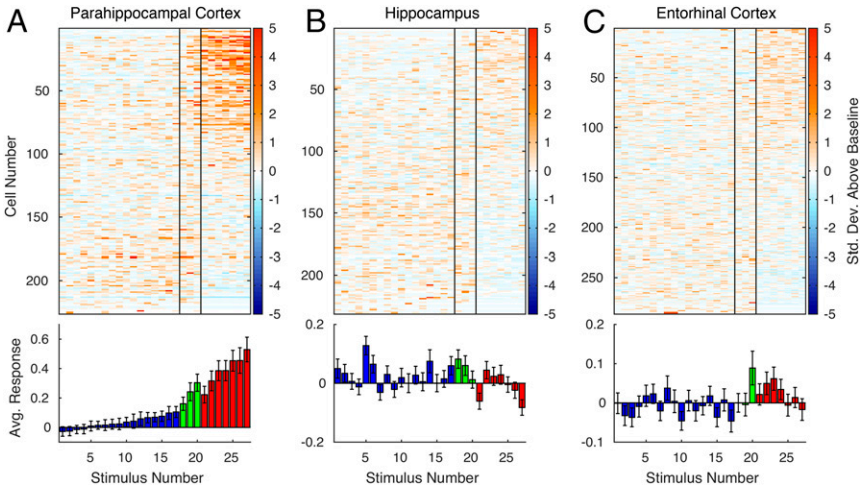

Fig. 2. Responses of 226 parahippocampal (*A*), 231 hippocampal (*B*), and 286 entorhinal (*C*) single and multiunits to the same 27 stimuli (see *Materials and Methods* for selection procedure) comprising persons (*Left*, blue), animals (*Middle*, green), and landscapes (*Right*, red), normalized by prestimulus baseline activity. Vertical bars in upper graphs separate stimulus categories. Units are sorted by scene selectivity index. Within each category, stimuli are sorted by average response. Error bars in lower graphs are ±SEM of responses averaged across units.

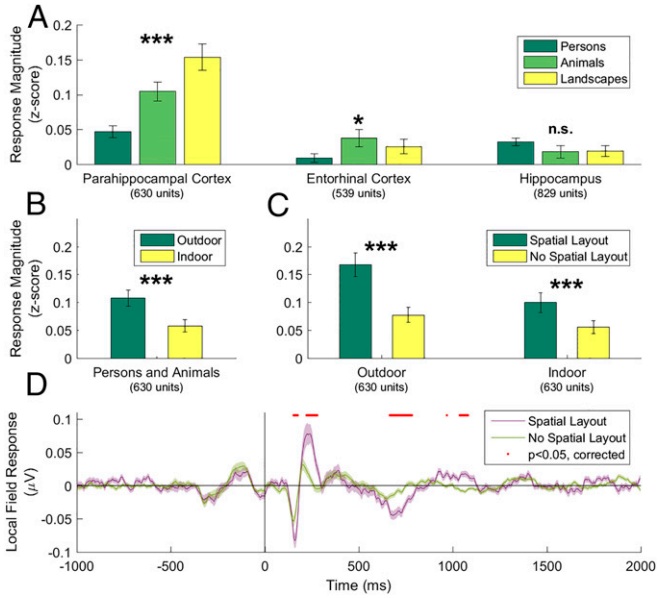

Fig. 3. (*A*) Response magnitudes of single and multiunits in different regions of the medial temporal lobe to three stimulus categories indicate that only parahippocampal neurons respond more strongly to pictures of landscapes than to pictures of persons or animals (repeated-measures ANOVAs). Error bars are ±SEM. (*B*) Even after excluding landscape stimuli, parahippocampal neurons respond more strongly to outdoor photographs than to indoor photographs (*t* test). (*C*) For both outdoor and indoor pictures (excluding landscapes), pictures with spatial layout (scenes) elicit stronger responses in parahippocampal neurons than those without spatial layout (nonscenes). (*D*) Averaged LFP ±SEM for stimuli with or without spatial layout across all 472 parahippocampal microelectrodes. LFP responses to stimuli containing spatial layout (scenes) significantly exceed those to stimuli without spatial layout (nonscenes). Red circles indicate $P < 0.05$ after multiple testing correction. *t* tests: ***$P < 0.001$; *$P < 0.05$; n.s., not significant.

stronger response than scenes classified as nonscenes ($P = 0.002$). Thus, although low-level features may account for some of the scene responsiveness, this analysis indicates that the presence or absence of spatial layout is the primary factor determining the response of parahippocampal neurons. This is furthermore supported by an analysis showing that PHC responses are conditionally independent of the low-level features, given the stimulus category (scene vs. nonscene), as shown in Fig. S4. In addition, the identity of individual landscape stimuli can be decoded from PHC neurons more accurately than from neurons in EC or hippocampus (Fig. S5).

Neurons Respond Faster to Landscapes. On average, PHC neurons responded faster to scenes than to nonscenes. A latency measure could be computed for 217 PHC single and multiunits (*SI Materials and Methods*). In the 121 units responding to at least one scene, the average response latency to scenes was 300 ms, which was significantly faster than the average response latency of 334 ms to persons and animals in the 185 PHC single and multiunits that responded to at least one person or animal ($P < 0.001$, independent samples unequal variance *t* test). Among the 89 single and multiunits that responded both to landscapes and to other stimulus categories, the response to landscapes was significantly faster (median difference = 15.8 ± 7.1 ms; $P = 0.03$, paired samples *t* test).

Neuronal Scene Responses in the PHC Are Spatially Clustered. The local field potential (LFP) measures the global activity of neuronal processes around the electrode tip (19, 20). Thus, an electrode located in a scene-selective region can measure a scene-selective LFP. LFPs from 28% (130/472) of PHC microelectrodes showed a significantly different response to images with and without spatial layout (significance threshold, $\alpha = 0.01$; *t* test). A significant difference was also visible in the average LFP across all electrodes (Fig. 3*D* and Fig. S6), with an onset of selectivity around 153 ms and peaking at 243 ms. Of the 168 single and multiunits recorded on these 130 microelectrodes, 43% ($n = 73$) were also scene-selective (individual units: $\alpha = 0.01$, Mann–Whitney *U* test; population: $P < 10^{-6}$, permutation test; *SI Materials and Methods*; chance median, 19%). Of the 630 PHC single and multiunits, 119 showed a significant category distinction between average responses to scenes and nonscenes ($\alpha = 0.01$, Mann–Whitney *U* test; Fig. S7). Of these, 61% ($n = 73$) single and multiunits were located on microelectrodes with a scene-selective LFP ($P < 10^{-6}$, permutation test; chance median, 28%; Fig. S8). In addition, microwire bundles that showed a scene-selective unit on one of the eight microwires had a significantly increased probability of having scene-selective units on the remaining wires ($P < 10^{-6}$, permutation test; *SI Materials and Methods*). These results indicate that scene-selective units as well as input signals (LFP) are spatially clustered within subjects, consistent with fMRI selectivity for scenes. Across subjects, no spatial clustering of scene-selective microwire bundles was observed (Fig. S9), which is in line with the interindividual variability of the PPA observed in fMRI studies.

Neuronal Scene Responses in the PHC Form a Distributed Code. Units that responded to at least one scene often responded to multiple scenes. A total of 176 (28%) of the 630 PHC single and multiunits responded to at least one image with spatial layout (scene). In comparison, on average, 106.6 of 630 PHC units (17%) responded to at least one image without spatial layout (nonscene) when randomly drawing a set of nonscene stimuli equal to the number of scene stimuli ($P < 10^{-6}$, permutation test; *Materials and Methods*). Of the 176 scene-responsive units, 49 (28%) responded to at least 25% of scenes (Fig. 4*A*); in contrast, of the 106.6 units responsive to nonscenes, on average, only 8.3 (7.8%) responded to at least 25% of a matched number of nonscenes

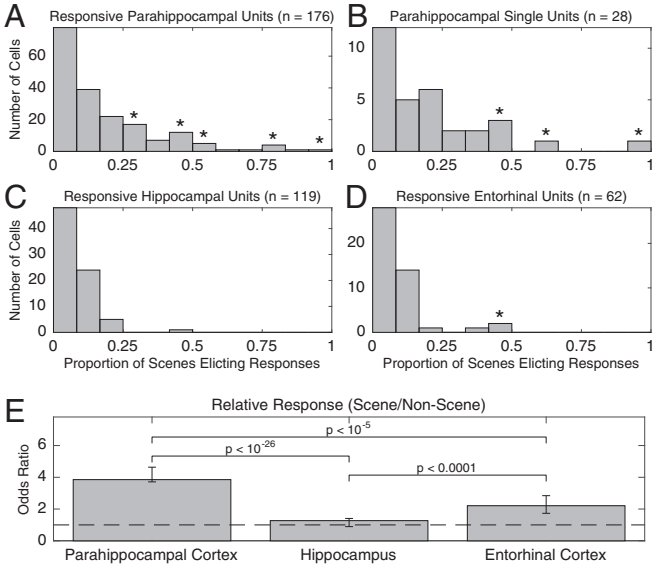

Fig. 4. Histograms of proportion of scene stimuli (i.e., stimuli possessing spatial layout) eliciting responses in cells responding to at least one scene stimulus for single and multiunits (*A*) and single units only (*B*) in PHC, as well as single and multiunits in the hippocampus (*C*) and EC (*D*). Asterisks indicate a significant difference ($P < 0.05$ corrected) from a null distribution, as calculated by drawing, with replacement, a number of nonscene stimuli equal to the number of scene stimuli presented for each session and computing the proportion of stimuli eliciting responses for those nonscene stimuli. (*E*) Mode of the conditional distribution of the common odds ratio $(\text{responses}_{\text{spatial}}/n_{\text{spatial}})/(\text{responses}_{\text{nonspatial}}/n_{\text{nonspatial}})$, the number of times more likely a unit is to respond to a stimulus with spatial layout than a stimulus without spatial layout, for units in each region, given the observed responses. Error bars are 95% CIs.

($P < 10^{-6}$). A comparable effect was present when only single units were included in the analysis (Fig. 4*B*): of the 22% (28/126) of single units responding to at least one scene, 32% ($n = 9$) responded to at least 25% of scenes compared with 4.1% responding to 25% of a matched number of nonscenes ($P < 10^{-6}$). Thus, PHC responses to stimuli with spatial layout were far more distributed than responses in other areas (and to other stimuli; Fig. S10). This category effect was much weaker or not present for the sparser response behavior of entorhinal and hippocampal neurons (Fig. 4 *C* and *D*). In the hippocampus, 14% ($n = 119$) of single and multiunits responded to at least one scene (to nonscenes, 14%; $P = 0.31$), and of those, 0.84% ($n = 1$) responded to at least 25% (to nonscenes, 2.41%; $P = 0.96$), whereas in EC, 12% ($n = 62$) of units responded to at least one scene (to nonscenes, 7%; $P < 10^{-6}$), and of those, 4.8% ($n = 3$) responded to at least 25% (to nonscenes, 0.63%; $P = 0.015$).

We additionally computed and compared the odds ratio (OR) of scene to nonscene responses across regions. This ratio measures how much more likely it is that a unit responds to a stimulus with spatial layout than to a stimulus without spatial layout. We found that in both PHC and EC, but not in the hippocampus, the OR was significantly greater than 1, suggesting units were more likely to respond to scene stimuli than nonscene stimuli (PHC: $P < 10^{-135}$; EC: $P < 10^{-9}$; hippocampus: $P = 0.33$, exact test of common odds ratio; Fig. 4*E*). However, the OR was significantly greater in PHC than in the other two regions (both $P < 10^{-4}$), indicating that parahippocampal units responded to a greater number of scenes relative to nonscenes compared with units elsewhere in the medial temporal lobe.

These findings show that neurons in PHC respond less selectively within their preferred category (i.e., scenes) than within other categories, as well as less selectively than neurons in hippocampus and EC, indicating a more distributed code for scenes in PHC compared with the sparser code in other MTL areas.

Neuronal Scene Responses in the MTL Are Independent of Familiarity. Some scenes, such as the picture of Mt. Rushmore in Fig. 1, were previously known to the subjects, whereas others such as the Victorian house were not. To test whether familiarity of scenes had a differential effect on neural responses in different MTL regions, we divided all scenes into the categories previously known vs. previously unknown and compared responsiveness to both categories by computing an OR. We found no significant differences in any of the three regions examined (PHC: $P = 0.24$; EC: $P = 0.13$; hippocampus: $P = 0.95$; exact test of common odds ratio). This is in good agreement with a previous study reporting no difference in the proportion of MTL cells responding to famous vs. unknown landscapes (21).

Discussion

Our results provide insight into how the human PHC encodes space at the level of individual neurons. The population response was significantly stronger to images that represent space than to images that do not, providing further evidence for correspondence between functional imaging and electrophysiology. PHC neurons prefer landscapes over persons and animals, but even among persons and animals, they prefer outdoor over indoor pictures, and even among these indoor pictures, stimuli with spatial layout are preferred over those without. Furthermore, PHC neurons were found to respond faster to landscapes than to other stimulus categories, indicating a facilitated processing of this stimulus class. In line with results obtained with fMRI and iEEG (1, 2, 5), comparison of LFP and unit responses indicated that scene-selective units were spatially clustered. Furthermore, we observed a distributed code in PHC, one in which the typical responsive unit was relatively specific to scenes in general, but not to any one particular scene, whereas, for example, in the hippocampus, a much sparser code was observed, one in which the typical responsive unit was not specific to scenes in general, but was specific to one particular image.

The scene-selective population response observed at the macroscopic level by fMRI could in principle be produced by three different types of single neuron selectivity profiles. First, units could exhibit sparse responses, each tuned to relatively few individual scenes, similar to the semantically invariant neurons observed in the human medial temporal lobe (13). In this case, the scene-selective responses observed macroscopically would reflect the spatial average of neurons with different responses. Second, each unit could be scene-selective, but respond to many scenes, thus representing a distributed code, as found in macaque face and scene patches (14, 15). Third, units might represent conjunctions of low-level features more prevalent in scene than in nonscene stimuli, such that population activity to scenes exceeds that to nonscenes.

The parahippocampal neurons analyzed in this study, similar to neurons in macaque scene areas (15), but unlike the sparse neurons elsewhere in the human medial temporal lobe (13), showed characteristics of a more distributed code (units that responded to one scene stimulus often responded to many), and these responses could not be attributed to low-level visual features, as established with the analysis using a feature classifier (the HMAX model). The second scenario described here therefore seems most realistic. Of note, a recent study reported a more distributed code for representation of visual stimuli in the human hippocampus (22). This study, however, used a less conservative response criterion, thus trading response selectivity for sensitivity, and contained no comparison with PHC representations.

Our study builds on previous intracranial studies of neural activity in human PHC. Although Ekstrom et al. (23) found no location selectivity in parahippocampal neurons during virtual navigation, they found that 15% of parahippocampal single and multiunits responded to views of shops in the environment vs. <5% in the hippocampus, amygdala, and frontal lobes. Several studies have reported stronger broadband gamma activity in human PHC to scenes vs. objects and earlier selectivity for scenes than for buildings (2, 5), but these studies recorded neural activity using macroelectrodes instead of microwires, and thus could not characterize the responses of individual neurons. Finally, Kraskov et al. (24) previously investigated the selectivity of spikes and LFPs across several medial temporal lobe regions, including the parahippocampal gyrus. However, they reported no parahippocampal electrodes with scene-selective LFPs, and few scene-selective units. Because the number of parahippocampal electrodes analyzed was small, it is possible that most or all were outside of the parahippocampal place area.

The response onset latencies of PHC neurons of ∼300 ms are similar to those reported previously (25) and are substantially shorter than those found in other MTL regions such as EC, hippocampus, and amygdala. In addition, this onset of neuronal firing occurs on average well after the peaking of the evoked LFP response at 243 ms, and even longer after the onset of selectivity in the LFP response at 153 ms. This difference in response latency between unit activity and LFP confirms the notion that neuronal action potentials represent the output activity of a neuron, whereas the LFP represents postsynaptic input activity and ongoing neuronal processing (20). These findings are in line with previous reports that LFP responses precede the onset of single-cell firing in the human MTL (26).

Our results demonstrate that a substantial proportion of PHC neurons respond not just to one scene but to multiple different scenes. This contrasts with neurons in other human MTL subregions, which, in this study and others (27, 25, 28), have been shown to exhibit much sparser responses. Because previous studies have shown that neurons in other MTL subregions show a high degree of invariance to specific concepts, it is natural to ask whether neurons in PHC, despite responding to a large number of scenes, might nonetheless encode the locations depicted in an invariant manner. Because we did not present the same scenes from multiple viewpoints, our data cannot rule out this possibility. However, fMRI studies have reported that activation in PHC is suppressed by repeated presentation of the same scene, but not when the same location is repeatedly presented from different viewpoints (29, 30). Assuming that fMRI adaptation measures underlying neural selectivity (ref. 31, but see ref. 32), these results suggest that most scene-selective PHC neurons do not respond invariantly to the same scene. Moreover, previous human single-neuron studies have shown that parahippocampal neurons exhibit little or no location selectivity

during virtual navigation (23, 33). Thus, available evidence indicates that scene representations in the PHC are neither as sparse nor as invariant as responses elsewhere in the MTL. It is, however, possible that PHC neurons possess some invariance to individual features, if not to individual locations.

Given the well-established role of the MTL in declarative memory (34) it is plausible to postulate that the PHC responses described in this study may have a relatively distributed code of space/location to provide contextual information to more specific items and associations coded in neurons higher up within the hierarchical structure of the MTL, in the hippocampus and EC. This notion is further supported by the fact that PHC is one of the primary inputs to EC and one of few cortical areas that project directly to the hippocampus (35). This denser distributed code may be necessary to rapidly form memories and contextual associations in novel environments. Given that scenes are defined by conjunctions of many features, it is implausible that the brain could possess sparse representations that are selectively and invariantly tuned to previously unseen environments. Moreover, rapidly forming such representations by integrating responses of neurons that respond sparsely to individual features would require a very high degree of connectivity. Instead, the brain may form sparse representations by integrating responses of neurons that respond to many features, but that are tuned along feature dimensions relevant to distinguishing scenes rather than to low-level features. In line with this hypothesis, the primary deficit observed in parahippocampal lesion patients is inability to navigate in novel environments (3, 4). This denser distributed representation may similarly be useful in forming novel contextual associations, a process in which the PHC has been shown to be involved (36–39).

The absence of a region-wide category preference in hippocampus and entorhinal cortex can in principle be attributed to their sparse and invariant representation (40). However, studies with a larger number of categories (compared with only landscapes, animals, and persons, as used here) might be necessary to rule out the presence of category preferences in these areas. In particular, functional imaging studies have shown that perirhinal and entorhinal cortex is preferentially activated by objects (41), a stimulus category underrepresented in this study.

Although our study gives insight into how individual neurons represent aspects of scenes, its retrospective nature makes it difficult to determine their exact nature. In accordance with neuroimaging studies, we show that single neurons in PHC responded more strongly to landscapes, outdoor scenes, images with spatial cues, images with a clearly recognizable background, images with greater depth, and larger real-world landmarks. However, our data are insufficient to determine what exactly about these images and features evokes a response in PHC neurons. fMRI studies suggest a wealth of parameters these neurons might encode, including scene category (42), spatial expanse (43, 6), texture (44), and clutter (11). Further studies will be necessary to determine the specific features to which individual PHC neurons are selective, and the role of these features in navigation and memory.

Materials and Methods

Subjects and Recordings. Twenty-four subjects undergoing treatment for pharmacologically intractable epilepsy were implanted with chronic depth electrodes (Fig. S9) to localize the epileptogenic focus for possible clinical resection (45). All studies conformed to the guidelines of the Medical Institutional Review Board of the University of California, Los Angeles, and the Institutional Review Board of California Institute of Technology. Informed written consent was obtained from each subject. Recordings were obtained from a bundle of nine microwires (eight high-impedance recording electrodes, one low-impedance reference) protruding from the end of each depth electrode. The voltage differences between the recording and reference electrodes were amplified, band-pass filtered from 1 to 9,000 Hz, and sampled at 28 kHz, using a Neuralynx Cheetah system. These recordings were stored digitally for further analysis. During each of 67 recording sessions, 23–190 images (median, 100; interquartile range, 95–124.5) were presented six times in pseudorandom order on a laptop computer, as described previously (13, 25, 46), while subjects sat comfortably in bed. Each image was presented for 1 s, at a random interspike interval (ISI) no less than 1.5 s, and subtended a visual angle of ∼5 degrees. To maintain attention, after image offset, subjects were ask to press the Y or N key on the keyboard to signal whether or not the presented image contained a face. Stimulus sets were composed of persons (grand average 75%), animals (9%), landscapes (13%), and stimuli from other categories (3%). Around 23% of the landscape pictures depicted contents the subjects had never seen before, whereas the others contained familiar landmarks and landscapes.

Image Classification. Before analysis, the authors categorized the types of all stimuli (747 pictures in total) and whether the pictures were indoors or outdoors. Stimuli were categorized as persons, animals, landscapes (with and without buildings; i.e., including landmarks), cartoons, food, abstract, or other. The latter four categories consisted of only a small number of stimuli and were excluded from further analysis. Indoor/outdoor discrimination was based on the visual properties of the image, and ambiguous cases were excluded from the analysis of this attribute. Spatial layout was defined as presence of elements relevant to navigation, such as topographical continuities in walls, room corners, and horizon lines. Because this distinction is sometimes ambiguous, one of the authors (S. Kornblith) and three additional individuals unrelated to this study also classified all images as possessing or not possessing spatial layout, blind to the neural responses to these stimuli. Three or more ratings agreed for 94% of stimuli. The remaining stimuli were excluded from analyses of spatial layout. Data for other classifications was collected using Amazon Mechanical Turk (*SI Materials and Methods*).

Population Response Plots. Because composition of the stimulus sets varied across patients and sessions, we used an automated, objective algorithm (*SI Materials and Methods*) to determine a set of 27 stimuli that had all been presented to the same 743 units to generate Fig. 2 and Figs. S1 and S8.

Spike Detection, Sorting, and Response Magnitude. After data collection, the signal recorded from the microelectrodes was band-pass filtered between 300 and 3,000 Hz and notch filtered at 2,000 Hz to remove artifacts produced by the clinical EEG system. The wave_clus software package was used to perform automated spike detection and sorting (47). To assess responsiveness, we calculated the average firing rate in the periods from 600 to 200 ms before stimulus onset (the baseline period) and from 200 to 600 ms after stimulus onset (the stimulus period). To measure the average population response of parahippocampal neurons, for each unit and stimulus, we computed a z-score-like normalized response as [mean(stimulus) − mean(baseline)]/standard deviation(baseline). We then averaged responses across units by stimulus category to yield the response magnitude values shown in Fig. 3 *B* and *C*. For the comparison between different MTL regions in Fig. 3*A*, we used wider periods from 1,000 to 0 ms before stimulus onset and from 0 to 1,000 ms after stimulus onset as the baseline and stimulus period, respectively, as average response latencies in EC and the hippocampus have been shown to be significantly longer than in PHC (25).

Response Onset Latencies. The latencies of PHC units responding to pictures from different categories were calculated using Poisson spike train analysis, as described in our earlier work (25). A detailed description of this procedure is given in the *SI Materials and Methods*. To compare latencies for two stimulus categories, we applied two different tests. First, we used an independent-sample unequal-variance *t* test to compare the groups of cells responding to each category. If a cell responded to both categories, then the median response latency for each category was used in each group. Second, we selected all cells that responded to both categories and ran a paired-sample Wilcoxon signed-rank test on these cells to compare response latencies.

Local Field Potentials. LFPs were band-pass filtered between 1 and 100 Hz and notch filtered at 60 Hz (4 Hz bandwidth) before downsampling to 365.5 Hz, using second-order Butterworth filters in the forward and reverse directions. To compute the average normalized LFP, we computed the trial-averaged response of each channel to stimuli with and without spatial layout and divided the result by the pooled SD of the 1-s interval before stimulus onset, and then averaged the per condition channel means across all channels. Population LFP selectivity to spatial layout was tested by averaging the response of each microelectrode across images with and without spatial layout and performing a paired *t* test at each of the 1,096 points in the interval from 1 s before stimulus onset to 2 s after. Selectivity in individual LFPs was determined by a two-sample *t* test, comparing the LFP amplitude for stimuli with and without spatial layout at each of the 365 points in the 1-s interval after stimulus onset. In both cases, to be considered significant, the LFP amplitude had to differ significantly between the two stimulus groups at a threshold of $P < 0.05$ after Holm-Bonferroni correction. To assess the spatial layout selectivity of individual units for comparison with the selectivity of the LFP and for computation of the spatial clustering statistic below, for each unit, we performed a Mann–Whitney *U* test on the firing rates during the interval from 200 to 600 ms after stimulus onset.

Proportion of Stimuli Eliciting a Response. The proportion of stimuli that elicited a response (PSER) was computed by dividing the number of stimuli eliciting a response according to the response criterion described earlier within a given category by the total number of images within the category. Because all sessions contained more images without spatial layout than with spatial layout, naive calculation of the PSER for images with and without spatial layout would lead to indices with different distributions, thus clouding interpretation. To make the indices directly comparable, for each cell, we computed the PSER for images with spatial layout and then randomly drew an equal number of images without spatial layout with replacement and computed a PSER for images without spatial layout based on this reduced set. Proportions of stimuli eliciting a response for matched numbers of nonscene stimuli and the null distribution, shown in Fig. S10, are based on 1,000,000 applications of this procedure. Because most cells did not respond to most stimuli, responses are rare events, and standard logistic regression is not applicable. Instead, we determined the conditional distribution of the common OR by convolving the corresponding hypergeometric distributions and found the corresponding confidence intervals by using a root solver (48). We then computed the mode of the conditional distribution. This procedure gives an estimate of the common OR, as well as exact CIs.

Analysis of Low-Level Features. We computed the response of the HMAX C1 layer to each stimulus in our stimulus set, using the Cortical Network Simulator package (49). Features were extracted from the original 160 × 160-pixel images presented at each subject at nine different scales, using the parameters described in Mutch and Lowe (50). After extracting the features, we trained a linear support vector machine on all but one stimulus and tested the remaining stimulus for each stimulus in our stimulus set. We used LIBLINEAR to train support vector machines (51), selected the regularization parameter C using 10-fold cross validation for each SVM trained, and inversely weighted training exemplars according to proportion in each category.

ACKNOWLEDGMENTS. We thank all our subjects for their participation; E. Behnke, T. Fields, E. Ho, V. Isiaka, E. Isham, K. Laird, N. Parikshak, and A. Postolova for technical assistance with the recordings; and D. Tsao for useful discussion. This research was supported by grants from the Volkswagen Foundation (Lichtenberg Program), the German Research Council (DFG MO930/4-1 and SFB 1089), the National Institute of Neurological Disorders and Stroke, the G. Harold and Leila Y. Mathers Foundation, the Gimbel Discovery Fund, the Dana Foundation, the Human Frontiers Science Program, and a National Science Foundation Graduate Research Fellowship (to S. Kornblith).

1. Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392(6676):598–601.
2. Bastin J, et al. (2013) Temporal components in the parahippocampal place area revealed by human intracerebral recordings. *J Neurosci* 33(24):10123–10131.
3. Aguirre GK, D'Esposito M (1999) Topographical disorientation: A synthesis and taxonomy. *Brain* 122(Pt 9):1613–1628.
4. Epstein R, Deyoe EA, Press DZ, Rosen AC, Kanwisher N (2001) Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cogn Neuropsychol* 18(6):481–508.
5. Mégevand P, et al. (2014) Seeing scenes: Topographic visual hallucinations evoked by direct electrical stimulation of the parahippocampal place area. *J Neurosci* 34(16): 5399–5405.
6. Harel A, Kravitz DJ, Baker CI (2013) Deconstructing visual scenes in cortex: Gradients of object and spatial layout information. *Cereb Cortex* 23(4):947–957.
7. 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.
8. Konkle T, Oliva A (2012) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74(6):1114–1124.
9. Mullally SL, Maguire EA (2011) A new role for the parahippocampal cortex in representing space. *J Neurosci* 31(20):7441–7449.
10. 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(2): 201–213.
11. Park S, Konkle T, Oliva A (2015) Parametric Coding of the Size and Clutter of Natural Scenes in the Human Brain. *Cereb Cortex* 25(7):1792–1805.
12. Logothetis NK (2008) What we can do and what we cannot do with fMRI. *Nature* 453(7197):869–878.
13. Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435(7045):1102–1107.
14. Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. *Science* 311(5761):670–674.
15. Kornblith S, Cheng X, Ohayon S, Tsao DY (2013) A network for scene processing in the macaque temporal lobe. *Neuron* 79(4):766–781.
16. Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30:123–152.
17. Rolls ET (2013) The mechanisms for pattern completion and pattern separation in the hippocampus. *Front Syst Neurosci* 7:74.
18. Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2(11):1019–1025.
19. Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. *Nat Rev Neurosci* 13(6):407–420.
20. Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. *Physiol Rev* 65(1): 37–100.
21. Viskontas IV, Quiroga RQ, Fried I (2009) Human medial temporal lobe neurons respond preferentially to personally relevant images. *Proc Natl Acad Sci USA* 106(50): 21329–21334.
22. Valdez AB, et al. (2015) Distributed representation of visual objects by single neurons in the human brain. *J Neurosci* 35(13):5180–5186.
23. Ekstrom AD, et al. (2003) Cellular networks underlying human spatial navigation. *Nature* 425(6954):184–188.
24. Kraskov A, Quiroga RQ, Reddy L, Fried I, Koch C (2007) Local field potentials and spikes in the human medial temporal lobe are selective to image category. *J Cogn Neurosci* 19(3):479–492.
25. Mormann F, et al. (2008) Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J Neurosci* 28(36):8865–8872.
26. Rey HG, Fried I, Quian Quiroga R (2014) Timing of single-neuron and local field potential responses in the human medial temporal lobe. *Curr Biol* 24(3):299–304.
27. Waydo S, Kraskov A, Quian Quiroga R, Fried I, Koch C (2006) Sparse representation in the human medial temporal lobe. *J Neurosci* 26(40):10232–10234.
28. Ison MJ, et al. (2011) Selectivity of pyramidal cells and interneurons in the human medial temporal lobe. *J Neurophysiol* 106(4):1713–1721.
29. Epstein R, Graham KS, Downing PE (2003) Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37(5):865–876.
30. Park S, Chun MM (2009) Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage* 47(4): 1747–1756.
31. Grill-Spector K, Malach R (2001) fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107(1-3):293–321.
32. Sawamura H, Orban GA, Vogels R (2006) Selectivity of neuronal adaptation does not match response selectivity: A single-cell study of the FMRI adaptation paradigm. *Neuron* 49(2):307–318.
33. Miller JF, et al. (2013) Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science* 342(6162):1111–1114.
34. Squire LR, Stark CEL, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27: 279–306.
35. Suzuki WA, Amaral DG (1990) Cortical inputs to the CA1 field of the monkey hippocampus originate from the perirhinal and parahippocampal cortex but not from area TE. *Neurosci Lett* 115(1):43–48.
36. Aminoff E, Gronau N, Bar M (2007) The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb Cortex* 17(7):1493–1503.
37. Buffalo EA, Bellgowan PSF, Martin A (2006) Distinct roles for medial temporal lobe structures in memory for objects and their locations. *Learn Mem* 13(5):638–643.
38. Ison MJ, Quian Quiroga R, Fried I (2015) Rapid encoding of new memories by individual neurons in the human brain. *Neuron* 87(1):220–230.
39. Law JR, et al. (2005) Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *J Neurosci* 25(24):5720–5729.
40. Quian Quiroga R, Kraskov A, Koch C, Fried I (2009) Explicit encoding of multimodal percepts by single neurons in the human brain. *Curr Biol* 19(15):1308–1313.
41. Litman L, Awipi T, Davachi L (2009) Category-specificity in the human medial temporal lobe cortex. *Hippocampus* 19(3):308–319.
42. 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(34):10573–10581.
43. Kravitz DJ, Peng CS, Baker CI (2011) Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *J Neurosci* 31(20):7322–7333.
44. Cant JS, Goodale MA (2011) Scratching beneath the surface: New insights into the functional properties of the lateral occipital area and parahippocampal place area. *J Neurosci* 31(22):8248–8258.
45. Fried I, MacDonald KA, Wilson CL (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18(5):753–765.
46. Mormann F, et al. (2011) A category-specific response to animals in the right human amygdala. *Nat Neurosci* 14(10):1247–1249.
47. Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput* 16(8):1661–1687.
48. Vollset SE, Hirji KF, Elashoff RM (1991) Fast computation of exact confidence limits for the common odds ratio in a series of 2 × 2 tables. *J Am Stat Assoc* 86(414):404–409.
49. Mutch J, Knoblich U, Poggio T (2010) *CNS: A GPU-based framework for simulating cortically-organized networks* (MIT CSAIL). Available at dspace.mit.edu/handle/1721.1/51839. Accessed May 17, 2013.
50. Mutch J, Lowe DG (2006) Multiclass object recognition with sparse, localized features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE, Washington, DC), pp 11–18.
51. Fan R-E, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J (2008) LIBLINEAR: A library for large linear classification. *J Mach Learn Res* 9:1871–1874.
52. Vinterbo SA (2002) *Maximum k-Intersection, Edge Labeled Multigraph Max Capacity k-Path, and Max Factor k-gcd Are All NP-Hard* (Decision Systems Group, Brigham and Women's Hospital, Boston), Decision System Group Technical Report 2002/12.