scientific data

OPEN

DATA DESCRIPTOR

A human single-neuron dataset for object recognition

Runnan Cao[1]✉, Peter Brunner[2], Nicholas J. Brandmeir[3], Jon T. Willie[2] & Shuo Wang[1]✉

Object recognition is fundamental to how we interact with and interpret the world around us. The human amygdala and hippocampus play a key role in object recognition, contributing to both the encoding and retrieval of visual information. Here, we recorded single-neuron activity from the human amygdala and hippocampus when neurosurgical epilepsy patients performed a one-back task using naturalistic object stimuli. We employed two sets of naturalistic object images from leading datasets extensively used in primate neural recordings and computer vision models: we recorded 1204 neurons using the ImageNet stimuli, which included broader object categories (10 different images per category for 50 categories), and we recorded 512 neurons using the Microsoft COCO stimuli, which featured a higher number of images per category (50 different images per category for 10 categories). Together, our extensive dataset, offering the highest spatial and temporal resolution currently available in humans, will not only facilitate a comprehensive analysis of the neural correlates of object recognition but also provide valuable opportunities for training and validating computational models.

Background & Summary

Object recognition is a crucial topic in both cognitive and computational neuroscience, as well as in computer vision. For years, neurons in the human amygdala and hippocampus have been associated with category-specific encoding of visual objects and faces[1–3], structured in a hierarchical fashion[4]. These category-selective may form highly selective (sparse) representations of objects[5–8] and are regarded as the building blocks of declarative memory[7–10]. The human amygdala and hippocampus have thus been recognized for their critical roles in object recognition and encoding. However, previous studies have typically used a limited number of object categories ($n = 9$ for[1], $n = 14$ for[11], $n = 14$ for[3], and $n = 10$ for[4]) to probe neural responses in the human amygdala and hippocampus. Furthermore, previous studies used a small number of single-units or multi-units for modeling object recognition in non-human primates ($n = 472$ for[12] using computer-generated stimuli, and $n = 107$ for[13] and $n = 32$ for[14] using naturalistic images). This limitation has hindered a comprehensive understanding of how these brain regions encode a wide variety of naturalistic objects. In particular, there are no publicly available datasets at the single-neuron level in the human amygdala and hippocampus that feature a large number of naturalistic images.

To address this gap, we present a large human single-neuron dataset to the research community. Our dataset offers several significant advantages. First, single-neuron recordings provide the highest spatial and temporal resolution currently available in the human brain. This level of detail is critical for training and validating computational models, bridging a key gap between primate neurophysiology and human neuroimaging studies. Second, we utilized naturalistic object images from two leading image datasets, ImageNet[15] and Microsoft COCO[16]. These datasets have been extensively used in primate neural recordings[13,14,17] and computer vision models[18–21], ensuring our data is well integrated with existing literature. Importantly, we not only recorded from a large number of neurons across two stimulus sets (1204 neurons for ImageNet and 512 neurons for Microsoft COCO), but also provided diversity in the stimuli: we included broader object categories (10 different images per category for 50 categories) from the ImageNet dataset and a higher number of images per category (50 different images per category for 10 categories) from the Microsoft COCO dataset. Computational models can be constructed based on single-units, multi-units, and neural populations; and thus the sample size of our current dataset is adequate to test computational models. Together, this extensive data contribution is expected to facilitate not only a deeper understanding of the neural mechanisms underlying object recognition but also the development of more accurate computational models.

[1]Department of Radiology, Washington University in St. Louis, St. Louis, MO, 63110, USA. [2]Department of Neurosurgery, Washington University in St. Louis, St. Louis, MO, 63110, USA. [3]Department of Neurosurgery, West Virginia University, Morgantown, WV, 26506, USA. ✉e-mail: r.cao@wustl.edu; shuowang@wustl.edu

Methods

Patients. We recruited six neurosurgical patients (2 males, aged 26-53; see Table 1 for detailed demographics) from the West Virginia University (WVU), nine patients (3 males, aged 24-45) from the Washington University in St. Louis (WUSTL), and one patient (female, age 45) from the University of Utah (Utah). All participants provided written informed consent to participate in the experiments (i.e., performing cognitive tasks) for single-neuron recordings and to share de-identified data with the research community. The protocols were approved by the Institutional Review Boards of WVU (protocol #: 1709745061), WUSTL (protocol #: 202201019), and the University of Utah (protocol #: 00144045).

All patients performed the one-back task with the ImageNet stimuli, yielding 1204 neurons (874 neurons had an overall mean firing rate greater than 0.15 Hz; Table 1). Five patients from WVU and eight patients from WUSTL performed the one-back task with the Microsoft COCO stimuli, yielding 512 neurons (453 neurons had an overall mean firing rate greater than 0.15 Hz; Table 1).

Stimuli. We used two sets of naturalistic stimuli:

For stimuli from the ImageNet dataset[15], we selected 50 categories of objects with 10 images for each object category. The object categories included battery, arachnid, bark, beverage, board, bread, brier, building, car, cat, collection, crustacean, dainty, dog, electrical device, electronic device, equipment, fare, fern, fish, flower, frog, fruit, fungus, furniture, game bird, gymnast, herb, hole, insect, light, man clothing, moped, musical instrument, needlework, nest, plate, reptile, ridge, rock, rodent, star, sugar maple, support, tool, utensil, vegetable, vessel, weapon, and young mammal (see Fig. 1a for examples).

For stimuli from the Microsoft COCO dataset[16], we selected 10 categories of objects with 50 images for each object category. The object categories included airplane, apple, bear, bird, car, chair, dog, elephant, person, and zebra (see Fig. 1b for examples).

While a key novelty of our current dataset lies in testing naturalistic object categories that have not been commonly explored in previous studies, we recognize the lack of data collection involving the same neurons under non-naturalistic image presentations[22,23]. As a result, we were unable to directly compare neural responses to naturalistic versus non-naturalistic stimuli within the same neural population, which could have provided valuable insights into the generalizability of our findings.

Experimental procedure. We used a one-back task for both sets of stimuli. In each trial, a single object image was presented at the center of the screen for a fixed duration of 1 s, with uniformly jittered inter-stimulus-interval (ISI) of 0.5-0.75 s. Each image subtended a visual angle of approximately 10°. Patients pressed a button if the present image was *identical* to the immediately previous image. 10% of trials were one-back repetitions. Each image was shown once unless repeated in one-back trials; and we excluded responses from the one-back trials to have an equal number of responses for each image because they may convey information related to decision-making or movement-related signals. This task kept patients attending to the images but avoided potential biases from focusing on a particular image feature[24]. It is worth noting that each neuron was recorded for all images, and thus the number of neurons recorded for each image was the same.

Patients performed well on this task (accuracy = 85% ± 18.44% [mean ± SD across all recorded sessions]; reaction time = 479 ± 30 ms [mean ± SD across all recorded sessions]). We did not impose any exclusion criteria based on patients' psychometric scores.

Electrophysiology. We recorded from implanted depth electrodes in the amygdala and hippocampus from patients with pharmacologically intractable epilepsy. At each site, we recorded from eight 40 μm microwires (Ad-Tech: WB09R-SP00X-014) inserted into a clinical electrode as described previously[25,26]. Recording locations (Fig. 2 and Table 2) were estimated based on post-operative CT scans (WVU: orientation = axial, slice number = 195, 0.485 × 485 mm in-plane resolution, 1 mm thick, tube current = 320 mA, matrix size = 512 × 512; WUSTL: orientation = axial, slice number = 192, 0.547 × 0.547 mm in-plane resolution, 1 mm thick, tube current = 245 mA, matrix size = 512 × 512; Utah: orientation = axial, slice number = 194, 0.453 × 0.453 mm in-plane resolution, 1 mm thick, tube current = 171 mA, matrix size = 512 × 512) that were performed after implantation of the electrodes. These scans were registered to pre-operative T1-weighted MRI scans (WVU: orientation = axial, slice number = 156, 0.976 × 0.976 mm in-plane, 1 mm thick, TR = 8248 ms, TE = 3.548 ms, matrix size = 256 × 256, flip angle = 12°; WUSTL: orientation = axial, slice number = 192, 0.972 × 0.972 mm in-plane, 1 mm thick, TR = 1700 ms, TE = 2.39 ms, matrix size = 288 × 288, flip angle = 9°; Utah: orientation = axial, slice number = 190, 0.972 × 0.972 mm in-plane, 1 mm thick, TR = 1900 ms, TE = 4.92 ms, matrix size = 288 × 216, flip angle = 25°;) to allow patient-specific localization. We used the function 'align_epi_anat.py' in AFNI to align the CT images to the MRI T1 images. The T1 images were then normalized to the MNI152_2009_template_SSW standard template to obtain a transformation matrix. We further applied the transformation matrix to the CT images to align them to the MNI space. Finally, we manually located each electrode position and derived the MNI coordinates for each recording location.

At each site, we recorded from eight 40 μm microwires inserted into a clinical electrode as described previously[25,26]. Bipolar wide-band recordings (0.1-9000 Hz), using one of the eight microwires as reference, were sampled at 32 kHz and stored continuously for off-line analysis with a Neuralynx (WVU) or Blackrock (WUSTL, Utah) system[27]. The raw signal was filtered with a zero-phase lag 300-3000 Hz bandpass filter and spikes were sorted using a semi-automatic template matching algorithm as described previously[28].

We employed an offline spike sorting procedure[28]. Specifically, the signal was continuously bandpass filtered from 300 to 3000 Hz using a zero-phase lag filter. Spikes were detected by applying a threshold to a local energy signal, which was continuously derived from the raw filtered signal. For example, spikes were detected

ID	Sex	Age	Race	Epilepsy diagnosis	Number of Amygdala Neurons			Number of Hippocampus Neurons		
					Total	LA	RA	Total	LH	RH
WVU14	M	26	Caucasian	Bilateral amygdala and hippocampus	16	12	4	9	4	5
					11	9	2	8	5	3
WVU16	F	29	Caucasian	Right temporal neocortex	34	18	16	37	24	13
					38	24	14	47	25	22
					30	17	13	41	27	14
WVU18	F	53	Caucasian	Left temporal	7	7	0	1	0	1
					54	26	28	8	3	5
					30	14	16	6	2	4
					41	19	22	8	0	8
WVU19	M	49	Caucasian	Right mesial temporal	1	0	1	18	10	8
					1	0	1	25	11	14
WVU20	F	31	Caucasian	Left temporal	38	34	4	52	52	0
					34	31	3	25	25	0
					34	31	3	27	27	0
					36	34	2	25	25	0
					34	32	2	32	32	0
WVU21	F	27	Caucasian	Left mesial temporal	31	31	0	0	0	0
					32	32	0	0	0	0
					31	31	0	0	0	0
UIC202214	F	45	Caucasian	Right mesial temporal	10	10	0	7	1	6
BJH024	F	38	Caucasian	Left temporal	13	0	13	8	0	8
					12	0	12	17	0	17
BJH025	F	45	Caucasian	Right mesial temporal	9	9	0	0	0	0
					10	10	0	0	0	0
BJH026	M	40	Caucasian	Left mesial temporal	19	19	0	13	13	0
					27	27	0	11	11	0
					22	22	0	14	14	0
					23	23	0	9	9	0
BJH027	M	37	Caucasian	Right temporal	0	0	0	18	18	0
					0	0	0	14	14	0
					0	0	0	21	21	0
					0	0	0	18	18	0
BJH028	F	36	Caucasian	Left frontal SMA, cingulate regions	21	21	0	0	0	0
					24	24	0	0	0	0
BJH029	F	34	Caucasian	Left amygdala and hippocampus	0	0	0	8	8	0
					0	0	0	5	5	0
					0	0	0	8	8	0
					0	0	0	6	6	0
BJH032	M	27	Caucasian/American Indian	Bilateral amygdala and hippocampus	0	0	0	6	0	6
					0	0	0	4	0	4
					0	0	0	8	0	8
					0	0	0	4	0	4
BJH033	F	24	Caucasian/American Indian	Right hippocampus	11	11	0	1	1	0
					3	3	0	0	0	0
					9	9	0	0	0	0
					12	12	0	0	0	0
Sum					758	602	156	569	419	150

Table 1. Patients and sessions. Each row of neurons represents a separate recording session using the ImageNet stimuli. Each session was recorded on a separate day. Total: all neurons recorded from an area. Left: neurons that were recorded from the left side of an area and had a firing rate greater than 0.15 Hz. Right: neurons that were recorded from the right side of an area and had a firing rate greater than 0.15 Hz. WVU: Patients from the West Virginia University. UIC: Patient from the University of Utah. BJH: Patients from the Washington University in St. Louis.

when the signal was greater than 5 times of the standard deviation of the noise. Once a spike was detected and properly re-aligned, a distance metric was calculated to measure the spike's distance from all known clusters at that moment. If the minimum distance was less than a threshold, the spike was assigned to that cluster. If not, a new cluster was created, and the spike was assigned to it. These thresholds were automatically and continuously

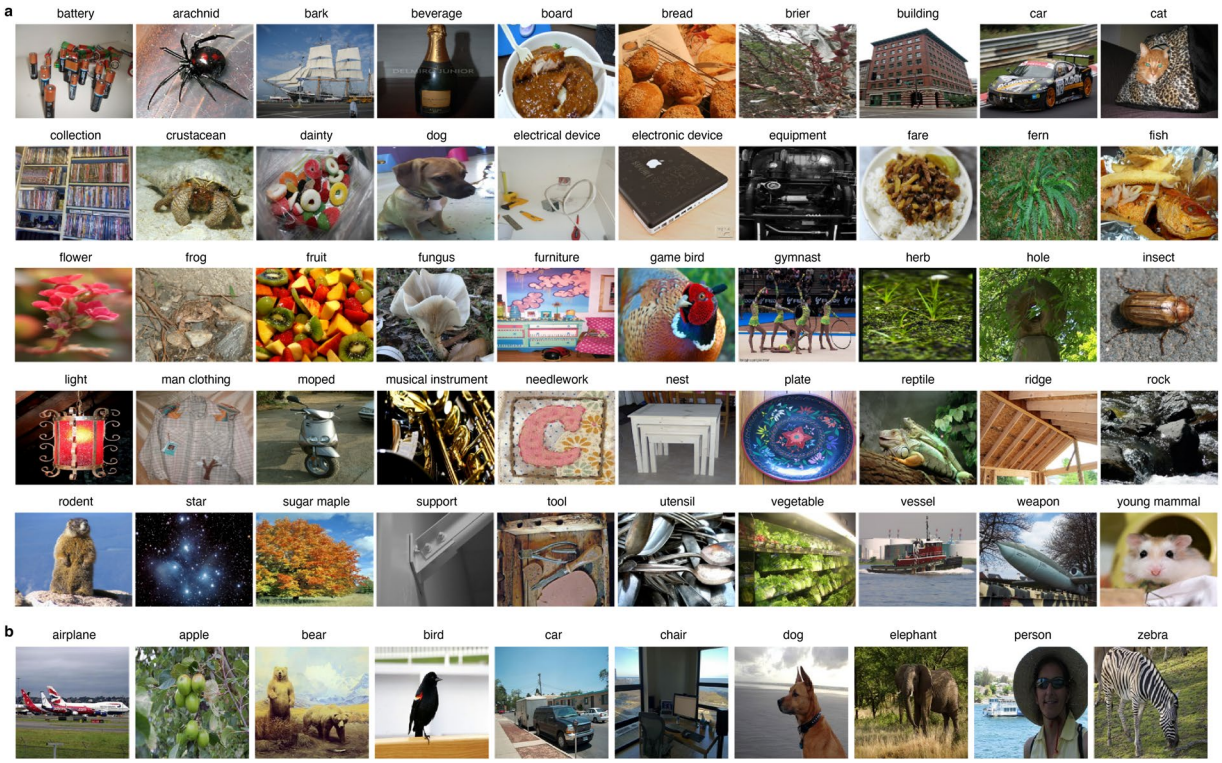

Fig. 1 Example stimuli. **(a)** ImageNet. **(b)** Microsoft COCO. Category labels are shown above each example.

determined based on the noise characteristics of the raw filtered signal. After assigning a spike to a cluster, the cluster's mean waveform was updated to reflect the new spike. This process allowed for tracking electrode movements and accommodating short-term changes due to bursts. If clusters overlap after updating the mean waveform, they were merged, and the spikes were reassigned. Regularly, the statistical evaluation criteria (such as the ISI distribution, power spectrum, and autocorrelation), along with projection tests for each cluster pair, were calculated. This continuous evaluation helped in discarding noise and multi-unit activity. It is worth noting that this sorting algorithm can handle all waveforms, including complex spike waveforms, and broad-spiking and narrow-spiking waveforms.

We manually examined the data during spike sorting and ensured that all neurons had a stable response throughout the task. It is also worth noting that our dataset includes neurons from epileptic seizure foci (see Table 1 for details). To test the potential influence of interictal activity, we selected category-selective neurons after excluding neurons ($n = 337$ for the ImageNet dataset and $n = 286$ for the COCO dataset) from the seizure onset areas. This revealed a comparable percentage of category-selective neurons for both the ImageNet (12.76%) and COCO (14.63%) datasets.

Spike sorting quality metrics. We quantitatively assessed the separation of clusters using a projection test[29] and the isolation distance metric[28]. First, we conducted the projection test to quantify the overlap between any pair of clusters identified from a single channel. For a given pair of clusters, the difference of every spike and the center of the cluster it is assigned to was projected onto the vector that connects the two centers of the clusters. The distance between these two centers was then considered as the measure of separation between the two clusters (Fig. 3f). Isolation distance (Fig. 3g) was calculated based on the methods described in[30,31]. If a cluster contains n_C cluster spikes, the isolation distance of the cluster is the D^2 value of the $n_C{}^{\text{th}}$ closest noise spike. Isolation distance is therefore the radius of the smallest ellipsoid from the cluster center containing all of the cluster spikes and an equal number of noise spikes. As such, isolation distance estimates how distant the cluster spikes are from the other spikes recorded on the same electrode. Isolation distance is not defined for cases in which the number of cluster spikes is greater than the number of noise spikes. Consistent with our previous studies[3,32–37], only single units with an average firing rate of at least 0.15 Hz throughout the entire task were considered.

Cross-session similarity of response pattern. We compared the response pattern similarity for neurons recorded across different sessions from the same participants with that of neurons recorded from different participants. For the within-participant condition, we estimated the Spearman correlation across the mean response of the 500 object stimuli between the first 2 sessions for each participant that had more than one session. Similarly, we estimated the correlation between any possible pairs of sessions from different participants for the between-participant condition. To balance the number of pairs for between-participant sessions ($n = 210$) and within-participant sessions ($n = 10$), we performed bootstrapping with subsampling (1000 times) for the between-participant condition (each time sampling 10 pairs).

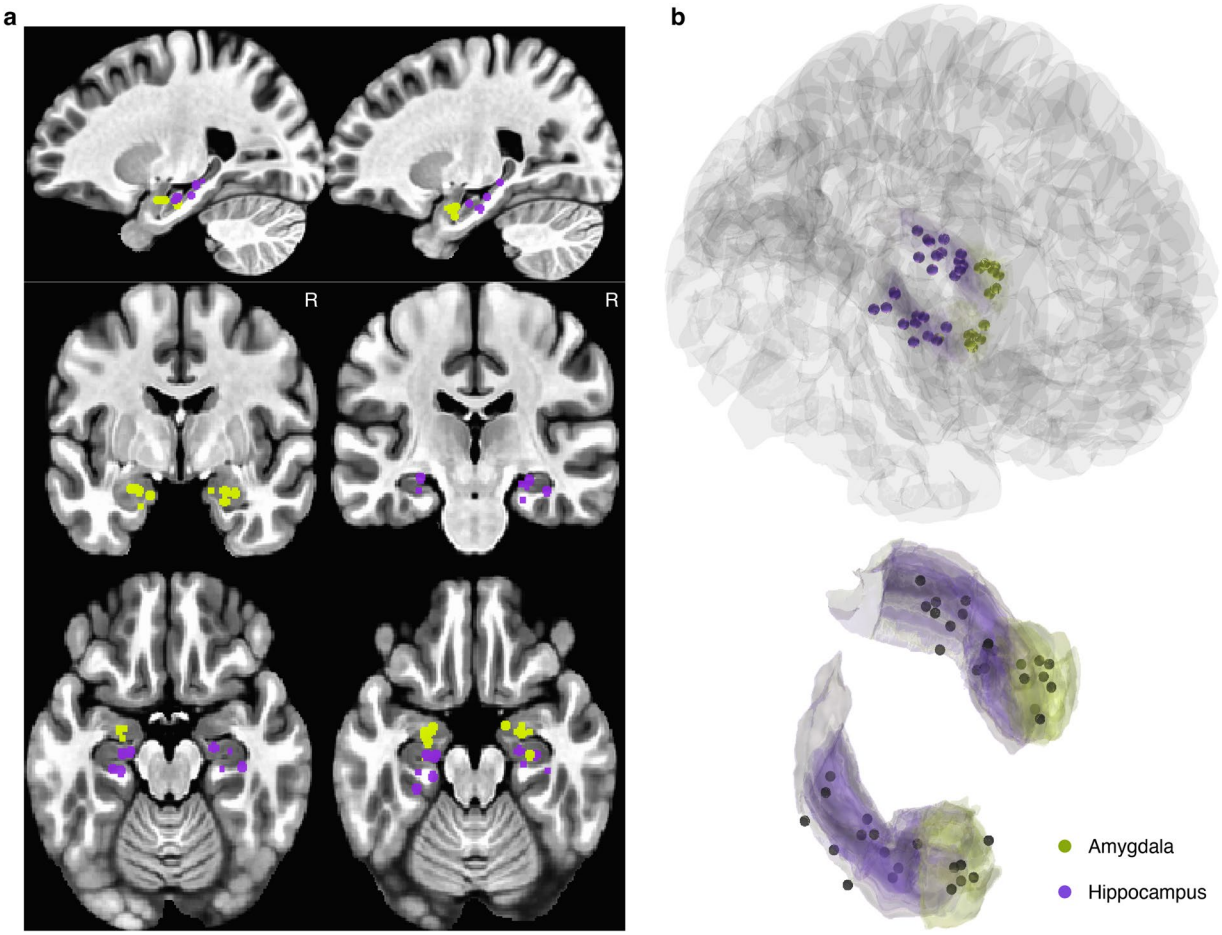

Fig. 2 Electrode locations in the MNI space. The recording sites of each patient were estimated according to the pre-implantation T1-weighted MRI and post-implantation CT. Illustration of electrodes from all participants. (**a**) Volumetric view. (**b**) Inflated view.

Depth of selectivity (DOS) index. To summarize the response of category-selective neurons, we quantified the depth of selectivity (DOS) for each neuron: $DOS = \frac{n - \left(\sum_{j=1}^{n} r_j\right) / r_{\max}}{n-1}$, where n is the number of categories ($n = 50$), r_j is the mean firing rate to category j, and r_{max} is the maximal mean firing rate across all categories. DOS varies from 0 to 1, with 0 indicating an equal response to all categories and 1 exclusive response to one category, but not to any of the other categories. Thus, a DOS value of 1 is equal to maximal sparseness of category coding. The DOS index has been used in many prior studies investigating visual selectivity[3,38,39].

Differential latency analysis. We conducted a spike train analysis to estimate the differential response latency of category-selective neurons in the amygdala and hippocampus using an established procedure[3]. We first binned spike trains into 1-ms bins and computed the cumulative sum. We then averaged the normalized cumulative sums for encoded objects and non-encoded objects, respectively, and compared, at every point of time, whether the normalized cumulative sums were different between encoded objects and non-encoded objects (two-tailed paired t-test; $P < 0.005$; FDR-corrected). The first point of time of the significant cluster (cluster size > 20 time points) was used as the estimate of the differential latency. Note that this method is not sensitive to differences in baseline firing rates between neurons because the latency estimate is pairwise for each neuron individually. To ensure a fair comparison between groups, we bootstrapped and subsampled category-selective neurons from the amygdala ($n = 84$ in total) 1000 times to match the number of category-selective neurons in the hippocampus ($n = 53$). To assess statistical significance, we compared the latency of the hippocampus with the bootstrapped distribution of the amygdala to obtain the p-value.

Data Records

The data is available at OSF[40], which consists of stimuli used in the experiments, behavior, events, and neuronal spikes. All data (except stimuli) are stored in MATLAB '.mat' files, as well as in other format ('.csv' or '.txt') for the usage on other platforms. Different parts of the data are organized in different directories and described in detail below. A detailed description of variables and usage can be found in 'readme.rtf' in each data directory. While the scope of the current dataset focuses on single-neuron activity, the raw time series data is available at the same OSF storage.

ID	Left Amygdala	Right Amygdala	Left Anterior Hippocampus	Right Anterior Hippocampus	Left Posterior Hippocampus	Right Posterior Hippocampus
WVU14	20, 6, −28	−23, 2, −22	20, 18, −22	−36, 23, −18	29, 32, −21	−35, 37, −7
WVU16	23, 10, −20	−22, 5, −21	27, 23, −12	−27, 21, −13	26, 32, −9	−27, 31, −5
WVU18	25, 7, −20	−23, 6, −21	19, 16, −19	−33, 27, −13	n.a.	−33, 33, −7
WVU19	n.a.	−22, 5, −26	28, 24, −18	−26, 24, −16	n.a.	−22, 32, −7
WVU20	25, 5, −20	n.a.	21, 16, −14	−23, 24, −15	24, 28, −11	n.a.
WVU21	Unknown	n.a.	Unknown	n.a.	Unknown	n.a.
UIC202214	15 3 −23	−24 20 −22	21 26 −21	n.a.	n.a.	n.a.
BJH024	n.a.	−27 5 −22	n.a.	−30 16 −20	n.a.	n.a.
BJH025	25 17 −22	−27 16 −21	n.a.	n.a.	n.a.	n.a.
BJH026	21 3 −21	n.a.	24 16 −20	n.a.	n.a.	n.a.
BJH027	n.a.	n.a.	n.a.	n.a.	24 25 −17	n.a.
BJH028	17 7 −15	n.a.	n.a.	n.a.	n.a.	n.a.
BJH029	n.a.	n.a.	n.a.	n.a.	21 27 −10	n.a.
BJH032	n.a.	−15 2 −21	n.a.	−22 14 −19	n.a.	n.a.
BJH033	22 5 −20	n.a.	23 17 −18	n.a.	n.a.	n.a.

Table 2. Estimated recording locations for each patient. Shown are radiological coordinates (also known as RAI: right-anterior-inferior) in the MNI space. The recording sites were confirmed by the neurosurgeon. n.a.: not available (no electrodes implanted at this location). Unknown: imaging data was unavailable for patient WVU21, preventing us from determining the precise coordinates of each electrode for this patient.

Stimuli. The stimuli used in the present study are located in the '/Stimuli/ImageNet/' and '/Stimuli/COCO/' directories. The name of the category for each object image is stored in the variable 'Cats' in '/Stimuli/Categories_ImageNet.mat' and '/Stimuli/Categories_COCO.mat'. We also present the object category names in '/Stimuli/Categories_ImageNet.csv' and '/Stimuli/Categories_COCO.csv' for the ease of usage on other platforms. The mapping between the filename of the object image and the index in the behavior data, as well as the order in the neural data, is stored in the variable 'vImg'.

Behavior. Behavioral data for all sessions is stored in the variable 'beh' in 'behaviorData/*_behavior.mat' (* indicates the task). The corresponding session index is stored in the variable 'session'. The timing parameters for each trial can be found in the variable 'iT'. Stimulus presentation order is in the variable 'code', which contains the index for each image. Indices for the repeated (i.e., one-back) trials are stored in the variable 'back_id'. If the patient responded (i.e., pressed the space bar) in the one-back trial, the value is 1 for this trial in the variable 'vResp'; otherwise, the value is 0. Response time is stored in the variable 'RT'.

Events. Timestamps of events (i.e., stimulus onsets) are stored in the '/Events Files/' directory, in both '.mat' and '.txt' formats. There is one file for each session. The variable 'periods' contains 3 columns. The first column denotes trial indices, the second column denotes the timestamps (in μs) corresponding to 500 ms before stimulus onset, and the third column denotes timestamps corresponding to 1500 ms after stimulus onset.

Neuronal spikes. Neuronal spikes were sorted and extracted using an established procedure[28]. We combined neurons from all sessions into a single file ('/Single Neuron/Sorted Data/ImageNet_Spikes.mat' and '/Single Neuron/Sorted Data/CoCo_Spikes.mat'). For each neuron, the timestamps (in μs), session ID, channel ID, cluster ID (i.e., randomly assigned neuron ID within a channel), and recording brain area are stored in the variables 'timestampsOfCellAll', 'vCell', 'vCh', 'vClusterID', and 'areaCell', respectively (note that these variables were all matched). Time stamps of each cell were also stored in '.txt' files under the directories '/Single Neuron/Sorted Data/ImageNet_TimeStamps/' and '/Single Neuron/Sorted Data/CoCo_TimeStamps/'. We present three variables that describe spike sorting quality. The variable 'IsolDist' shows the isolation distance value for each cluster (i.e., a single neuron; see Methods). The variable 'statsSNR' contains 6 columns: 1—session ID, 2—channel ID, 3—cluster ID, 4—the average signal-to-noise ratio (SNR), 5—inter-spike intervals (ISI) that are below 3 ms, and 6—peak SNR. The variable 'statsProjAll' contains 5 columns: 1—session ID, 2—channel ID, 3 and 4—IDs for a cluster pair, and 5—pairwise distance between two clusters.

Technical Validation

Quality metrics for spike sorting. We recorded from the human amygdala and hippocampus (Fig. 2). 874 units had an overall firing rate greater than 0.15 Hz and we restricted our analysis to this subset of units. We first assessed the quality of our recordings and showed that the quality was comparable to previous studies[3,32,33,37,41]. Specifically, the number of units recorded per wire was 2.67 ± 1.51 (mean ± SD; in wires with at least one unit; Fig. 3a). The mean firing rate was 2.17 ± 3.50 Hz (Fig. 3b). The percentage of inter-spike interval (ISI) below 3 ms was 0.46% ± 0.67% (Fig. 3c). The ratio between the peak amplitude of the mean waveform of each cluster and the standard deviation of the noise (peak signal-to-noise ratio [SNR]) was 6.62 ± 4.54 (Fig. 3d). The mean SNR was 2.57 ± 1.36 (Fig. 3e). The pairwise distance (projection test) between all possible pairs of units on all wires with more than one cluster was calculated, yielding a mean of 4.44 ± 9.13 (Fig. 3f). The median isolation distance was 18.88 (Fig. 3g).

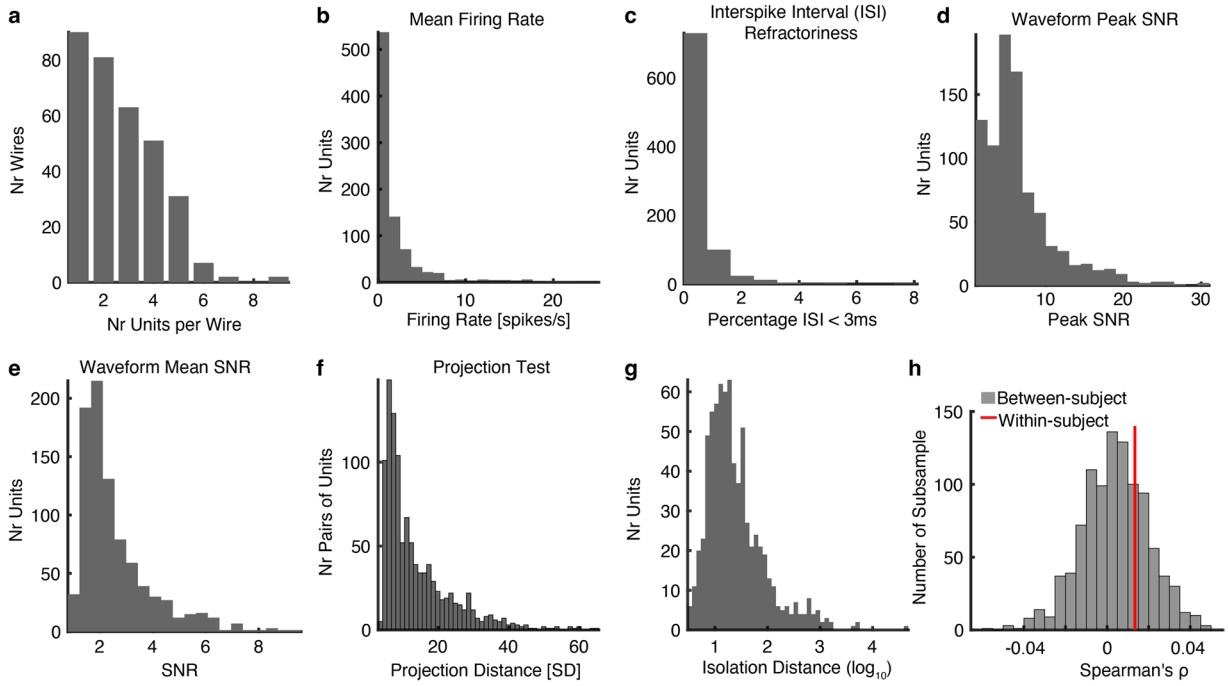

Fig. 3 Assessment of recording and spike sorting quality. **(a)** Histogram of the number of units identified on each active wire (only wires with at least one unit identified are counted). The average yield per wire with at least one unit was 2.67 ± 1.51 (mean $\pm$ SD). **(b)** Histogram of mean firing rates. **(c)** Histogram of proportion of inter-spike intervals (ISIs) which are shorter than 3 ms. The large majority of clusters had less than 0.5% of such short ISIs. **(d)** Histogram of the signal-to-noise ratio (SNR) of the mean waveform peak of each unit. **(e)** Histogram of the SNR of the entire waveform of all units. **(f)** Pairwise distance between all possible pairs of units on all wires where more than 1 cluster was isolated. Distances are expressed in units of standard deviation (SD) after normalizing the data such that the distribution of waveforms around their mean is equal to 1. **(g)** Isolation distance of all units for which this metric was defined ($n = 874$, median $= 3.50$). **(h)** Cross-session similarity within versus between participants.

It is worth noting that neurons from different sessions of the same patient were not recorded on the same day. While the neurons from different sessions are thus likely to be different, it is important to note that the same neuron can persist across days in chronic recordings[42]. To quantitatively address this question, we conducted two control analysis. First, we compared the response (i.e., firing rate) pattern similarity across 2 sessions within each participant versus those from different participants (see Methods). This revealed that the correlation between sessions was very small for both conditions and the within-participant cross-session similarity was not significantly greater than that for the between-participant condition (Fig. 3h). Second, we measured the peristimulus time histogram (PSTH) similarity using Spearman correlation across clusters within the same channels. Only channels that were recorded in the first two sessions from the same participant that had identified clusters were included ($n = 110$). Consistently, the similarity score was not significantly above zero (one-sample right-tailed t-test against 0, $t(109) = 1.08$, $P = 0.14$). Together, our results suggest that the neurons recorded from the same channels across different sessions (days) may not be the same.

Summary of category-selective neurons. To select *category-selective neurons*, we first used a one-way ANOVA to identify neurons with a significantly unequal response to different object categories ($P < 0.05$) in a window 250–1250 ms following stimulus onset. We next imposed an additional criterion to identify the selective categories: the neural response of a category was 1.5 standard deviations (SD) above the mean of neural responses from all categories. If a neuron satisfied both criteria, it was defined as a *category-selective neuron*. We identified category-selective neurons from each stimulus set. Specifically,

In the ImageNet dataset, we found 144 category-selective neurons (16.48%, binomial $P < 10^{-20}$; see Fig. 4a,b for examples). The most preferred object categories were *arachnid* and *needlework* (Fig. 4c). The depth of selectivity (DOS)[3,38,39] index was 0.53 ± 0.16 across neurons (Fig. 4d). Category-selective neurons had a significantly higher DOS index than non-category-selective neurons ($t(872) = 4.41$, $P = 4.48 \times 10^{-5}$, $d = 0.40$, 95% CI = [0.04, 0.09]).

In the Microsoft COCO dataset, we found 50 category-selective neurons (11.04%, binomial $P = 8.22 \times 10^{-8}$; see Fig. 5a,b for examples). The most preferred object categories were *person* and *car* (Fig. 5c). The DOS index was 0.29 ± 0.15 across neurons (Fig. 5d). Category-selective neurons had a significantly higher DOS index than non-category-selective neurons ($t(451) = 8.81$, $P = 2.63 \times 10^{-17}$, $d = 1.32$, 95% CI = [0.15, 0.23]).

To characterize the temporal profile of category-selective neurons across different areas, we estimated the onset latency of category selectivity for amygdala and hippocampal neurons separately. This was achieved by

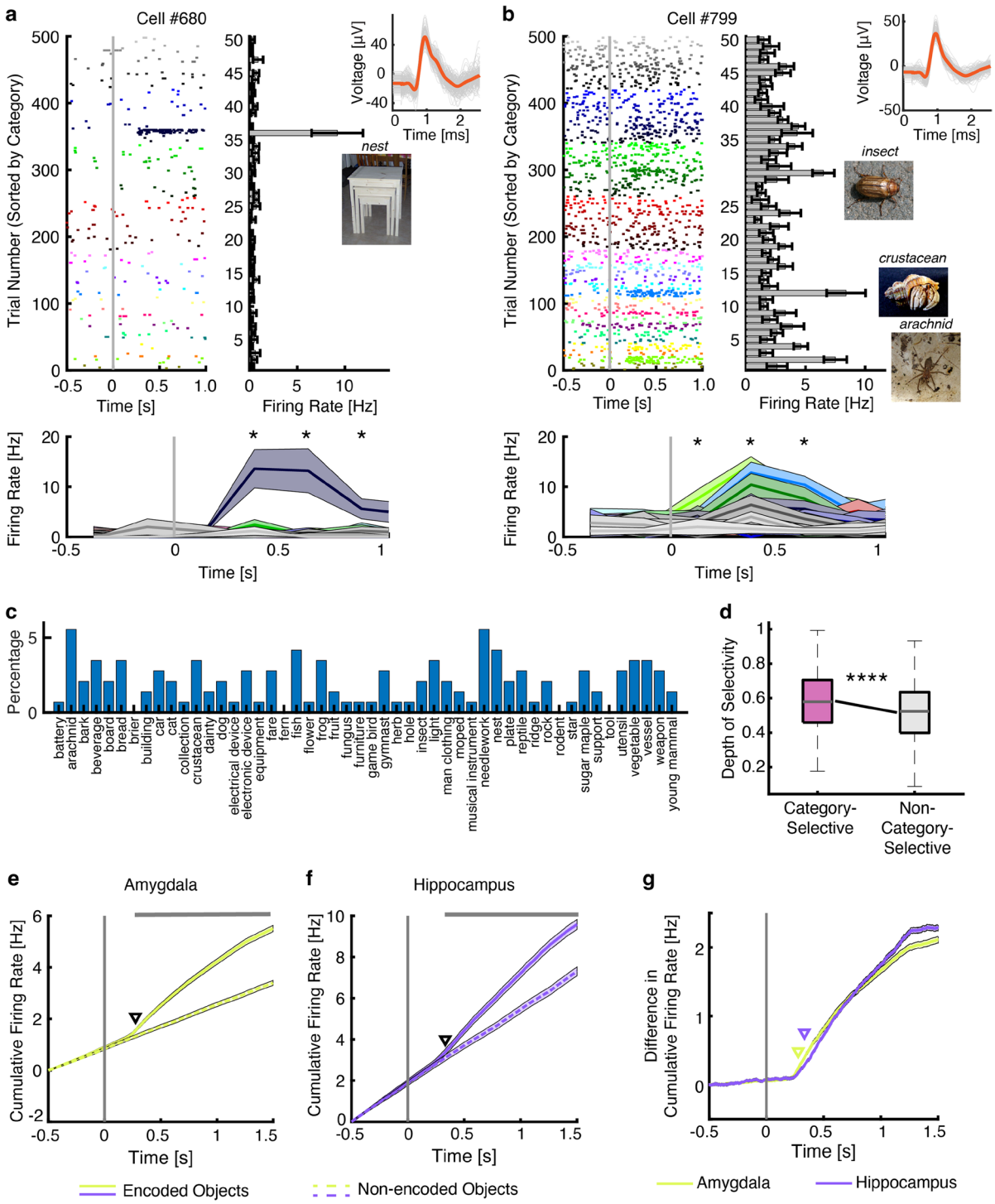

Fig. 4 Category-selective neurons from the ImageNet dataset. **(a,b)** Example neurons. Shown are neuronal responses to 500 images. Trials are aligned to stimulus onset (gray line) and are grouped by categories. Asterisk indicates a significant difference between object categories in that bin ($P < 0.05$, one-way ANOVA, after Bonferroni correction for multiple comparisons; bin size = 250 ms). Average waveforms for the example neurons were shown in the upper right. Shaded area denotes ± SEM across trials. Error bars denote ± SEM across images. **(c)** Summary of selected object categories. Shown are the percentages of neurons with the most preferred category. **(d)** Depth of selectivity (DOS) index. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points the algorithm considers to be not outliers, and the outliers are plotted individually. Asterisks indicate a significant difference using two-tailed two-sample *t*-test. ****$P < 0.0001$. **(e,f)** Cumulative firing rate for encoded objects (solid line) and non-encoded objects (dashed lines) of category-selective neurons from the amygdala (**e**; n = 90 neurons) and hippocampus (**f**; n = 54 neurons). Shaded area denotes ± SEM across neurons. Top bar shows clusters of time points with a significant difference (one-tailed pairwise *t*-test; $p < 0.005$; FDR-corrected; cluster size > 20 time points). Arrows indicate the first time point of the significant cluster. **(g)** Difference in cumulative firing rate (same data as shown in **[e,f]**). Shaded area denotes ± SEM across neurons.

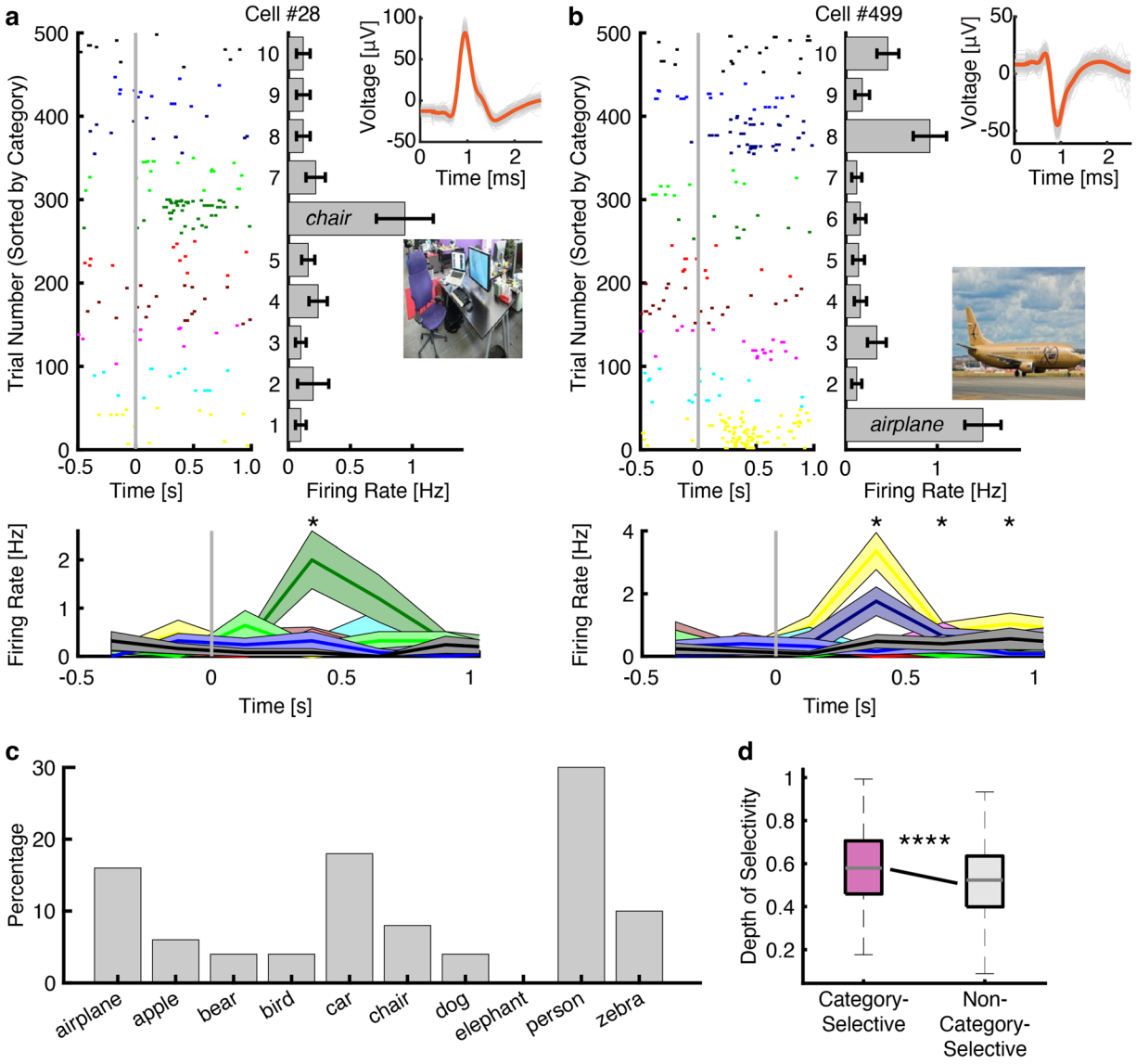

Fig. 5 Category-selective neurons from the Microsoft COCO dataset. Legend conventions as in Fig. 4.

comparing the cumulative firing rates for encoded versus non-encoded objects (see Methods) within each group of neurons. Our results revealed that the amygdala ($t = 278$ ms; Fig. 4e,g) displayed category-selective responses earlier ($P = 0.008$) than the hippocampus ($t = 332$ ms; Fig. 4f,g). Because the quantification of latency depends on multiple variables, such as the number of trials and response amplitudes[43], the absolute latency should be interpreted with caution.

Usage Notes

Code for preprocessing each part of the data is located in the corresponding directory '*/Code/' (e.g., to extract neuron firing rate, see '/SingleNeuron/Code/SU_getFiringRate.m'). Demonstration code for analyzing each part of the data is located in the '/Analysis/' directory. 'plotSortingQuality.m' performs analysis and plotting of spike sorting quality. '*NeuronSelection.m' (* denotes the neuron category) performs statistics and neuron selection. 'plotRasterSampleCell.m' plots activity of example neurons. 'ComputeLatency.m' calculates response latency and performs statistics between recording areas. Variables that need to be adjusted are explained at the beginning of each code. Functions that are common to all analyses are located in the 'Analysis/tools/' directory.

Code availability

The source code is included as part of the dataset[40]. All code is implemented using MATLAB (MathWorks Inc.)

Received: 24 June 2024; Accepted: 9 December 2024;
Published online: 15 January 2025

References

1. Kreiman, G., Koch, C. & Fried, I. Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat Neurosci* **3**(9), 946–953 (2000).
2. Rutishauser, U. *et al*. Representation of retrieval confidence by single neurons in the human medial temporal lobe. *Nat Neurosci* **18**(7), 1041–1050 (2015).
3. Wang, S. *et al*. Encoding of Target Detection during Visual Search by Single Neurons in the Human Brain. *Current Biology* **28**(13), 2058–2069.e4 (2018).
4. Reber, T. P. *et al*. Representation of abstract semantic knowledge in populations of human single neurons in the medial temporal lobe. *PLOS Biology* **17**(6), e3000290 (2019).
5. Barlow, H. B. Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology? *Perception* **1**(4), 371–394 (1972).
6. Valentine, T. A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A* **43**(2), 161–204 (1991).
7. Quian Quiroga, R. *et al*. Invariant visual representation by single neurons in the human brain. *Nature* **435**(7045), 1102–1107 (2005).
8. Quian Quiroga, R. Concept cells: the building blocks of declarative memory functions. *Nature Reviews Neuroscience* **13**, 587 (2012).
9. Quian Quiroga, R. *et al*. Sparse but not 'Grandmother-cell' coding in the medial temporal lobe. *Trends in Cognitive Sciences* **12**(3), 87–91 (2008).
10. Rutishauser, U. *et al*. The Architecture of Human Memory: Insights from Human Single-Neuron Recordings. *The Journal of Neuroscience* **41**(5), 883 (2021).
11. Mormann, F. *et al*. A category-specific response to animals in the right human amygdala. *Nat Neurosci* **14**(10), 1247–9 (2011).
12. Bao, P. *et al*. A map of object space in primate inferotemporal cortex. *Nature* **583**(7814), 103–108 (2020).
13. Bashivan, P., Kar, K. & DiCarlo, J. J. Neural population control via deep image synthesis. *Science* **364**(6439), eaav9436 (2019).
14. Ponce, C. R. *et al*. Evolving Images for Visual Neurons Using a Deep Generative Network Reveals Coding Principles and Neuronal Preferences. *Cell* **177**(4), 999–1009.e10 (2019).
15. Deng, J. *et al*. Imagenet: A large-scale hierarchical image database. in *2009 IEEE conference on computer vision and pattern recognition*. 2009.
16. Lin, T.-Y. *et al*. Microsoft COCO: Common Objects in Context. Cham: Springer International Publishing (2014).
17. Kar, K. *et al*. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat Neurosci* **22**(6), 974–983 (2019).
18. Joseph Redmon, S. D., Ross G. & Ali F., You Only Look Once: Unified, Real-Time Object Detection, in *CVPR*. (2016).
19. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. CVPR (2016).
20. Schrimpf, M. *et al*. Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like? (2020).
21. Achal Dave, P. D., Ramanan, D., Kirillov, A. & Girshick, R. Evaluating Large-Vocabulary Object Detectors: The Devil is in the Details, in *CVPR*. (2022).
22. Todorov, A. Data-driven methods for Modeling social perception. (2011).
23. Chang, L. & Tsao, D. Y. The Code for Facial Identity in the Primate Brain. *Cell* **169**(6), 1013–1028 e14 (2017).
24. Grossman, S. *et al*. Convergent evolution of face spaces across human face-selective neuronal groups and deep convolutional networks. *Nature Communications* **10**(1), 4934 (2019).
25. Rutishauser, U., Mamelak, A. N. & Schuman, E. M. Single-Trial Learning of Novel Stimuli by Individual Neurons of the Human Hippocampus-Amygdala Complex. *Neuron* **49**(6), 805–813 (2006).
26. Rutishauser, U. *et al*. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* **464**(7290), 903–907 (2010).
27. Schalk, G. *et al*. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering* **51**(6), 1034–1043 (2004).
28. Rutishauser, U., Schuman, E. M. & Mamelak, A. N. Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, *in vivo*. *Journal of Neuroscience Methods* **154**(1-2), 204–224 (2006).
29. Pouzat, C., Mazor, O. & Laurent, G. Using noise signature to optimize spike-sorting and to assess neuronal classification quality. *J Neurosci Methods* **122**(1), 43–57 (2002).
30. Harris, K. D. *et al*. Temporal Interaction between Single Spikes and Complex Spike Bursts in Hippocampal Pyramidal Cells. *Neuron* **32**(1), 141–149 (2001).
31. Schmitzer-Torbert, N. *et al*. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131**(1), 1–11 (2005).
32. Wang, S. *et al*. Neurons in the human amygdala selective for perceived emotion. *Proceedings of the National Academy of Sciences* **111**(30), E3110–E3119 (2014).
33. Wang, S. *et al*. The human amygdala parametrically encodes the intensity of specific facial emotions and their categorical ambiguity. *Nature Communications* **8**, 14821 (2017).
34. Cao, R. *et al*. Encoding of facial features by single neurons in the human amygdala and hippocampus. bioRxiv (2020).
35. Cao, R. *et al*. Feature-based encoding of face identity by single neurons in the human medial temporal lobe. bioRxiv, p. 2020.09.01.278283 (2020).
36. Cao, R. *et al*. A neuronal social trait space of faces in the human amygdala and hippocampus. *bioRxiv* (2021).
37. Cao, R. *et al*. Neural mechanisms of face familiarity and learning in the human amygdala and hippocampus. *Cell Reports* **43**(1), 113520 (2024).
38. Rainer, G., Asaad, W. F. & Miller, E. K. Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* **393**(6685), 577–579 (1998).
39. Minxha, J. *et al*. Fixations Gate Species-Specific Responses to Free Viewing of Faces in the Human and Macaque Amygdala. *Cell Reports* **18**(4), 878–891 (2017).
40. Cao, R. *et al*. Data for "Single-Neuron Recordings in the Medial Temporal Lobe for Natural Objects". OSF https://doi.org/10.17605/OSF.IO/VH5KQ (2025).
41. Faraut, M. C. M. *et al*. Dataset of human medial temporal lobe single neuron activity during declarative memory encoding and recognition. *Scientific Data* **5**(1), 180010 (2018).
42. Yuan, A. X. *et al*. Multi-day neuron tracking in high-density electrophysiology recordings using earth mover's distance. *eLife* **12**, RP92495 (2024).
43. Tang, H. *et al*. Spatiotemporal dynamics underlying object completion in human ventral visual cortex. *Neuron* **83**(3), 736–48 (2014).

Acknowledgements

We thank all patients for their participation and staff from WVU Ruby Memorial Hospital and Barnes-Jewish Hospital for support with patient testing. This research was supported by the AFOSR (FA9550-21-1-0088), NSF (BCS-1945230), NIH (K99EY036650, R01MH129426, R01MH120194, R01EB026439, U24NS109103, U01NS108916, U01NS128612, R21NS128307, P41EB018783), McDonnell Center for Systems Neuroscience, Fondazione Neurone, and Dana Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

R.C. and S.W. designed research. R.C., P.B., and S.W. performed research. N.J.B. and J.T.W. performed surgery. R.C. analyzed data. R.C. and S.W. wrote the paper. All authors discussed the results and contributed toward the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.C. or S.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.