© 2000 Nature America Inc. • http://neurosci.nature.com

Category-specific visual responses of single neurons in the human medial temporal lobe

Gabriel Kreiman[1], Christof Koch[1] and Itzhak Fried[2]

[1] *Computation and Neural Systems Program, California Institute of Technology, 139-74, Pasadena, California 91125, USA*

[2] *Division of Neurosurgery and Department of Psychiatry and Biobehavioral Sciences, Box 957039, UCLA School of Medicine, 740 Westwood Plaza, Los Angeles, California 90095-7039, USA*

Correspondence should be addressed to I.F. (ifried@mednet.ucla.edu)

© 2000 Nature America Inc. • http://neurosci.nature.com

The hippocampus, amygdala and entorhinal cortex receive convergent input from temporal neocortical regions specialized for processing complex visual stimuli and are important in the representation and recognition of visual images. Recording from 427 single neurons in the human hippocampus, entorhinal cortex and amygdala, we found a remarkable degree of category-specific firing of individual neurons on a trial-by-trial basis. Of the recorded neurons, 14% responded selectively to visual stimuli from different categories, including faces, natural scenes and houses, famous people and animals. Based on the firing rate of individual neurons, stimulus category could be predicted with a mean probability of error of 0.24. In the hippocampus, the proportion of neurons responding to spatial layouts was greater than to other categories. Our data provide direct support for the role of human medial temporal regions in the representation of different categories of visual stimuli.

Visual recognition of objects is a key function of the primate brain. There is a progression in the complexity of the representation of the visual scene by single neurons. Neurons in early visual areas in monkeys are tuned to simple features of the stimuli, such as the orientation of bars in area V1 or direction of motion in area V5. In the monkey inferotemporal cortex (IT), neurons respond to complex stimuli, including faces and hands, but also abstract patterns or common, everyday objects[1–3]. There are strong projections from IT to higher association areas in the temporal lobe, including the parahippocampal gyrus, perirhinal cortex, entorhinal cortex, hippocampus and amygdala[4,5]. Single neurons in these polymodal areas in monkeys also show visual selectivity for complex stimuli[6–8].

Temporal lobe lesions lead to profound category-specific deficits in visual recognition in both macaques and humans[9–12]. There is evidence from electrical stimulation studies in epileptic patients that current injection in the temporal lobe can interfere with visual recognition[13] and elicit visual memories and hallucinations[13,14]. Functional brain imaging and event-related potentials (ERP) also show a correlation between brain activity and visual recognition of specific categories of stimuli such as human faces and spatial layouts or places[15–20].

We reported that neurons in the human medial temporal lobe discriminate objects from faces[21]. Here we further investigated visual response properties and showed that single neurons respond selectively to different stimulus categories.

Results

We recorded the activity of 427 single neurons in 11 patients with pharmacologically resistant epilepsy who had intracranial depth electrodes implanted to determine the location of the seizure focus for possible surgical resection. Based on clinical criteria, electrode probes, containing several microwires each, were placed in medial temporal lobe targets bilaterally (Table 1). Based on MRI confirmation (Fig. 1), 149 of the neurons recorded were in the amygdala, 153 neurons in the entorhinal cortex and 125 neurons in the hippocampus. (Eighty-five percent of the sites were in the anterior segment of the hippocampus.) Most of the electrodes we recorded from were in the right temporal lobe (79%). During single-neuron recording, subjects were presented with visual stimuli (Fig. 2) and performed a simple discrimination task indicating whether the picture was a human face or not. Of the 427 neurons, 85 (20%) showed changes in firing rate during presentation of the visual stimuli, and 61 (14%) showed visually selective responses that were category specific.

Visual responses

Most neurons showed maintained firing rates below 10 spikes per second (Table 1). No significant response differences were observed between the right and left hemispheres, and therefore the data were pooled. The average overall firing rate was 3.6 ± 5.6 spikes per second, similar to observations in rats and monkeys[6–8,22].

We studied the responses of each neuron to the 1000-ms presentation of the visual stimuli by averaging the activity for all images within each category. For each neuron and each category, a post-stimulus time histogram was computed, showing the neuronal response starting 1000 ms before stimulus onset and ending 1000 ms after the stimulus disappeared.

A neuron was considered visually selective for a specific category if the activity during stimulus presentation for that category was significantly different from the baseline activity and from the responses to other categories of stimuli (Methods). For example, a visually selective neuron in the entorhinal cor-

© 2000 Nature America Inc. • http://neurosci.nature.com

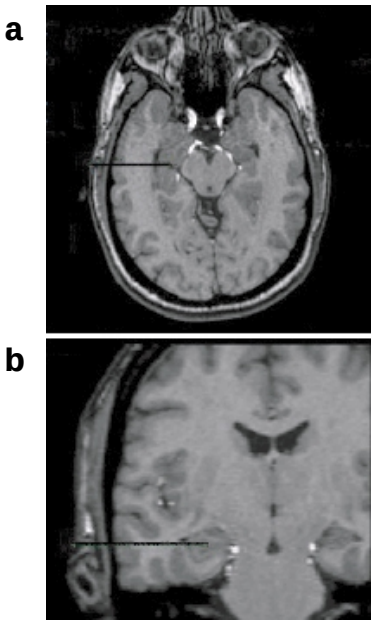

Fig. 1. Electrode placement. The trajectory of an electrode placed in the hippocampus is depicted in axial (**a**) and coronal (detail, **b**) structural MR images (1.5 Tesla scanner). Post-operative CT and MRI were used to confirm the location of the electrode. The CT was co-registered with MRI structural information for anatomic verification. The distal end of the electrode included platinum-iridium microwires from which single neurons were recorded. The microwires extended about 4 mm from the tip, lying on a cone with an opening angle of less than 45 degrees.

Fig. 2. Sample of stimuli presented in each category. Figures (mostly color) were drawn from a group of nine categories that included faces denoting emotional expressions by unknown actors[21], household objects, spatial layouts (including house exteriors, interiors and natural scenes), animals, cars, drawings of famous people or cartoon characters, photographs of famous people, food items and abstract patterns. Stimuli were presented for 1000 ms. Subjects had to indicate by pressing a button whether the image was a human face or not.

tex had increased firing rate in response to pictures of animals (Fig. 3a). The activity of this neuron during the 100–1000 ms interval after stimulus onset was different from baseline for animal stimuli ($p < 10^{-4}$, Wilcoxon rank-sum test), but not for the other stimulus categories ($p > 0.1$). A one-way ANOVA comparing firing rates between categories yielded $p < 0.001$, and comparing the activity for animals to all other categories using a pair-wise non-parametric Wilcoxon test yielded statistically significant differences ($p < 0.001$). The latency of response for this neuron was 219 ms, and the duration of increased response over baseline was 752 ms.

How specific was the response of the neuron within the selective category? If the average increased response to animals were due to enhanced firing for only a few pictures of animals, one might expect to observe a bimodal (or multimodal) distribution of firing rates. However, the distribution of firing rates for this neuron during presentation of different pictures of animals did not show any clear signs of multimodality (Fig. 3b). Although there was variability in the response of the neuron to individual instances of animals, the neuron responded above baseline for all pictures of animals (Fig. 3c; $p < 0.005$). We also compared the variability for different presentations of the same animal to the variability across different pictures of animals using an analysis of variance (ANOVA). Both parametric and non-parametric tests failed to show differences among individual stimuli ($p > 0.4$).

One visually selective neuron in the anterior hippocampus (Fig. 4) showed an increased firing rate over baseline in response to drawings of famous people as well as, to a lesser degree, to photos of famous people ($p < 0.001$). A one-way ANOVA yielded $p < 0.001$, and subsequent across-categories, pair-wise comparisons also showed that the activity during stimulus presentation was significantly higher for these two categories. Although the peak response was larger for drawings than for photographs (13.9 spikes/s versus 9.6 spikes/s), the average

Table 1. Number of neurons and response properties.

	Amygdala		Entorhinal cortex		Hippocampus		Total	
	Left	Right	Left	Right	Left	Right	Left	Right
Number of channels	21	57	13	58	12	54	215	
Number of neurons	50	99	24	129	44	81	427	
f (spikes/s)	2.1 ± 3.5		4.7 ± 5.4		4.0 ± 7.2			
Responsive neurons	6 (12)	12 (12)	6 (25)	29 (22)	16 (36)	16 (20)	28 (24)	57 (18)
Selective neurons	5 (10)	9 (9)	5 (21)	20 (16)	13 (30)	9 (11)	23 (19)	38 (12)
Latency (ms)	240 ± 145		209 ± 119		239 ± 132			
Durations (ms)	459 ± 322		507 ± 356		568 ± 325			

Number of recorded channels and isolated neurons in each location in each hemisphere from the 11 patients. On average, we recorded approximately two neurons per microwire. For each neuron, we computed the mean firing rate (*f*) over the entire experimental session. The firing rates in spikes/s ranged from 0.03 to 27 in the amygdala, 0.02 to 38 in the entorhinal cortex and 0.03 to 29 in the hippocampus. The criteria used to classify a neuron as visually responsive or selective are described in the text. The numbers in parenthesis indicate the percentages with respect to the total number of neurons in each location. Neurons with late responses were not included in the mean latency computation (Methods).

© 2000 Nature America Inc. • http://neurosci.nature.com

© 2000 Nature America Inc. • http://neurosci.nature.com

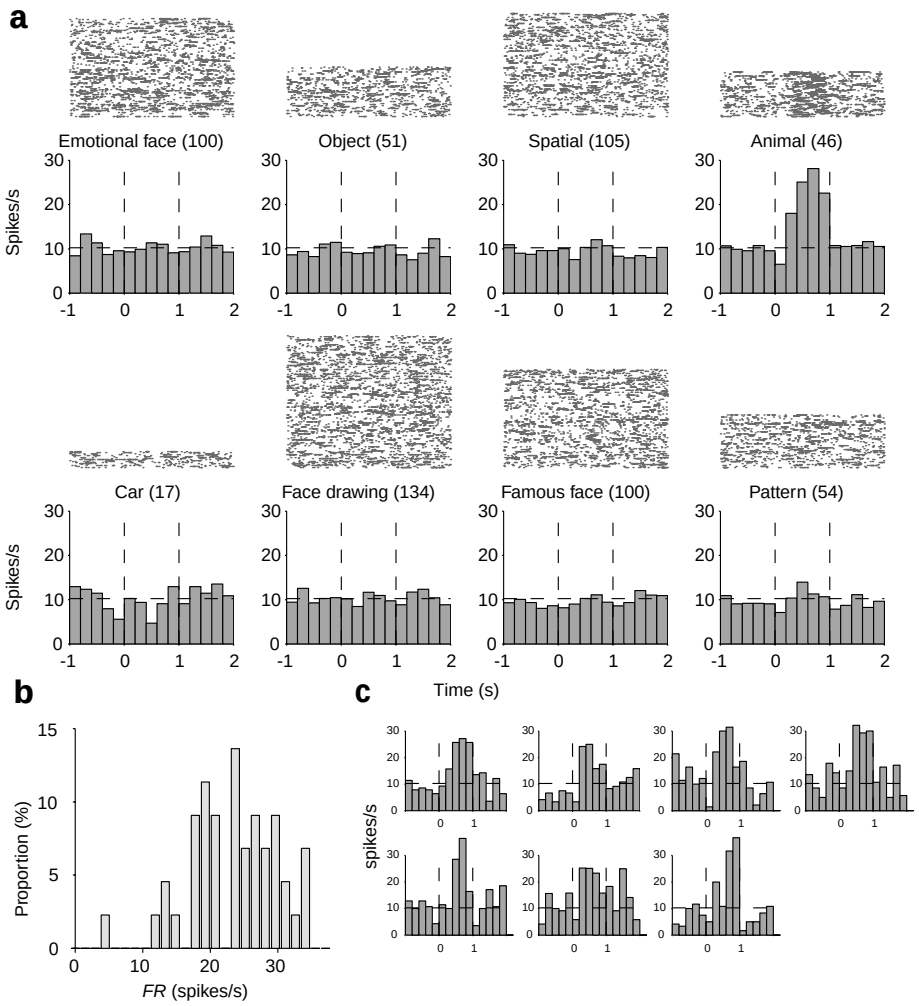

Fig. 3. Visually selective neuron in the entorhinal cortex. (**a**) Post-stimulus time histogram (PSTH) of the responses of a neuron in the right entorhinal cortex. The rasters and histograms are aligned to the onset of the stimulus. The stimulus was presented between $t = 0$ and $t = 1000$ ms (indicated by dashed vertical lines in each histogram). Responses were averaged for all stimuli within a given category using a bin size of 200 ms. The dashed horizontal line indicates the mean firing rate over the whole experiment (10.3 spikes/s). The category and the number of stimuli presented in each category are indicated at the top of each histogram. The firing rate in the 100–1000 ms interval upon presentation of a picture of an animal was significantly different from that in the –1000 to 0 ms baseline preceding the stimulus onset ($p < 10^{-4}$). The probability of error (p_e) from the ROC analysis (**Fig. 6**) was 0.21. There were only five presentations of food items in this experiment, and they were not included in the graph. The neuron did not respond to food items based on these five repetitions. (**b**) Distribution of firing rates during presentation of pictures of animals. Histogram distribution of mean firing rate of response in each trial, bin size of 1.5 spikes/s. There is no clear sign of bimodality in the distribution. (**c**) PSTHs showing the responses of this neuron to each individual stimulus within the category of animals. Although the responses vary from one stimulus to another, the neuron responds to all stimuli within this category (comparison with baseline, $p < 0.01$). An analysis of variance comparing the responses to different individual animals did not yield significance ($p > 0.4$). The scale is the same as in (**a**).

© 2000 Nature America Inc. • http://neurosci.nature.com

activity was not significantly different (Wilcoxon rank-sum test, $p > 0.05$). This neuron did not respond to faces *per se*, as indicated by the lack of change in the activity for emotional faces of unknown actors. The distribution of firing rates in response to photos of famous people for this neuron did not show clear signs of displaying more than one mode (Fig. 4b). Variability in the responses to distinct individual photos (Fig. 4c) was not higher than variability across different presentations of the same photograph (ANOVA, $p > 0.2$). Similar results held for drawings of famous people. The response to all individual stimuli within the selective categories was significantly different from baseline ($p < 0.01$).

Although we used a significance criterion of 0.05 in the statistical analysis, most of the actual p values were below 0.01. For the visually selective neurons, 69% of the p values were less than 0.01 (average p value, 0.01 ± 0.02). A χ^2 test rejected the hypothesis that these responses could be due to chance ($p < 0.001$).

Some of the neurons showed changes in firing rate in response to more than one of the categories (for example, Fig. 4). The proportion of neurons, relative to the number of selective neurons, that responded to more than one category was 21% in the amygdala, 10% in the entorhinal cortex and 25% in the hippocampus. Most of these neurons responded to two categories.

We also observed neurons that showed significant but nonselective changes in firing rate during stimulus presentation (Methods). The proportion of nonselective neurons, relative to the total number of responsive neurons, was 28% in the hippocampus, 29% in the entorhinal cortex and 17% in the amygdala.

Although most of the visually selective neurons showed increases in the firing rate on presentation of visual stimuli, some cells had reduced firing rate from baseline (three neurons in the amygdala, two in the entorhinal cortex and two in the hippocampus). Decreases were also observed in the visually responsive but nonselective neurons (one neuron in the amygdala, five in entorhinal cortex and three in hippocampus).

Most of the neurons responded during the stimulus-presentation period, but there were some that responded when the stimulus was removed. To address this, we computed for all neurons, within the 2000 ms after stimulus onset, the number of spikes in a 600-ms interval centered on the peak of the response and statistically analyzed the responses as described above. There were eight selective neurons (four in hippocam-

© 2000 Nature America Inc. • http://neurosci.nature.com

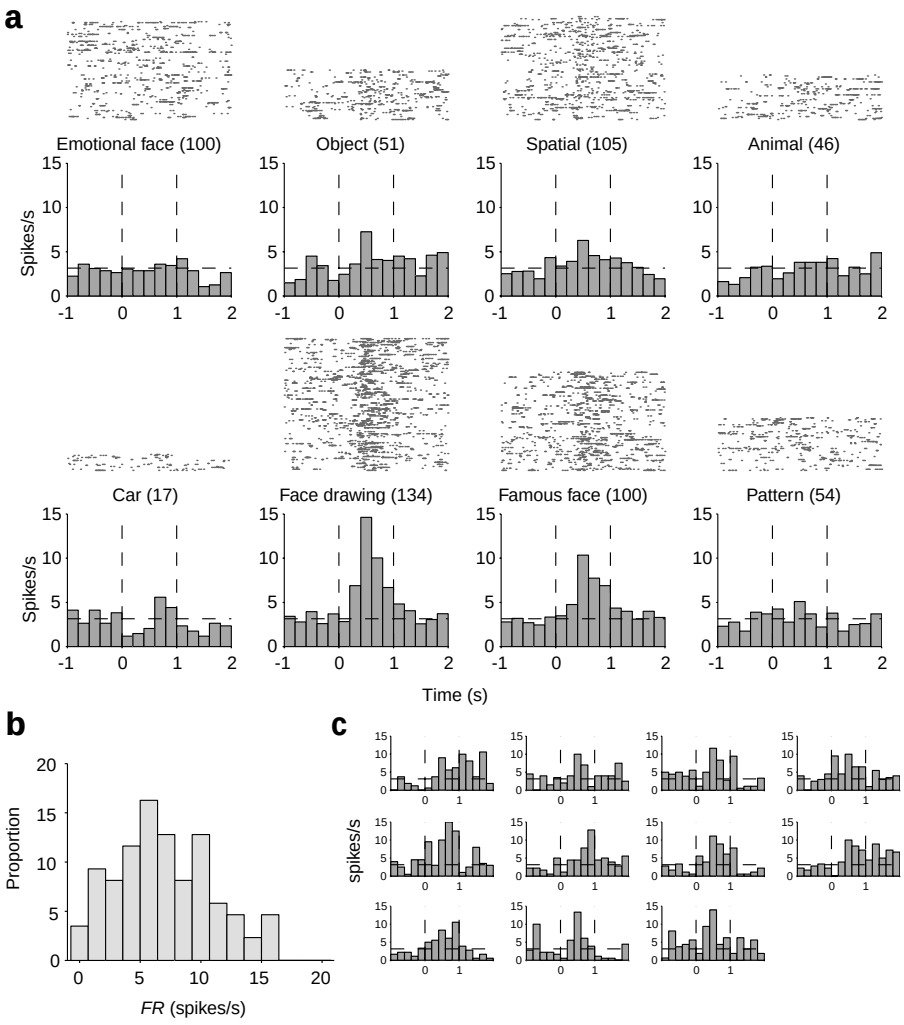

Fig. 4. Visually selective neuron in the hippocampus. (**a**) PSTH of the responses of a neuron in the right anterior hippocampus. The notation and symbols are the same as in the previous figure. The firing rate in the 100–1000 ms interval on presentation of a drawing or a photo of a famous face was significantly different from that in the –1000 to 0 ms baseline preceding stimulus onset ($p < 0.001$). Note that this neuron does not respond to just any face, as it fails to change its activity to the unknown actors depicting emotional expressions (top, left histogram). The mean firing rate over the whole experiment was 3.1 spikes/s, and the p_e was 0.19. There were only five presentations of food items in this experiment, and they were not included in the graph. The neuron did not respond to food items based on these five repetitions. (**b**) Distribution of firing rates during presentation of famous faces. Histogram distribution of mean firing rate of responses in each trial, bin size of 1.5 spikes/s. There is no clear sign of bimodality in the distribution. (**c**) PSTHs showing the responses of this neuron to each individual stimulus within the category of photos of famous faces. Although the responses vary from one stimulus to another, the neuron responds to all stimuli within this category (comparison with baseline, $p < 0.05$). An ANOVA comparing the responses to different individual famous faces did not yield significance ($p > 0.2$). The scale is the same as in (**a**).

© 2000 Nature America Inc. • http://neurosci.nature.com

pus, three in entorhinal cortex and one in the amygdala) that showed a statistically significant late response and were not detected with the previous analysis.

Under the assumption that there is no preference in the prevalence of selectivity for any of the nine different stimulus categories, the number of selective neurons within any one area should be uniformly distributed among these categories. We tested this hypothesis using a χ^2 test[23]. We obtained χ^2 values of 6.0, 9.6 and 29.5 for the amygdala, entorhinal cortex and hippocampus respectively (Fig. 5). The p values were over 0.25 for the amygdala and entorhinal cortex but less than 10^{-3} for the hippocampus. In the hippocampus, we observed a small relative proportion of responses to animals, food items and patterns and, interestingly, a relatively high number of neurons responding to spatial layouts (Fig. 5). No significant difference was observed in a direct comparison of the response of any of these neurons to house facades versus natural scenes (Wilcoxon, $p > 0.05$).

Among the selective neurons, there were 2 that yielded $p < 0.05$ (and 3 more with a p value between 0.05 and 0.1) in the ANOVA analysis of specificity to individual stimuli within the selective category. These neurons responded more strongly to one to three of the individual stimuli within the selective category. These were excluded from the number of selective neurons in Table 1.

We computed the latency and the duration of the evoked activities for all neurons showing a visual response (Methods). The latencies ranged from 52 to 695 ms and the durations from 53 to 1190 ms (Table 1). There was no significant difference among the amygdala, entorhinal cortex and hippocampus in either of these two variables (ANOVA, $p > 0.1$).

Classification by an ideal observer

We assessed how well an ideal observer could discriminate, based on the response of an individual neuron, whether a stimulus belonged to the category for which the neuron was selective or not. We computed p_e, the probability of misclassifying a stimulus based on the firing rate, using a classical optimal decision procedure (ROC analysis; Methods and Fig. 6a–c). The value of p_e can range from 0 to 0.5, with $p_e = 0.5$ indicating chance performance and $p_e = 0$ indicating perfect classification. Our p_e values ranged from 0.13 to 0.32 (0.22 ± 0.06, mean ± s.d.) in the amygdala, 0.04 to 0.44 (0.23 ± 0.10) in the entorhinal cortex and 0.08 to 0.47 (0.23 ± 0.10) in the hippocampus (Fig. 6d–f).

Discussion

Increasingly complex stimulus attributes are represented from the retina to the higher visual areas. Evidence from neurology[11,12,24], functional brain imaging[15–17,20] and evoked-potential

© 2000 Nature America Inc. • http://neurosci.nature.com

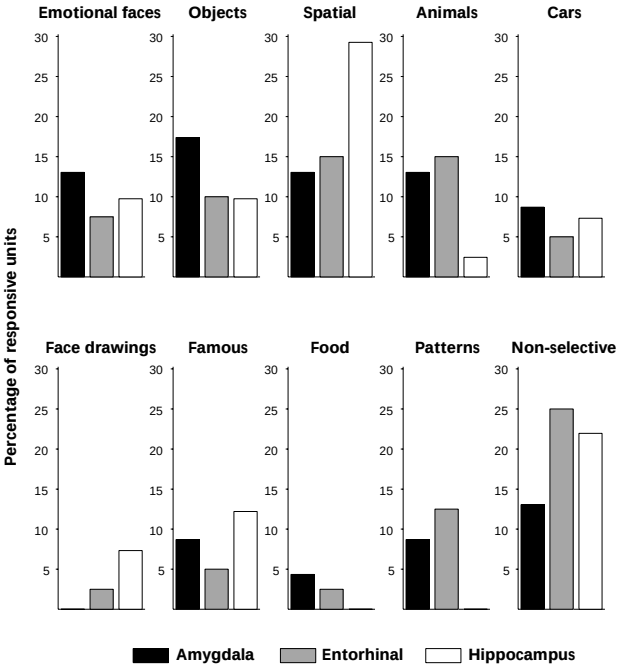

Fig. 5. Distribution of selective neurons for each category. The percentage of neurons selective for each category of stimuli and the percentage of nonselective neurons are shown for each location (black, amygdala; gray, entorhinal cortex; white, hippocampus). The percentages are based on the total number of responsive neurons in each location. Data from the right and left hemisphere have been pooled. A neuron was considered to be visually selective if the activity during stimulus presentation for a given category was significantly different from the baseline and from the neuronal response to other categories. If the response was different from baseline but not among the different categories, the neuron was defined as nonselective.

studies in humans[18], and from single-neuron electrophysiology[1–3] and lesions in monkeys[9,10,25] suggests a fundamental role for the medial temporal lobe in visual object recognition.

Category-specific knowledge deficits occur in which neurological patients show impairments in identifying living things, objects, food items or faces[11,12,24]. Functional imaging studies show activation of different areas in the temporal lobe that correlates with subjects' observing pictures belonging to different categories[15–17]. In particular, there are areas specialized for faces[20,26], spatial layouts[27,28], objects and animals[17]. Specific changes in activity on presentation of faces, objects and letter strings can also be observed by evoked potentials in humans[29–31].

Single inferotemporal cortex (IT) neurons in monkeys respond to complex visual stimuli, including faces, objects and abstract patterns[1–3]. Neurons in human temporal neocortex respond to faces and words[32,33]. Information from these neocortical neurons is conveyed to polymodal association areas in the temporal lobe[4,5].

We showed that in the relatively small matrix of hippocampus, entorhinal cortex and amygdala, there is a remarkable degree of segregation of categories at the level of single neurons. Neurons in these regions show visual object discrimination among at least nine stimulus categories. Based on the firing rate of individual neurons, it was possible to predict with a mean probability of error of 0.24 whether the preferred stimulus category was presented or not (Fig. 6). This, by itself, shows a striking degree of category-specific firing on a trial-by-trial basis. By combining the activity of multiple neurons, it is likely that an even higher level of accuracy can be achieved. Such category-specific processing may be important not only in object recognition, but also in the representation and retrieval processes that have been closely linked with the medial temporal lobe[34–36].

Twenty percent of the neurons that we recorded showed a visual response and 14% a visually selective one. These percentages are comparable to those reported in monkeys. In the entorhinal cortex, 11% of the neurons show selective visual responses[7], compared to 16% in our study. In the monkey amygdala, 12% of the neurons show visual responses and about 33% of those are selective for faces[6], compared to 12% and 20%, respectively, in our data.

Most of the selective neurons responded to only a single stimulus category, rather than weakly responding to a large fraction of all stimuli. Our data thus support the existence of sparse coding in the medial temporal lobe. Sparsely coded representation has been suggested for information processing in the rodent and primate hippocampus[37–39] and for processing of faces and objects in IT[40,41].

Whereas a significant proportion of neurons are selective for faces in the superior temporal sulcus[1–3], responses in the entorhinal cortex[7], hippocampus[8] and amygdala[6] are much more varied. Responses were also diverse in our sample of entorhinal cortex and amygdala neurons (Fig. 5). However, in the hippocampus, we observed more responses to images showing spatial layouts, including houses, natural scenes and interiors. The rat hippocampus contains place cells that respond selectively to the position of a rat while it is navigating a maze[22,42]. Neurons in the monkey hippocampus respond selectively depending on the position of the stimulus in a conditioned spatial response task[8]. Functional MRI studies report parahippocampal[43] and hippocampal[44] activation associated with navigational tasks as well as while observing images similar to the ones shown in our study[27,28]. This area is posterior to the medial temporal recording sites in our study, and likely projects to the hippocampus.

It is possible that some neurons may change their activity more specifically than to the broad categories used in our study, responding, for instance, only to one specific example that we did not present[45]. Our experimental setting did not enable investigation of this issue for several reasons. First, because of clinical considerations, an electrode location remained fixed once placed, and we did not change electrode location in search for an optimal stimulus, as is commonly done in animal experiments. Also, because the data for all the channels were analyzed off-line, we could not determine, via immediate feedback, the 'optimal' stimulus for any of the cells. For most of the selective neurons (61 of 63), the analysis of variance showed that variability for different stimuli within the selective category was comparable to variability due to different presentations of the same stimulus.

Data from very different experiments and using distinct techniques are converging to show an important role for the human medial temporal lobe in visual object recognition. This study establishes that single neurons in humans explicitly respond to specific categories of stimuli, which may be relevant to the representation and retrieval of visual information.

METHODS

Patients. Subjects were patients with pharmacologically resistant epilepsy. Extensive non-invasive evaluation did not yield concordant data corresponding to a single resectable epileptogenic focus, and therefore the patients were stereotactically implanted with up to 12 chronic

© 2000 Nature America Inc. • http://neurosci.nature.com

© 2000 Nature America Inc. • http://neurosci.nature.com

intracranial depth electrodes for one to two weeks to determine the focus of their seizures for possible surgical resection[21,46]. Through the lumen of the electrodes, up to 8 microwires (40 μm diameter) were inserted. The surgeries were performed by I.F. All studies described here conformed with the guidelines of the Medical Institutional Review Board at UCLA. The present study describes data from 11 subjects (6 males and 5 females; 8 right-handed and 3 left-handed; 24 to 48 years old).

The sites of implantation of the electrodes were based exclusively on clinical criteria. The location of the electrodes was verified by structural MR images obtained before removing the electrodes (Fig. 1). Individual microwires extended approximately 4 mm from the tip, lying in a cone with an opening angle of less than 45 degrees.

We report here the activity of neurons for all the probes located in the medial temporal lobe. Neurons from anterior (85%), middle (10%) and posterior (5%) parts of the hippocampus were pooled together as hippocampus neurons. Most neurons in the amygdala were in the basolateral nuclear complex. Our MR resolution did not allow us to accurately determine in which CA fields the hippocampal probes were placed or what layer of the entorhinal cortex we recorded from.

The information recorded during seizures from the depth electrodes was used to localize the seizure focus[46]. Eighty-five percent of recorded neurons were outside the clinically determined zone of seizure onset (that is, either in the other hemisphere or in a different brain area on the same side). Ninety-four percent of responsive neurons were outside the seizure focus. Because we did not observe any differences in their waveforms, firing rates, interspike interval distributions or response properties, all neurons were included in Table 1. The percentage of responsive neurons would increase from 20% to 22% if we excluded neurons within the seizure focus.

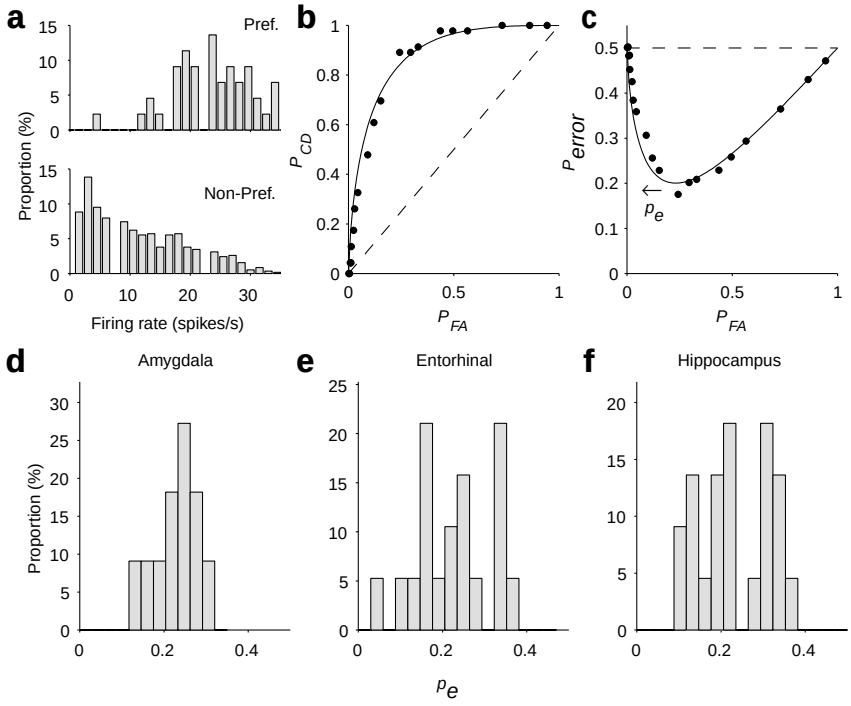

Fig. 6. ROC analysis and p_e distribution. (**a**) Distribution of firing rates for the neuron shown in Fig. 3 for the preferred category (top, animals) and non-preferred category (bottom, all other categories). Bin size, 1.5 spikes/s. These represent the conditional probability distributions P(*f* | *stim* ∈ *preferred category*) and P(*f* | *stim* ∉ *preferred category*), where *f* denotes the firing response computed in the 100–1000 ms interval and *stim* indicates the stimulus. (**b**) ROC analysis. Each stimulus is classified into the preferred category if the firing rate during the 100–1000 ms interval is above a given threshold *T*. The probability of correctly classifying a stimulus (P_{CD}) is plotted as a function of the probability of making a false alarm (P_{FA}). This was calculated for successive values of the threshold T by integrating the tails of the two distributions: P(*f* > *T* | *stim* ∈ *preferred cat.*) and P(*f* > *T* | *stim* ∉ *preferred cat.*). The dashed line indicates chance performance ($P_{CD} = P_{FA}$). The dots indicate the actual data using numerical integration, whereas the continuous line indicates the values after assuming a Gaussian distribution of firing rates. (**c**) The overall probability of error, $p_{error}(T) = 1/2\ P_{FA}(T) + 1/2\ (1–P_{CD}(T))$ is plotted as a function of the probability of false alarm. The classification performance of the neuron based on the firing rate was characterized by the minimum in this curve, p_e (indicated by an arrow). (**d–f**) Distribution of p_e for neurons in each of the three locations. Bin size, 0.03.

Experimental protocol. A series of images was shown on a monitor at an approximate size of 5 degrees of visual angle (Fig. 2). Each picture was repeated 4–10 times (depending on time constraints), and there was a total of up to 600 presentations; the order of presentation of the stimuli was random. The number of different individual stimuli per category ranged from 3 to 25 (7.2 ± 4.6, mean ± s.d.). Each picture was presented for 1000 ms. In the first two patients, stimuli from only three different categories were presented (emotional faces, objects and spatial layouts). Immediately after the picture disappeared, there was a tone that indicated that the subject had to respond whether the picture was a human face or not by pressing a button. This was done to engage the subject's attention and to verify that he was seeing the pictures. Trials in which subjects made an error (mean percentage correct 97 ± 1%) or in which the subjects moved were discarded from subsequent analysis. The behavioral response occurred on average 472 ± 201 ms after stimulus offset.

Recordings. Data from each of the recorded microwires were amplified and high-pass filtered (with a corner frequency of 300 Hz), A/D converted and stored for off-line spike sorting using Experiment Workbench data acquisition software (Datawave, Denver, Colorado). Note that because the microelectrodes were chronically implanted, we could not further select neurons by moving the electrodes, as is common in monkey experiments. We analyzed all spikes from all neurons we detected.

Spikes from single neurons were discriminated from the extracellular recordings based on the height, width, peak voltage and other parameters of the waveforms using a manual cluster-cutting method implemented in Datawave. For each isolated neuron, we determined the fraction of all spikes that were within 2 ms of each other. If this fraction exceeded 2%, the data were discarded because of possible contamination by firing from more than one neuron.

In some cases, we recorded from the same microwires on separate days (presenting different sets of pictures from the same categories). Because of various constraints, we often could not be sure that a neuron recorded on one day was identical to a neuron recorded on another day. Of our 427 neurons, 128 (30%) were recorded on consecutive days. If we assume that all 128 neurons are identical across days, the total number of neurons reduces to 299, and the percentage of responsive neurons increases from 20% (85 of 427) to 23% (70 of 299).

Data analysis. For each neuron, we computed histograms locked to the stimulus presentation times (post-stimulus time histograms, PSTHs) by averaging the neuronal responses for all stimuli within each category. We also compared all color versus non-color pictures; we found only two neurons with a statistically significant enhanced response to color stimuli. To assess the significance of the response to different categories of

© 2000 Nature America Inc. • http://neurosci.nature.com

© 2000 Nature America Inc. • http://neurosci.nature.com

visual stimuli, we ran a series of statistical comparisons. For a neuron to be considered visually selective for a specific category, it had to meet three requirements. First, the neuronal response pooled over all stimuli belonging to the same category had to show a different firing rate, that is, lower or higher, during the time of presentation of the image than during the preceding baseline. The firing rate during image presentation was computed in the interval 100 ms ≤ t < 1000 ms. (The lower boundary of 100 ms was chosen because most neurons showed latencies above 100 ms.) The baseline before stimulus presentation was computed in the interval –1000 ms ≤ t < 0 ms. Significance was assessed both by a two-tailed t-test and a non-parametric Wilcoxon rank-sum test[23]. Because the results from both tests were very similar, the p values reported throughout the text for pairwise comparisons correspond to those from the non-parametric test. Second, a one-way analysis of variance (ANOVA) during the 100–1000 ms interval after stimulus onset, assessing whether there were differences in firing rate among the different categories, had to show a significant p-value (< 0.05). In this analysis, stimuli were pooled according to the category they belonged to. Third, subsequent pairwise comparisons between the activity during this interval for the putative selective category and the rest of the categories had to show a statistically significant change (Wilcoxon rank sum test, < 0.05).

Although we used a significance criterion of 0.05, most responsive neurons showed p values that were less than 0.01 (Results). Our analysis encompassed both increases and decreases in firing rate.

If the ANOVA test failed to indicate a significant difference between categories, but the activity for 75% of the stimulus categories was significantly different from the baseline activity (that is, the first criterion was met but not the second), we labeled the neuron as visually responsive but nonselective.

For the neurons that showed visual selectivity, we further studied the degree of specificity of the responses. We computed the distribution of firing rates for the selective category to assess whether it was a polymodal distribution. We also computed an analysis of variance (both a parametric analysis[23] as well as a non-parametric analysis using a bootstrap procedure[47]) to compare the variance to different individual stimuli within the same category to the variance to repeated presentations of the same stimulus.

It is reported that neurons can respond to visual stimuli with a long delay even after the stimulus disappears[21,48,49]. To study these cases, we analyzed the responses in an interval of 600 ms centered on the response peak. To estimate the time of occurrence of the peak, we computed an estimation of the spike density function (*sdf*) for each category of stimuli by convolving the spike trains with a Gaussian of fixed width of 100 ms and then averaging over repetitions.

We computed the latency and duration of the responses from the *sdf* for each category. The latency was defined as the first time value on which five consecutive bins of the *sdf* yielded a value that was beyond two s.d. of the mean response before stimulus presentation. In analogous fashion, the end of the response was defined as the first time point at which five consecutive bins of the *sdf* yielded a value not different from baseline after the latency value.

To discriminate whether the neurons showed significant activity related to the behavioral response, we did two additional analyses. First, we computed histograms of all the neuronal responses locked to the time at which the subjects pressed the button by averaging over all stimuli. We then computed whether there was a significant difference in the response in the interval 300 ms before and 300 ms after the button was pressed. This interval was chosen so that it would not overlap on average with the periods of visual presentation. We found four neurons (three in the amygdala, one in the hippocampus) for which the response before the button press was significantly different from that after the button press and also from that during the baseline –1000 to 0 ms interval (1-way ANOVA, p < 0.05; pairwise comparisons, p < 0.05). None of these neurons was visually responsive. We also found 2 neurons that showed an increased activity in a 600-ms window around the response compared to the baseline (Wilcoxon test, p < 0.05). Neither of these neurons was visually responsive.

When computing the average neuron activity in response to all human faces (emotional faces, drawings and photos of famous faces), we observed 10 neurons (5 in the hippocampus, 3 in the entorhinal cortex, 2 in the amygdala) for which the activity in the 100–1000 ms interval for all face stimuli was significantly larger than the baseline activity and than the response for non-face stimuli (1-way ANOVA, p < 0.05; pairwise comparisons, p < 0.05). However, for some of these neurons (for example, Fig. 4a), the response was selective for some faces and not others.

For those neurons that showed visual selectivity for stimuli within specific categories, we also addressed the question of how well an ideal observer could estimate which category the stimulus belonged to by observing the firing rate. To quantify the classification performance of the neurons, we used an ROC analysis as used in signal-detection theory and psychophysics experiments[50]. For each visually selective neuron, we computed the distribution of firing rates for the preferred stimuli and the non-preferred stimuli (the remaining categories; Fig. 6a). From the distribution of firing rates we evaluated, by sliding a threshold T over the whole range of firing rates, the probability of correct detection (P_{CD}) and the probability of false alarm (P_{FA}). Assuming chance performance, $P_{CD} = P_{FA}$ (Fig. 6b, dashed line). The departure from the diagonal shows the possibility of discriminating between the preferred and non-preferred categories. The probability of misclassification plotted against the probability of false alarm can be obtained as

$$p_{error}(T) = 1/2\, P_{FA}(T) + 1/2\,(1 - P_{CD}(T))$$

The overall probability of error, p_e, is then defined as the minimum value of this function (arrow in Fig. 6c). A value of p_e = 1/2 corresponds to chance performance, whereas a value of p_e = 0 indicates that it is possible to predict with perfect accuracy based on the number of spikes whether the specified category was presented or not.

Throughout the manuscript, values are given as mean ± standard deviation (s.d.).

Acknowledgements

This work was supported by grants from NIH (to I.F.), the Keck Foundation (to C.K.) and the Center for Consciousness Studies, University of Arizona (to I.F. and C.K.). We would like to thank D. Rozenfarb and M. Zirlinger for suggestions and reading the manuscript and Peter Steinmetz for general discussion. We wish to acknowledge Tony Fields, Jack Morrow, Eve Isham, Charles Wilson, Rick Staba, Eric Behnke and Anatol Bragin for help with the recordings.

Received 16 June; accepted 18 July 2000

1. Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* 19, 577–621 (1996).
2. Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* 19, 109–139 (1996).
3. Gross, C. G. How inferior temporal cortex became a visual area. *Cereb. Cortex* 5, 455-469 (1994).
4. Suzuki, W. A. Neuroanatomy of the monkey entorhinal, perirhinal and parahippocampal cortices: Organization of cortical inputs and interconnections with amygdala and striatum. *Sem. Neurosci.* 8, 3–12 (1996).
5. Saleem, K. S. & Tanaka, K. Divergent projections from the anterior inferotemporal area TE to the perirhinal and entorhinal cortices in the macaque monkey. *J. Neurosci.* 16, 4757–4775 (1996).
6. Leonard, C. M., Rolls, E. T., Wilson, F. A. W. & Baylis, G. C. Neurons in the amygdala of the monkey with responses selective for faces. *Behav. Brain. Res* 15, 159–176 (1985).
7. Suzuki, W., Miller, E. & Desimone, R. Object and place memory in the macaque entorhinal cortex. *J. Neurophysiol.* 78, 1062–1081 (1997).
8. Miyashita, Y., Rolls, E. T., Cahusac, P. M. B., Niki, H. & Feigenbaum, J. D. Activity of hippocampal formation neurons in the monkey related to a conditional spatial response task. *J. Neurophysiol.* 61, 669–678 (1989).
9. Mishkin, M. Memory in monkeys severely impaired by combined but not separate removal of amygdala and hippocampus. *Nature* 273, 297–298 (1978).
10. Meunier, M., Bachevaoer, J., Mishkin, M. & Murray, E. A. Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *J. Neurosci.* 13, 5418–5432 (1993).
11. Tranel, D., Damasio, H. & Damasio, A. Face agnosia and the neural substrates of memory. *Annu. Rev. Neurosci.* 13, 89–109 (1990).
12. Warrington, E. & McCarthy, R. Categories of knowledge — Further fractionations and an attempted integration. *Brain* 110, 1273–1296 (1987).
13. Fried, I., Mateer, C., Ojemann, G., Wohns, R. & Fedio, P. Organization of visuospatial functions in human cortex. *Brain* 105, 349–371 (1982).

© 2000 Nature America Inc. • http://neurosci.nature.com

© 2000 Nature America Inc. • http://neurosci.nature.com

14. Penfield, W. & Jasper, H. *Epilepsy and the Functional Anatomy of the Human Brain* (Little, Brown, Boston, 1954).
15. Treisman, A. M. & Kanwisher, N. G. Perceiving visually presented objects: recognition, awareness, and modularity. *Curr. Opin. Neurobiol.* 8, 218–226 (1998).
16. Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L. & Haxby, J. V. Distributed representation of objects in the human ventral visual pathway. *Proc. Natl. Acad. Sci. USA* 96, 9379–9384 (1999).
17. Martin, A., Wiggs, C. L., Ungerleider, L. G. & Haxby, J. V. Neural correlates of category-specific knowledge. *Nature* 379, 649–652 (1996).
18. Allison, T. *et al.* Face recognition in human extrastriate cortex. *J. Neurophysiol.* 71, 821–825 (1994).
19. Nobre, A., Allison, T. & McCarthy, G. Word recognition in the human inferior temporal lobe. *Nature* 372, 260–263 (1994).
20. Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O. & Kanwisher, N. Response properties of the human fusiform face area. *Cognit. Neuropsychol.* 17, 257–279 (2000).
21. Fried, I., MacDonald, K. A. & Wilson, C. Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18, 753–765 (1997).
22. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* 261, 1055–1058 (1993).
23. Keeping, E. S. *Introduction to Statistical Inference* (Dover, New York, 1995).
24. Adolphs, R., Tranel, D., Damasio, H. & Damasio, A. Impaired recognition of emotion in facial expressions following bilateral damage to the amygdala. *Nature* 372, 669–672 (1994).
25. Buckley, M. J. & Gaffan, D. Perirhinal cortex ablation impairs object identification. *J. Neurosci.* 18, 2268–2275 (1998).
26. Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311 (1997).
27. Epstein, R. & Kanwisher, N. A cortical representation of the local visual environment. *Nature* 392, 598–601 (1998).
28. Epstein, R., Harris, A., Stanley, D. & Kanwisher, N. The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23, 115–125 (1999).
29. Allison, T., Puce, A., Spencer, D. & McCarthy, G. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430 (1999).
30. Klopp, J., Halgren, E., Marinkovic, K. & Nenov, V. Face-selective spectral changes in the human fusiform gyrus. *Clin. Neurophysiol.* 110, 676–682 (1999).
31. Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V. & Hari, R. Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb. Cortex* 10, 69–81 (2000).
32. Ojemann, J. G., Ojemann, G. A. & Lettich, E. Neuronal activity related to faces and matching in human right nondominant temporal cortex. *Brain* 115, 1–13 (1992).
33. Ojemann, G. A., Ojemann, S. G. & Fried, I. Lessons from the human brain: neuronal activity related to cognition. *The Neuroscientist* 4, 285–300 (1998).
34. Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry.* 20, 11–21 (1957).
35. Penfield, W. & Milner, B. Memory deficit produced bilateral lesions in the hippocampal zone. *Arch. Neurol. Psychol.* 79, 475–479 (1958).
36. Zola-Morgan, S. & Squire, L. R. Neuroanatomy of memory. *Annu. Rev. Neurosci.* 16, 547–563 (1993).
37. McNaughton, B. L. & Nadel, L. in *Neuroscience and Connectionist Theory* (ed. Seifert, W.) 1–64 (Academic, New York, 1990).
38. Alvarez, P. & Squire, L. R. Memory consolidation and the medial temporal lobe: a simple network model. *Proc. Natl. Acad. Sci. USA* 91, 7041–7045 (1994).
39. Treves, A. & Rolls, E. T. Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4, 374–391 (1994).
40. Young, M. P. & Yamane, S. Sparse population coding of faces in the inferior temporal cortex. *Science* 256, 1327–1331 (1992).
41. Logothetis, N. K. & Pauls, J. Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb. Cortex* 3, 270–288 (1995).
42. Muller, R. A quarter of a century of place cells. *Cell* 17, 813–822 (1996).
43. Aguirre, G. K., Detre, J. A., Alsop, D. C. & D'Esposito, M. The parahippocampus subserves topographical learning in man. *Cereb. Cortex* 6, 823–829 (1996).
44. Maguire, E. A. *et al.* Knowing where and getting there: a human navigation network. *Science* 280, 921–924 (1998).
45. Heit, G., Smith, M. E. & Halgren, E. Neural encoding of individual words and faces by the human hippocampus and amygdala. *Nature* 333, 773–775 (1988).
46. Fried, I. *et al.* Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients. *J. Neurosurg.* 91, 697–705 (1999).
47. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (Chapman & Hall/CRC, London, 1993).
48. Yakovlev, V., Fusi, S., Berman, E. & Zohary, E. Inter-trial neuronal activity in inferior temporal cortex: a putative vehicle to generate long-term visual associations. *Nat. Neurosci.* 1, 310–317 (1998).
49. Miyashita, Y. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335, 817–820 (1988).
50. Green, D. & Swets, J. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).

© 2000 Nature America Inc. • http://neurosci.nature.com