

Single-Trial Learning of Novel Stimuli by Individual Neurons of the Human Hippocampus-Amygdala Complex

Report

Ueli Rutishauser,¹ Adam N. Mamelak,^{2,3}
and Erin M. Schuman^{1,*}

¹Computation and Neural Systems
Division of Biology and
Howard Hughes Medical Institute
California Institute of Technology
Pasadena, California 91125

²Epilepsy and Brain Mapping Unit
Huntington Memorial Hospital
Pasadena, California 91105

³Maxine Dunitz Neurosurgical Institute
Cedars-Sinai Medical Center
Los Angeles, California 90048

Summary

The ability to distinguish novel from familiar stimuli allows nervous systems to rapidly encode significant events following even a single exposure to a stimulus. This detection of novelty is necessary for many types of learning. Neurons in the medial temporal lobe (MTL) are critically involved in the acquisition of long-term declarative memories. During a learning task, we recorded from individual MTL neurons in vivo using microwire electrodes implanted in human epilepsy surgery patients. We report here the discovery of two classes of neurons in the hippocampus and amygdala that exhibit single-trial learning: novelty and familiarity detectors, which show a selective increase in firing for new and old stimuli, respectively. The neurons retain memory for the stimulus for 24 hr. Thus, neurons in the MTL contain information sufficient for reliable novelty-familiarity discrimination and also show rapid plasticity as a result of single-trial learning.

Introduction

One prominent feature of nervous systems is the ability to distinguish novel from familiar stimuli. A rapid assessment of stimulus novelty is a prerequisite for certain kinds of learning (Davis et al., 2004; Kohonen and Lehtio, 1981; Li et al., 2003; Stark and Squire, 2000; Yamaguchi et al., 2004). For instance, conditioned taste aversions (CTA) and some forms of conditioned fear can be acquired in a single learning trial. Crucially, successful conditioning depends on the novelty of the conditioned stimulus (CS) (see Welzl et al., 2001, for a review). Pre-exposure to the CS severely diminishes associative learning (a.k.a. “latent inhibition”). Further, conditioning is also reduced if only some aspects of the CS are novel while others are familiar. The sensitivity to CS novelty, but not the taste aversion itself, is blocked by hippocampal lesions (Gallo and Candido, 1995). The novelty dependence of single-trial learning in the CTA paradigm is a good example of the importance of a rapid assessment of stimulus novelty or familiarity.

The medial temporal lobe (MTL) is crucial for the acquisition of declarative memories, and some functional imaging techniques have shown activation of MTL structures associated with either novel or familiar stimuli (Stark and Squire, 2000; Stern et al., 1996; Tulving et al., 1996; Yamaguchi et al., 2004). Lesion studies have repeatedly demonstrated that MTL damage impairs or abolishes behavioral, electrographic, and skin responses to novel stimuli (Kishiyama et al., 2004; Knight, 1996; Yonelinas et al., 2002). While these studies suggest a role of the MTL in novelty detection, the cellular basis for this discrimination has yet to be described. We report here that single neurons in the human MTL can alter their firing behavior to discriminate between novel and familiar complex stimuli following a single trial, thus exhibiting rapid plasticity as a result of single-trial learning.

Results

Task Paradigm and Behavioral Results

We recorded single neuron activity using microwires implanted in the human hippocampus-amygdala complex (Figures 1A and 1B; see Table S1 in the Supplemental Data for electrode locations) while subjects performed an object learning and recognition task. The delay between the learning and the initial recognition period was ~30 min, during which time the subject performed a different, cognitively demanding task. During learning, subjects were shown 12 different visual images. Each image was presented once, randomly in one of four quadrants on a computer screen (Figure 1C). Subjects were instructed to remember both the identity and the position of the image(s) presented. During the recognition period, subjects saw either previously viewed (familiar) or new images (novel) presented at the center of the screen (Figure 1D). For each image, the subject was asked to indicate whether the stimulus was new (novel) or old (familiar). Note that the novelty of a stimulus is only defined by whether it has been seen before or not (contextual). No other attributes of the stimulus changed. For each image identified as familiar, the subject was also asked to identify the quadrant in which the stimulus was originally presented (spatial recollection). Subjects correctly identified, on average, 88.5% ± 2.8% of all familiar and novel items during recognition (Figure S1). Subjects correctly recalled the quadrant location for 49.5% ± 8.0% of the familiar stimuli.

Neural Representations of Single-Trial Learning, Novelty, and Familiarity

We analyzed the response of every neuron recorded (total number of neurons across all subjects = 244) during the baseline, stimulus presentation, and poststimulus delay period. A neuron was considered selective if it exhibited an altered firing rate as a function of the stimulus (novel versus familiar) ($p < 0.05$, bootstrap, see Experimental Procedures) and as a function of the task (learning versus recognition phase). Neurons that increased their firing when exposed to novel versus

*Correspondence: schumane@caltech.edu

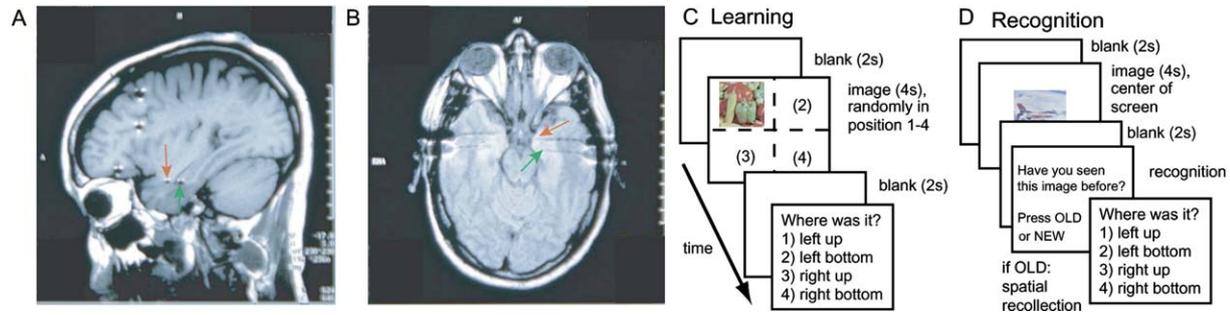


Figure 1. Electrode Placement and Task Design

(A) Saggital and (B) axial postimplantation structural MRI of one patient. The electrodes implanted in the amygdala (red) and the hippocampus (green) are indicated with arrowheads. The experiment has a learning (C) and recognition block (D). Learning trials consisted of 12 images presented in one of four quadrants on the screen. Two seconds after the stimulus was removed and replaced by a blank screen, the subject was asked to report in which quadrant the stimulus was presented. During recognition trials (30 min later), the subject was shown the 12 old images mixed with a set of 12 new images and was asked to indicate whether the image had been viewed before (old) or not (new). After classifying an image as “old,” the subject was also asked to indicate where the picture was during learning (spatial recognition).

familiar stimuli were classified as signaling “novelty,” whereas neurons that increased their firing to familiar stimuli were classified as signaling “familiarity” (Figure 2). Additionally, we classified responding neurons according to *when* they increased their firing: during the stimulus presentation of the stimulus or during the poststimulus period (Figure S2D). Note that neurons signaling novelty increased their firing to new stimuli during the learning phase *and* also increased their firing to new stimuli presented during the recognition phase.

Are individual neurons capable of signaling that learning has occurred? If this is the case, then once the subject learns something about a stimulus (e.g., that it has been seen before) the firing properties of the neuron should reflect this knowledge. In our task, any knowledge about whether the specific stimulus presented has been seen before must result from a single trial experience. We indeed found subsets of neurons that showed enhanced or depressed firing rates on the second of two stimulus presentations, indicating the capacity for single-trial learning of familiarity. There are two different patterns of responses we observed that indicate single-trial learning. One set of neurons (“familiarity detectors”) exhibited enhanced firing when previously viewed stimuli were presented a second time during the recognition phase of the experiment. An example of this type of response is shown in Figure 2 where the neuron does not exhibit any appreciable response to the stimuli when first presented (Figure 2C), but when these same stimuli are presented a second time a dramatic increase in firing rate was observed (Figure 2E). These cells, which form a class of “familiarity” detectors, thus exhibit single-trial learning, exhibiting memory for a stimulus that was presented only one time. The other class of cells increased firing only for the first presentation of the stimulus (“novelty detectors,” see Figure S4 for an example). All told, 40 neurons consistently signaled either novelty ($n = 23$) or familiarity ($n = 17$) (Figures 3A and 3B). To characterize the firing differences of all neurons, we used two measures: (1) average firing rate increase relative to baseline for new or old stimuli (depending on type of neuron) and (2) the average firing rate difference between new versus old stimuli. For both measures, spikes were counted in the entire

6 s period following stimulus onset. We found that neurons increase firing on average 47% relative to baseline, and the average firing difference between old versus new stimuli was 76% (Figure 3C). The larger difference when comparing new versus old firing indicates that in addition to increasing firing to the preferred stimulus (e.g., familiar), neurons decrease firing for the other stimulus type (e.g., novel). The large change in firing rate observed was induced by a single presentation of the stimulus, and as such, these neurons provide a potential source for the rapid single-trial memory exhibited behaviorally by the subjects.

Do the observed neuronal changes reflect either a priming or a habituation response, or alternatively, do they reflect a form of long-term memory? If the former is the case, one would expect that, if presented with the same familiar stimuli (as well as new stimuli) 24 hr later, the neuronal response to the familiar stimulus would be diminished. On the other hand, if the response reflects long-term memory, the altered firing pattern should still be observed the next day. To address this, we conducted a recognition session on the second and/or third day of recording, presenting subjects with the stimuli learned the previous day (four sessions total in three patients) as well as a new set of stimuli. The time delay between the learning and the second recognition session was ~24 hr (including one night of sleep). The behavioral performance (recognition and recollection) of these three patients did not differ significantly after a 30 min or 24 hr time delay. Unfortunately, single-unit microwire recordings do not allow one to unambiguously determine whether the same individual neurons can be recorded on two sequential days. As such, we asked whether individual neurons, recorded 30 min or 24 hr after the stimulus presentation showed differences in firing to old versus new stimuli. We then compared the average response strength per neuron after 30 min and 24 hr time delays. We found that neither the average response strength per neuron nor the average increase in firing rate relative to baseline (Figure 3D) differed significantly for the two different time delays (two-way ANOVA with groups neuron type [Novelty/Familiarity] and time delay [30 min/24 hr], $p < 0.05$). These neurons thus reflect the memory of the stimulus learned 24 hr earlier

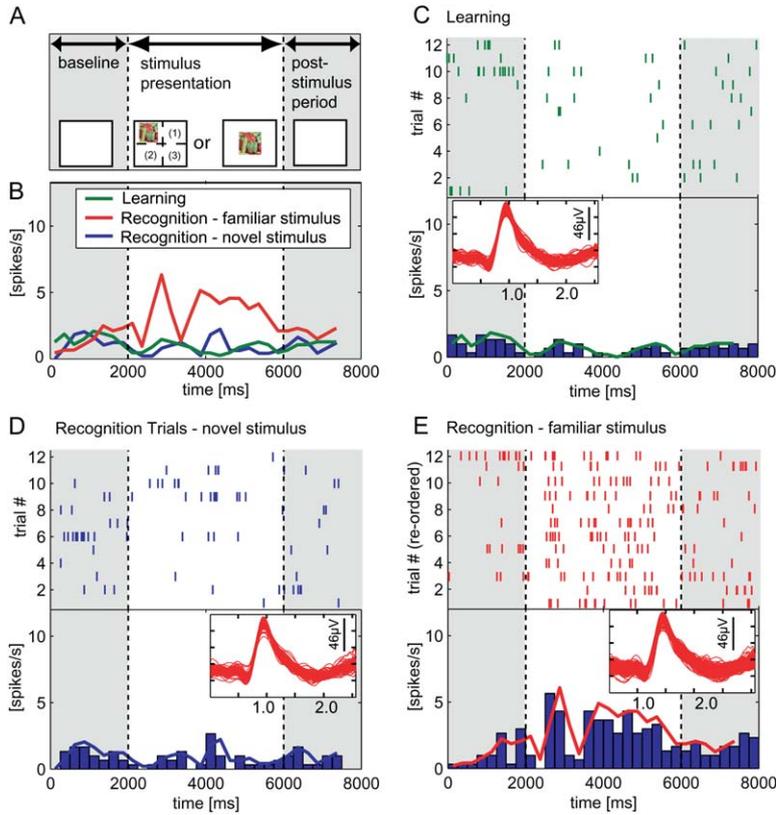


Figure 2. Response of a Single Hippocampal Neuron during Learning and Recognition

(A) Schematic representation of the experiment. Baseline (blank screen) from 0 to 2 s, stimulus presentation from 2 to 6 s, and post-stimulus period (blank screen) from 6 to 8 s. (B) Average responses (spikes/s). (C–E) The top portion of each figure shows the rasters depicting individual spikes. The stimulus was presented during the epoch defined by the dashed vertical lines. The bottom portion of each figure shows the binned histograms across all trials. Insets show overlays of all spike waveforms during the phase of the experiment depicted. (C) Responses during each learning trial. (D) Responses during the recognition phase for all new (not previously viewed) stimuli. (E) Responses during the recognition phase for all previously viewed (old) stimuli. Trials were randomly ordered during the experiment but are shown in (E) in the same order as during learning (C). This neuron increases its firing rate for stimuli seen before (E) but not for stimuli viewed for the first time (novel during both learning and recognition) (C and D). Note that in (C) and (E), the exact same visual stimuli are presented to the subject (12 images). When the stimuli are presented the first time (C), the neuron does not respond, whereas for the second presentation (E) it responds strongly.

but do not exhibit any further increases in firing rate (see Discussion). The majority of neurons (37 of 40) exhibited a significant response within the first 2 s after stimulus onset (Figure S3C). Does the response strength decrease as a function of trial number? We found that neither novelty nor familiarity neurons significantly reduce

their response strength over the duration of the experiment, during either learning or recognition (one-way ANOVA with block-nr and $p < 0.05$ reveals no significant effects for blocks of 1, 2, 3, or 4 trials). In addition, we found both types of neurons, familiarity and novelty detectors, in the amygdala as well as the hippocampus

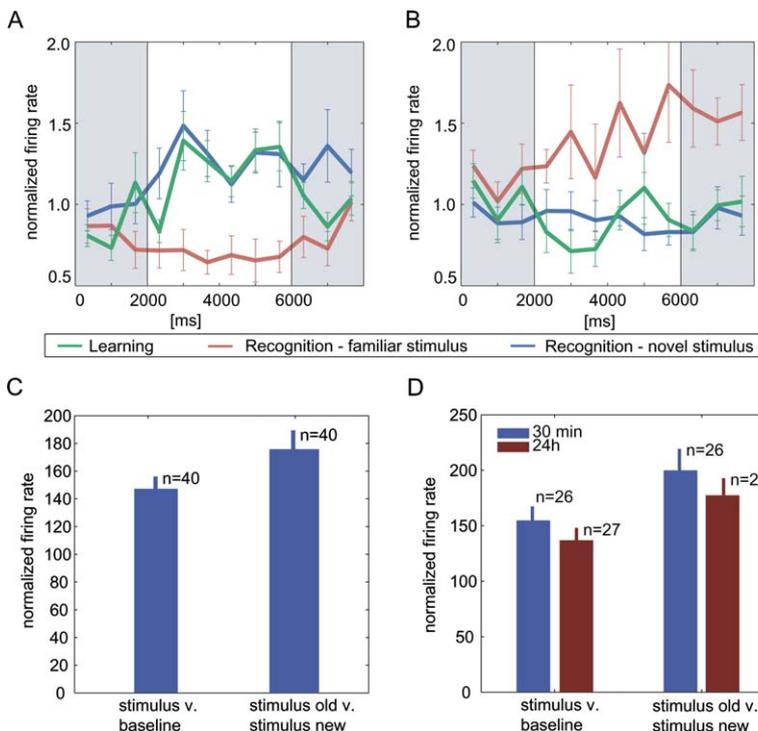


Figure 3. Population Summary of All Responding Neurons

Learning trials are in green, recognition old (familiar) trials are in red, and recognition new (novel) trials are in blue. We classified neurons based on the type of stimulus they increased their firing to (old or new) and when they increased their firing (during either the stimulus or poststimulus period or both). (A and B) Population average of all novelty ($n = 18$) and familiarity neurons ($n = 10$) that signal during the stimulus period. (C) Summary of response, quantified either as normalized firing rate difference during the 6 s poststimulus period for old versus new stimuli (right) or relative to baseline (left). Note that the average rate increase of 75% is the result of a single stimulus exposure—the stimulus is learned after one trial. (D) Comparison of response for different time delays between learning and recognition. Shown is the average response strength with 30 min and 24 hr delay. There is no significant difference in response strength for 30 min and 24 hr delay (t test, $p < 0.05$) nor is there a difference for novelty and familiarity detectors (data not shown, two-way ANOVA, $p < 0.05$). All error bars are \pm SE, and n specifies the number neurons.

(Figure S2). However, the overall incidence of these neurons was significantly less in the amygdala when compared to the hippocampus: $19.7\% \pm 4.9\%$ ($n = 11$) of all hippocampal neurons and $8.3\% \pm 2.7\%$ ($n = 12$) of all amygdala neurons were classified as either novelty or familiarity neurons (n is number of sessions, $p < 0.05$).

Single Neuron and Population Decoding

We analyzed how reliably these neurons can signal novelty or familiarity with an ideal-observer model. The model has access to the number of spikes fired during the 6 s period following stimulus onset. Using this information, a “decision” is made as to whether the subject is viewing a novel or a familiar stimulus. By parametrically varying the threshold (number of spikes) above which a single trial was considered novel or familiar, we conducted a receiver operator characteristic (ROC) analysis for each single neuron (Figure S3) and compared the true and false positives ratio at different thresholds. As a summary measure, we computed the area-under the curve (Britten et al., 1996), which is the probability of correctly predicting whether the subject is currently viewing a novel or familiar stimulus (probability is between 0 and 1.0; 0.5 represents chance performance). We found that our neurons have an average single-trial single-neuron prediction probability of 0.72 ± 0.02 . The population average is significantly above the chance level, which is determined by randomly shuffling the novel/familiar labels while keeping the spike trains intact. An observer that only has access to a single neuron’s firing pattern can thus predict with on average 72% success whether a subject is seeing a familiar or novel stimulus.

How much information does the population of all recorded neurons contain about the familiarity of a stimulus? While ROC analysis quantifies how much information a single neuron conveys about the stimulus, it remains to be investigated how well this information can actually be decoded from a population of neurons on a single-trial basis. Single trials are highly variable and noisy. Does combining multiple neurons allow more accurate decoding than observing only a single neuron? Only if the signal or the noise were uncorrelated among neurons would one expect an improvement in decoding accuracy.

To address these questions, we used a simple population decoder that has access to all simultaneously recorded neurons that were previously identified as signaling novelty or familiarity. The decoder does not know the identity (novelty or familiarity detector) of the neurons. The only information available to the decoder is the number of spikes each neuron fired in the 6 s period following stimulus onset. The weighted sum (Figure 4A) of all spike counts is used to predict whether, for a given trial, an old or new stimulus was presented. The weights are estimated from a set of labeled trials (old or new) using multiple linear regression (see Experimental Procedures).

We evaluated the properties of the classifier by considering only behaviorally correct recognition trials. For each recording session, we trained the classifier with all trials except a randomly chosen one (the “left out trial”). Afterward, we tested the classifier’s performance by using it to predict whether the left out trial was old or new. Repeating this procedure many times for each ses-

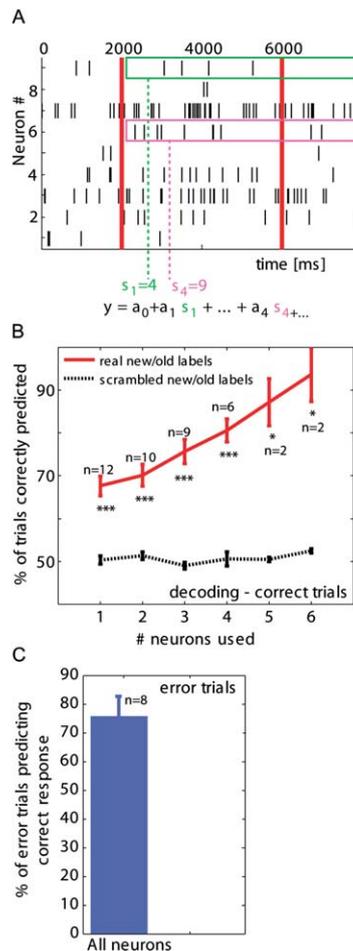


Figure 4. Population Decoding from Simultaneously Recorded Neurons

(A) Illustration of the decoding approach. Spikes of each neuron that signals novelty/familiarity (nine neurons in this example) are counted in the 6 s period following stimulus onset (first red line). Each neuron is assigned a weight determined by multiple linear regression. For a given trial, y predicts whether the trial is “old” or “new.” (B) Performance of the single-trial predictor as a function of the number of simultaneously recorded neurons. Decoding performance increases when information from multiple recorded neurons is considered. The number of neurons used for decoding has a significant effect on performance of the decoder (one-way ANOVA, $p < 0.001$). Error bars are SE, and n = the number of recording sessions. (C) The population decoder as trained in (B) applied to error trials. For 75% of all error trials in each session it predicts the correct response, that is, the neurons have better memory than the patient has behaviorally. The maximum number of available neurons is used for each session (mean number of neurons = 4.5). Only sessions that have at least two error trials are included (eight sessions). Error bars are SE per session ($n = 8$), and the mean per session is significantly different from chance ($p < 0.01$).

sion gives an accurate estimate of classifier performance (leave-one-out cross validation, see Experimental Procedures). Additionally, we restricted the number of neurons that the classifier has access to. We found that the average single-trial classification performance increases from 67% correct for one neuron to 93% when six simultaneously recorded neurons are considered (Figure 4B, red line). A one-way ANOVA reveals a significant effect of number of neurons ($F = 6.6$, $p = 0.0001$). Repeating the same procedure using randomly

scrambled labels for the test trial results in a chance (50%) level performance (Figure 4B, black line). This analysis shows that it is beneficial for an “ideal” decoder to look at multiple neurons simultaneously. This indicates that the spikes fired by individual neurons signaling familiarity are uncorrelated in the sense that each of them contributes additional information that can be used to increase the accuracy of decoding.

Relations between Neural Responses and Behavior

What is the relationship between the familiarity/novelty responses of individual neurons and the behavioral performance of the subject? The neuronal activity associated with behavioral errors allows us to answer this question. In our experiments, there were two kinds of error trials: (1) recognition (novel versus familiar) errors and (2) spatial recollection (which quadrant?) errors. Below, we investigate each type of error separately, beginning with spatial recollection errors.

There have been conflicting accounts as to whether retrieval-related activity in the hippocampus is related to familiarity recognition or recollection (Cameron et al., 2001; Stark and Squire, 2000; Yonelinas et al., 2002). One hypothesis states that the hippocampus is not involved in the retrieval of pure recognition memory, that is, memory without a recollective component. To investigate this issue, we examined neural activity during trials with successful recognition but failed recollection (spatial location of stimulus). We found that the subsequent successful spatial recollection is not required for neurons to exhibit familiarity responses. In fact, we observe novelty and familiarity selective neurons in subjects who perform at chance levels for spatial recollection: In four (of 12) sessions, spatial recollection performance was at chance ($21.7\% \pm 15.8\%$), and yet we found that 12 of the total 68 recorded units (17%) signaled novelty or familiarity. Thus, despite the fact that these patients weren't able to correctly recollect the spatial location in any of the trials, the same percentage of cells signaled novelty as in the other sessions. Also, for the sessions in which spatial recollection performance was above chance, we repeated our analysis including only trials associated with failed spatial recollection. Of the original 30 neurons, 26 remained significant (see [Experimental Procedures](#) for details). We thus conclude that successful recollection is not required to observe a novelty/familiarity response in the hippocampus.

How is the neuronal activity during the stimulus presentation related to errors in recognition? Recognition of pictures is a highly automatic and reliable form of memory, and subjects are usually very confident in their responses. This results in a small number of errors even when a large stimulus set is used, which has prevented analysis of such error trials in the past (Xiang and Brown, 1998). In our experiments, however, we record from many neurons simultaneously and can thus use a population decoder that allows accurate single-trial decoding (see discussion above). For each recording session, we trained the population decoder using all behaviorally successful trials. Afterward, we used it to investigate what it would predict for the spiking activity observed during error trials. What might the population decoder (classifier) predict for an error trial? The classifier could

(1) be at chance, (2) mimic the subject's (incorrect) response, or (3) predict the (correct, but not chosen) response. Each outcome would be informative: (1) if it is at chance, these neurons do not contain any information about the stimulus on error trial; (2) if it predicts the behavioral response given, these neurons would likely represent some form of decision taken by the patient or motor planning activity related to the key the patient used to indicate the response; (3) if it predicts the correct response, these neurons would likely represent some form of high-fidelity memory. The third possibility is intriguing because it would suggest that these neurons exhibit “better memory” than the subject's behavioral response indicated. Since we are interested in the fraction of error trials per session that predict a certain outcome, we consider only sessions that contain at least two error trials (8 out of 12 sessions with a total of 33 error trials). For each session, we trained a classifier with all available neurons (on average 4.5) that signaled novelty/familiarity using all behaviorally correct trials and used it to predict the outcome of each error trial. We find that the classifier predicts the actual correct response for $75\% \pm 7\%$ of all error trials. The classifier is thus able to correctly predict the correct response in 75% of all cases even when the subject responded incorrectly (Figure 4C). These neurons thus have better memory than the patient exhibited behaviorally. This also suggests that the neuronal activity reported here does not represent some form of motor activity related to the subject's intended or actual response.

Discussion

Novelty and Familiarity Detectors in the Human Brain

We identified single neurons in the human hippocampus and amygdala that signal novelty or familiarity with an increase in firing rate. Several other groups have described nonhuman primate neurons that gradually (over many trials) decrease their response magnitude as specific stimuli become more familiar (Asaad et al., 1998; Fahy et al., 1993; Li et al., 1993; Rainer and Miller, 2000; Rolls et al., 1993). These types of neurons have also been observed in rodents (Berger et al., 1976). The opposite pattern, neurons that increase their response magnitude for familiar stimuli, have largely not been observed in the primate brain (Fahy et al., 1993; Heit et al., 1990; Rolls et al., 1993; Xiang and Brown, 1998) and only rarely in humans (Fried et al., 1997). Also, studies investigating the relative proportion of novelty/familiarity-selective neurons in different areas of the MTL have usually failed to find any such neurons in the nonhuman primate hippocampus (Riches et al., 1991; Xiang and Brown, 1998) or, in one case, found only a very small proportion of such cells (Rolls et al., 1993). In contrast, we found a large proportion (17%) of familiarity/novelty-sensitive neurons, with an approximately equal number of neurons that increased firing for novelty or familiarity in the human hippocampus and amygdala. It has been speculated that the apparent absence of novelty/familiarity neurons in the primate hippocampus can be attributed to the lack of a spatial component in the tasks used (Riches et al., 1991; Xiang and Brown, 1998). We found that the responses observed do not depend on successful spatial recollection. Another

crucial difference is the behavioral task. Our task consists of a learning and recognition block with an interposed time delay. During the delay, other tasks are conducted. Others have used a serial recognition task where learning and recognition trials are intermixed, and as such, there is no time delay that would permit a diversion of cognitive resources. It is possible that the emergence of the neuronal response requires time to develop. In our experiments, the firing rate increase can be observed after an initial delay of 30 min and remains equally strong for at least 24 hr. This indicates that these neuronal responses represent some form of long-term memory. Also note that the response strength does not increase further between 30 min and 24 hr delays. The ability to correlate neuronal responses with human behavior may also be critical: we used an abstract task that can be rapidly learned, thus facilitating the detection of these rapidly changing neuronal responses. In contrast, in nonhuman primates a simple associative memory task can take many trials for animals to reach criterion, and learning-induced changes in hippocampal activity show a similar prolonged temporal profile (Wirth et al., 2003).

Could it be that the different findings are caused by eye movements? Most primate studies require the animal to fixate. In our experiments, subjects are free to move their eyes as they like. This is to make the task as natural as possible. Owing to clinical constraints, we were unable to record eye movements, but there are several pieces of evidence that argue that eye movements cannot explain our results. The first few fixations made on any picture are mostly dominated by the statistics of the stimulus and do not change as a function of the familiarity of the stimulus (Noton and Stark, 1971). Also, a previous study of human MTL neurons found no influence of the fixated location of the picture on the visual response properties (Kreiman et al., 2002).

Others have reported that some neurons in the human MTL (Kreiman et al., 2000) and the primate cortex (Li et al., 1993) are sharply tuned to the visual category of stimuli. Here, we used stimuli from many different visual categories with one example per category. While the small stimulus set required for this kind of memory experiment prevents us from testing large numbers of stimuli from different categories, the response observed is invariant to at least a majority of the visual categories we have used. Thus, the neuronal responses we describe here are capable of signaling the familiarity of the stimulus regardless of its visual category. They could thus serve as “general” novelty detectors that establish the significance of behavioral stimuli during the acquisition of new or consolidation of existing memories (Lisman and Otmakhova, 2001).

Recognition and recollection are two largely distinct memory processes. Importantly, we find that the response to the second presentation of the stimulus does not depend on whether spatial recollection is successful. This is in agreement with an earlier study of recollective memory which found that recall success is not correlated with the response of hippocampal neurons (Cameron et al., 2001). Also note that Cameron et al. (2001) used the same stimuli many times during learning, so that the resulting neuronal changes cannot be related to any specific stimulus presentation.

Neurons that Remember Better than Subjects

The finding that the neuronal activity during a majority of the error trials predicts the correct response represents an interesting disassociation between behavior and neuronal activity. If one examines the successful recognition trials exclusively, one might conclude that the neuronal responses represent the outcome of the decision taken (is it old or new) or a consequence of that decision, e.g., planning and/or pre- or post-motor activity. If this were the case, however, activity during error trials would have to predict the response that was actually observed. However, we observed the opposite: activity during error trials predicts the correct response. We thus conclude that the neuronal responses reported here represent some form of memory. In addition, the proportion of trials correctly identified by the neuronal responses is higher than what we observed behaviorally. Our data do not address at what point in the circuit the accurate neuronal responses on error trials fail to translate into correct behavioral responses. However, it is likely that information from multiple brain areas must be integrated to decide about the novelty of a stimulus. Any system of this nature requires an internal threshold for what is considered sufficient cumulative evidence for a stimulus to be classified as familiar. One could thus imagine situations where some brain areas provide input indicating familiarity but the cumulative evidence does not pass this threshold. Such a system would be maximally robust because it integrates multiple sources of information, perhaps trusting some more than others (Pouget et al., 2003).

It has previously been observed that the average firing rate of some MTL neurons differs for successful versus unsuccessful retrieval (Fried et al., 1997, 2002). However, in these studies, activity of the same neuron was not recorded during learning, and it has thus remained impossible to determine whether these neurons changed their firing as a function of previous stimulus exposure or as a function of the task. In contrast, here we demonstrate that these changes result from a single stimulus exposure.

Relationship to fMRI and ERP Findings

It has proven difficult to find human MTL fMRI activity correlated with behavioral success in recognition memory tasks (Manns et al., 2003; Stark and Squire, 2000). Using single-unit recordings we find evidence for the coexistence of novelty and familiarity cells recorded at the same time in the same brain region. On half of all macroelectrodes (18 of 36), we detected both novelty and familiarity neurons. On 2 of 6 microwires with more than one novelty/familiarity neuron both types were found. Since fMRI methods have limited spatial and temporal resolution and often rely on subtractive techniques, it is likely that the presence of both classes of neurons prevented their detection (Logothetis et al., 2001).

Scalp and intracranial event-related potentials (ERP) recorded during a serial recognition task have revealed a prominent potential (P300) to novel as well as target stimulus items (McCarthy et al., 1989; Sutton et al., 1965). That is, there is a potential to both novel as well as familiar (task relevant) items, but not to distractors. In hippocampal lesion patients, it has been observed that the P3a component of the P300 is reduced (Knight,

1996). While we did not record ERPs in this study, the P300 response has been observed previously with intracranial electrodes in similar locations (McCarthy et al., 1989). It is thus of interest to note that the identified subpopulations of novelty and familiarity neurons we identified here could contribute to the P300.

Interaction with Other Brain Systems

What is driving the response of these neurons? Neurons from multiple other brain areas can signal novelty or, more generally, the behavioral relevance of stimuli encountered in the environment. These include noradrenergic neurons in the locus coeruleus, cholinergic neurons in the basal forebrain, as well as dopaminergic neurons in the midbrain (see Schultz and Dickinson, 2000, for a review). Their response to novel events habituates with brief delays, evidence for short-term memory. Common to all these areas is the modulatory nature of their output—it is thus unlikely that their output is sufficient to account for the MTL responses we observe. These modulatory systems are known to regulate the strength of hippocampal-dependent learning, however (Frey et al., 1990; Neuman and Harley, 1983; Williams and Johnston, 1988), raising the possibility that the rapid plasticity we describe is related to the simultaneous release of neuromodulators that help induce long lasting memories.

It is well known that animal behavior can be modified by a single exposure to a relevant stimulus (Sokolov, 1963). One instance of such memory is episodic memory, which is, by definition, memory of a single experience (Tulving et al., 1996). Other instances of single-trial learning include object recognition (Standing et al., 1970), spatial learning, and food caching (Clayton et al., 2001). In contrast, other forms of learning, like classical conditioning or rule learning (Wirth et al., 2003; Yanike et al., 2004), require many learning trials. The neurons that underlie or participate in the rapid behavioral plasticity have, for the most part, evaded detection. Here we find that MTL neurons exhibit remarkable plasticity: a single exposure to a stimulus was sufficient to induce a dramatic and significant change in the spiking pattern. The observation of single-trial learning in MTL neurons indicates that, at least in principle, the rapid learning that human subjects exhibit has an electrophysiological correlate that occurs at the level of individual neurons.

Experimental Procedures

Subjects and Electrophysiology

Subjects were six patients (three male, three female; mean age 37.5 ± 5.5 years; all native English speakers) diagnosed with drug-resistant temporal lobe epilepsy and implanted with intracranial depth electrodes to record intracranial EEG and single-unit activity. Patients underwent stereotactic placement of hybrid depth electrodes containing both clinical field potential contacts and micro-wire (50 μm) single-unit contacts, as described by Fried et al., 1999. Briefly, electrodes were placed using orthogonal trajectories through the dorsolateral cortex, with the tip of the electrode targeting the amygdala, anterior hippocampus, orbitofrontal region, supplementary motor area, or anterior cingulate gyrus. The commercially available electrodes (Behnke hybrid depth electrode, Adtech Inc, Racine, MN) contain four to six platinum-iridium 5 mm long circular electrodes, with a hollow center. After insertion of the electrode to the target, the inner cannula was removed, and a bundle

of microwires was passed through the center of the electrode, extending 5 mm beyond the tip of the electrode in a “flower spray” design. The electrodes were secured in place via a skull anchor bolt. All electrodes were placed based on clinical criteria alone. Patients were recruited for the research study after surgery was completed and EEG monitoring was initiated. Participation was voluntary and patients could withdraw from the study at any time. Informed consent was obtained, and the protocol was approved by the Institutional Review Boards of the Huntington Memorial Hospital and the California Institute of Technology. For further details regarding the electrophysiological recordings, please see the Supplemental Data.

Data Analysis

Spikes were sorted with a template-matching method (Rutishauser et al., 2006). Only well separated single neurons were used (see the Supplemental Data for details). We used a nonparametric bootstrap statistical test (Efron and Tibshirani, 1993) to assess significance at $p < 0.05$ (see the Supplemental Data for a discussion of why a t test was not used). To determine whether a neuron responds to new or old stimuli, we compared the number of spikes fired for old versus new stimuli during the stimulus on (4 s) and the poststimulus (2 s) period. For bootstrapping, 10,000 randomly resampled (with replacement) sets of spike counts were generated and tested for equality of means (Efron and Tibshirani, 1993). A second statistical test was performed to determine whether the firing of a neuron between old stimuli during recognition and all stimuli during learning (which are, by definition, new) was different. Only if both statistical tests were passed with $p < 0.05$ was the neuron determined to function as a novelty or familiarity detector. We randomly shuffled the start/endpoints of trials (in time) while keeping everything else the same to establish chance performance for this statistical procedure. We repeated this procedure ten times and found a chance performance of 4.4% of all neurons (Figure S2D). Error trials during learning (incorrect position) and recognition (new/old wrong) were excluded from this analysis.

All errors are standard error (SE), unless noted otherwise.

Population Analysis

To quantify how well we were able to decode information about the novelty of the stimulus for a single trial, we used a population decoder. This also allowed us to analyze whether and how the decoding performance depends on the number of simultaneously recorded neurons. We used a simple weighted sum classifier of the form $y = a_0 + a_1s_1 + \dots + a_n s_n$, where s_x represents the number of spikes in the 6 s period following stimulus onset for neuron x and a_x is the weight of this neuron. The weights are determined from labeled training data using multiple linear regressions (Johnson and Wichern, 2002). The label y is either set to 1 (new) or -1 (old). Only neurons that were previously found to be signaling novelty/familiarity were considered for this analysis.

For verification purposes, we trained the classifier on behaviorally correct trials using leave-one-out cross validation. The performance of this classifier was then verified by evaluating its prediction for the left out trial. Repeating this procedure many times gives an accurate estimate of the true performance of the estimator. We repeated the same analysis by restricting the number of neurons the classifier had access to. In cases where more neurons were available than the classifier could consider, a random subset of the available neurons was chosen and the procedure was repeated multiple times so that all possible combinations were explored. All error bars in the population analysis are given as SE, with n being the number of sessions, to demonstrate the variance over multiple patients and recording sessions rather than over multiple neurons.

Supplemental Data

The Supplemental Data for this article, including Supplemental Experimental Procedures, can be found online at <http://www.neuron.org/cgi/content/full/49/6/805/DC1/>.

Acknowledgments

We thank all patients for their participation in this study. We would like to acknowledge the support we have received from the Staff of the Epilepsy and Brain Mapping Department at the Huntington

Memorial Hospital. We thank Farshad Moradi, Melissa Saenz, and Dinah Thylerlei for advice regarding the processing of structural MRIs. This research was funded by the Howard Hughes Medical Institute and the Gimbel Discovery Fund.

Received: December 15, 2005

Revised: January 27, 2006

Accepted: February 9, 2006

Published: March 15, 2006

References

- Asaad, W.F., Rainer, G., and Miller, E.K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21, 1399–1407.
- Berger, T.W., Alger, B., and Thompson, R.F. (1976). Neuronal substrate of classical conditioning in the hippocampus. *Science* 192, 483–485.
- Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebri, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13, 87–100.
- Cameron, K.A., Yashar, S., Wilson, C.L., and Fried, I. (2001). Human hippocampal neurons predict how well word pairs will be remembered. *Neuron* 30, 289–298.
- Clayton, N.S., Griffiths, D.P., Emery, N.J., and Dickinson, A. (2001). Elements of episodic-like memory in animals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 356, 1483–1491.
- Davis, C.D., Jones, F.L., and Derrick, B.E. (2004). Novel environments enhance the induction and maintenance of long-term potentiation in the dentate gyrus. *J. Neurosci.* 24, 6497–6506.
- Efron, B., and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap* (London: Chapman & Hall).
- Fahy, F.L., Riches, I.P., and Brown, M.W. (1993). Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Exp. Brain Res.* 96, 457–472.
- Frey, U., Schroeder, H., and Matthies, H. (1990). Dopaminergic antagonists prevent long-term maintenance of posttetanic LTP in the CA1 region of rat hippocampal slices. *Brain Res.* 522, 69–75.
- Fried, I., MacDonald, K.A., and Wilson, C.L. (1997). Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18, 753–765.
- Fried, I., Wilson, C.L., Maidment, N.T., Engel, J., Behnke, E., Fields, T.A., MacDonald, K.A., Morrow, J.W., and Ackerson, L. (1999). Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients - Technical note. *J. Neurosurg.* 91, 697–705.
- Fried, I., Cameron, K.A., Yashar, S., Fong, R., and Morrow, J.W. (2002). Inhibitory and excitatory responses of single neurons in the human medial temporal lobe during recognition of faces and objects. *Cereb. Cortex* 12, 575–584.
- Gallo, M., and Candido, A. (1995). Reversible inactivation of dorsal hippocampus by tetrodotoxin impairs blocking of taste aversion selectively during the acquisition but not the retrieval in rats. *Neurosci. Lett.* 186, 1–4.
- Heit, G., Smith, M.E., and Halgren, E. (1990). Neuronal activity in the human medial temporal lobe during recognition memory. *Brain* 113, 1093–1112.
- Johnson, R.A., and Wichern, D.W. (2002). *Applied Multivariate Statistical Analysis* (New York: Prentice Hall).
- Kishiyama, M.M., Yonelinas, A.P., and Lazzara, M.M. (2004). The von Restorff effect in amnesia: the contribution of the hippocampal system to novelty-related memory enhancements. *J. Cogn. Neurosci.* 16, 15–23.
- Knight, R. (1996). Contribution of human hippocampal region to novelty detection. *Nature* 383, 256–259.
- Kohonen, T., and Lehtio, P. (1981). Storage and Processing of Information in Distributed Associative Memory Systems. In *Parallel Models of Associative Memory*, G. Hinton and J. Anderson, eds. (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.), pp. 105–143.
- Kreiman, G., Koch, C., and Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.* 3, 946–953.
- Kreiman, G., Fried, I., and Koch, C. (2002). Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proc. Natl. Acad. Sci. USA* 99, 8378–8383.
- Li, L., Miller, E.K., and Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *J. Neurophysiol.* 69, 1918–1929.
- Li, S., Cullen, W.K., Anwyl, R., and Rowan, M.J. (2003). Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nat. Neurosci.* 6, 526–531.
- Lisman, J.E., and Otmakhova, N.A. (2001). Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine. *Hippocampus* 11, 551–568.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.
- Manns, J.R., Hopkins, R.O., Reed, J.M., Kitchener, E.G., and Squire, L.R. (2003). Recognition memory and the human hippocampus. *Neuron* 37, 171–180.
- McCarthy, G., Wood, C.C., Williamson, P.D., and Spencer, D.D. (1989). Task-dependent field potentials in human hippocampal formation. *J. Neurosci.* 9, 4253–4268.
- Neuman, R.S., and Harley, C.W. (1983). Long-lasting potentiation of the dentate gyrus population spike by norepinephrine. *Brain Res.* 273, 162–165.
- Noton, D., and Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science* 171, 308–311.
- Pouget, A., Dayan, P., and Zemel, R.S. (2003). Inference and computation with population codes. *Annu. Rev. Neurosci.* 26, 381–410.
- Rainer, G., and Miller, E.K. (2000). Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27, 179–189.
- Riches, I.P., Wilson, F.A., and Brown, M.W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *J. Neurosci.* 11, 1763–1779.
- Rolls, E.T., Cahusac, P.M., Feigenbaum, J.D., and Miyashita, Y. (1993). Responses of single neurons in the hippocampal formation related to recognition memory. *Exp. Brain Res.* 93, 299–306.
- Rutishauser, U., Schuman, E., and Mamelak, M. (2006). Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *J. Neurosci. Methods*, in press. Published online February 17, 2006. 10.1016/j.jneumeth.2005.12.033.
- Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23, 473–500.
- Sokolov, E.N. (1963). Higher nervous functions; the orienting reflex. *Annu. Rev. Physiol.* 25, 545–580.
- Standing, L., Conezio, J., and Haber, R.N. (1970). Perception and memory for pictures - single-trial learning of 2500 visual stimuli. *Psychon. Sci.* 19, 73–74.
- Stark, C.E., and Squire, L.R. (2000). Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. *J. Neurosci.* 20, 7776–7781.
- Stern, C.E., Corkin, S., Gonzalez, R.G., Guimaraes, A.R., Baker, J.R., Jennings, P.J., Carr, C.A., Sugiura, R.M., Vedantham, V., and Rosen, B.R. (1996). The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proc. Natl. Acad. Sci. USA* 93, 8660–8665.
- Sutton, S., Braren, M., Zubin, J., and John, E.R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science* 150, 1187–1188.

- Tulving, E., Markowitsch, H.J., Craik, F.E., Habib, R., and Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cereb. Cortex* 6, 71–79.
- Welzl, H., D'Adamo, P., and Lipp, H.P. (2001). Conditioned taste aversion as a learning and memory paradigm. *Behav. Brain Res.* 125, 205–213.
- Williams, S., and Johnston, D. (1988). Muscarinic depression of long-term potentiation in CA3 hippocampal neurons. *Science* 242, 84–87.
- Wirth, S., Yanike, M., Frank, L.M., Smith, A.C., Brown, E.N., and Suzuki, W.A. (2003). Single neurons in the monkey hippocampus and learning of new associations. *Science* 300, 1578–1581.
- Xiang, J.Z., and Brown, M.W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology* 37, 657–676.
- Yamaguchi, S., Hale, L.A., D'Esposito, M., and Knight, R.T. (2004). Rapid prefrontal-hippocampal habituation to novel events. *J. Neurosci.* 24, 5356–5363.
- Yanike, M., Wirth, S., and Suzuki, W.A. (2004). Representation of well-learned information in the monkey hippocampus. *Neuron* 42, 477–487.
- Yonelinas, A.P., Kroll, N.E., Quamme, J.R., Lazzara, M.M., Sauve, M.J., Widaman, K.F., and Knight, R.T. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nat. Neurosci.* 5, 1236–1241.