Supporting Information

Wixted et al. 10.1073/pnas.1408365111

SI Materials and Methods

Microwire Implantation and Recordings. Electrode implantation was performed stereotactically (Medtronic StealthStation), and the position was confirmed by coaligning the postoperative CT or MRI (using the Statistical Parametric Mapping toolkit, www.fil.ion.ucl.ac.uk/spm/) with the preoperative structural MRI. This procedure localizes the tips of the microwires to within 2 mm (1). Bundles of nine platinum-iridium microwires 38 μm in diameter (California Fine Wire) were introduced through a lumen within the clinical intraparenchymal electrode during surgery. The implantation sites were chosen according to clinical criteria, limiting the potential recording sites. For the nine patients studied here, the sites included the hippocampus and amygdala, bilaterally. In the hippocampus, the wires usually were targeted to be in the midbody of the hippocampus, just behind the head of the hippocampus, opposite the apex of the cerebral peduncle. All patients received a postimplantation CT scan as a check to ensure there was no bleeding after the operation. These scans do not have sufficient resolution to resolve hippocampal subfields.

The extracellular potentials corresponding to single-unit activity and multiunit activity were recorded from the tips of the microwires. At each site, the potential difference between eight of the microwires was recorded relative to a ninth microwire in the same bundle using a headstage amplifier of custom design. This amplifier provided a 400× gain and was connected to signal-conditioning electronics and analog-to-digital converters (model DT9834; Data Translation) via a 1-m tether cable. Each signal channel was preconditioned with a high-pass filter (0.5-Hz corner) followed by a 10-kHz antialiasing filter and a computer-controlled 1–16× adjustable gain amplifier (custom-designed signal-conditioning board). The conditioned signal was digitized at 29,412 Hz with 16-bit resolution.

Data Analysis. Possible action potentials (APs) were detected by filtering twice (forward and backward, acausally) with a 24th-order digital IIR bandpass filter, 300–3,000 Hz, with a −100-dB stop band and −12-dB notches at 1, 2, and 3 kHz followed by a two-sided threshold detector (threshold 2.8 times each channel's SD) to identify AP times. The original signal then was high-pass filtered (100 Hz, single-pole Butterworth, applied causally) to capture the shape of the AP waveform in windows of 32 samples (1.1 ms) with the absolute peak value aligned at the ninth sample.

Because more than one neuron may be recorded near any given electrode, APs were grouped into several clusters of similar waveform shape. This clustering was performed using the open-source clustering program KlustaKwik (Klustakwik.sf.net), which is a modified implementation of the Classification Expectation-Maximization clustering algorithm (2). The first principal component of all waveform shapes recorded from a channel was the waveform feature used for sorting. After sorting, each cluster was graded as being noise, multiunit activity, or single-unit activity based on criteria including the waveform shape, size of the waveform relative to noise, evidence of a refractory interval, and lack of powerline interference, as previously described (3). Fig. S3 illustrates a typical cluster of single-unit activity after spike sorting.

SI Results

In the hippocampus, the mean normalized spike count in response to targets (μ_{Target}) averaged across the nine patients was 0.11 [which marginally exceeded the baseline value of 0, $t(8) = 2.18$, $P = 0.061$], and the mean normalized spike count in response to foils (μ_{Foil}) averaged across the nine patients was 0.04 (which was not significantly different from 0, $P = 0.43$). As noted in the main text, however, the average difference score, D', was significantly greater than 0 ($P < 0.01$). In the amygdala, the mean normalized spike count in response to targets (μ_{Target}) averaged across the nine patients was 0.13, and the mean normalized spike count in response to foils (μ_{Foil}) averaged across the nine patients was 0.11. Neither value was significantly different from 0, and the average difference score, D', also was not significantly different from 0 ($P > 0.16$).

Our analyses were based on all recorded clusters, not on a subset of clusters that were deemed to be task-relevant based on any indication of responsiveness to the study or test stimuli. However, because such subset analyses have become common practice, we also asked whether a higher percentage of significant clusters would be identified when the analysis was limited to only those clusters deemed to be responsive based on a significant change in average spike counts (relative to baseline) across the 32 items presented during the study phase. Using an alpha level of 0.10 to identify responsive units, we found that 15 of 205 clusters exhibited a significant change in firing relative to the prestimulus baseline during the initial presentation of the list items, a number that is not greater than would be expected from chance (expected = 0.10 × 205 = 20.5) (Although 220 clusters were analyzed during the recognition test, only 205 of those clusters yielded spike counts during study). Thus, we did not find evidence of neurons that were generally responsive to study items.

The quantile–quantile (Q–Q) plots for all clusters combined (Fig. 4*E*) and for the single units considered separately (Fig. 4*F*) show points associated with normalized spike counts of 10 or less. Five of the 1,088 values from single units (0.5%) exceeded 10 and were deemed to be too large to reflect true responses (and therefore were excluded from analysis). Some of the values less than 10 also may reflect measurement error. For example, the apparent return to the diagonal in Fig. 4*E* reflects the fact that, on rare occasions, extremely high values (e.g., 7 SDs or more above baseline firing) occurred for both targets and foils with approximately equal frequency. It is not clear how these extremely high scores should be interpreted, but they may simply reflect measurement error (which would occur for targets and foils with equal frequency).

The overall pattern of results is consistent with the bimodal target distribution predicted by the sparse distributed account, but the same Q–Q pattern could be generated by a continuous target distribution that is extremely skewed compared with the foil distribution. An extremely skewed distribution would reflect the strong response generated by a few items in the tail of the target distribution. This distribution, too, would correspond to sparse distributed coding. That is, no matter whether the data reflect a bimodal target distribution or an extremely skewed unimodal distribution, the results are consistent with the idea that a small percentage of targets generated a strong response in a small fraction of hippocampal neurons. Note that visual evidence for a bimodal target distribution was apparent when normalized spike counts were examined, but not when raw spike counts were examined. This is not surprising because large differences in the baseline firing rates of the recorded units swamp any evidence of the modestly elevated firing that occurred in response to a few targets for each unit.

1. Mehta AD, et al. (2005) Frameless stereotactic placement of depth electrodes in epilepsy surgery. *J Neurosurg* 102(6):1040–1045.
2. Celeux G, Govaert G (1995) Gaussian parsimonious clustering models. *Pattern Recognition* 28(5):781–793.
3. Valdez AB, Hickman EN, Treiman DM, Smith KA, Steinmetz PN (2013) A statistical method for predicting seizure onset zones from human single-neuron recordings. *J Neural Eng* 10(1):016001, 10.1088/1741-2560/10/1/016001.

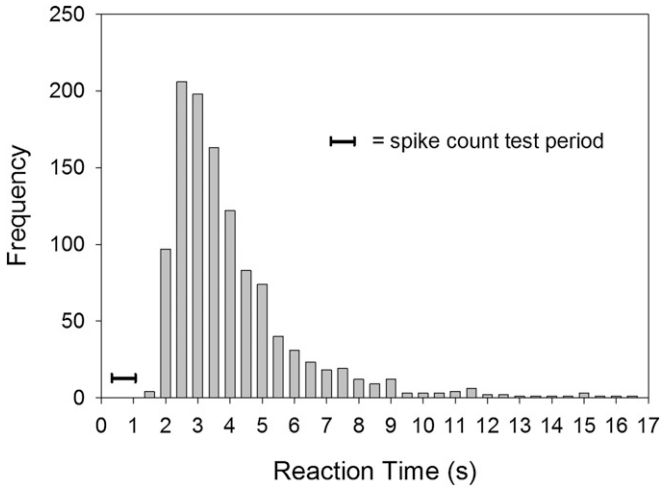

Fig. S1. Reaction time (RT) frequency distribution pooled over 18 recognition tests, where RT = the interval between the onset of a test item and the mouse click indicating the confidence rating for that test item. Spike counts were recorded from 200–1,000 ms after stimulus onset, before all overt responses.

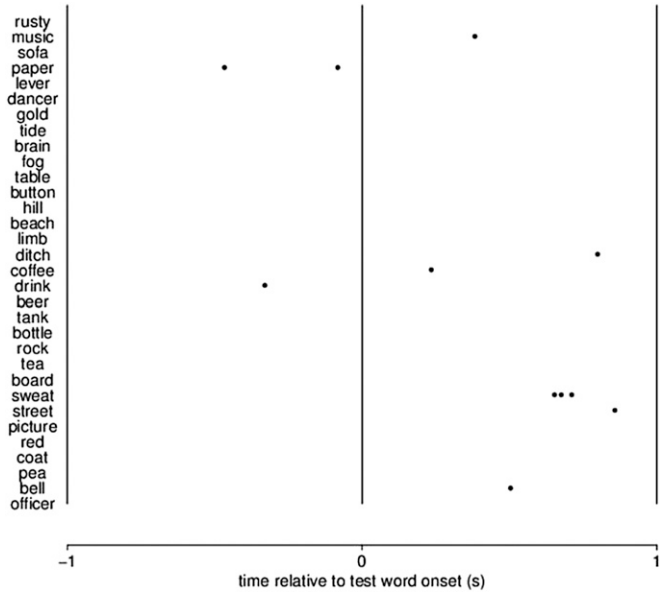

Fig. S2. Representative example of a raster plot of prestimulus and poststimulus activity for one of the single units that was responsive to one target. Although this unit yielded a strong normalized response to the target item "sweat" ($z = 7.67$), it yielded only a moderately strong response when measured in absolute terms. More specifically, three spikes occurred in response to that target during the 800-ms test period, reflecting a modest elevation in firing for a unit that had a baseline spike count mean and SD of 0.11 and 0.36, respectively. This result is fairly typical of the 30 targets that yielded notably elevated (off-diagonal) responses evident in the single-unit Q–Q plot (Fig. 4*F*). These results suggest that episodic memory may not be characterized by the kind of conspicuously elevated single-unit responding (in absolute terms) that is observed when single units exhibit an elevated response to repeatedly presented images of famous people and landmarks (e.g., ref. 1). It is important to emphasize that, using our single-presentation design, no single instance of bursting that coincides with the presentation of a test item (such as the response to "sweat" illustrated here) can be attributed confidently to the presentation of that item. The raster plot is intended only to illustrate the kind of bursting that occurs significantly more often in response to targets than to foils, not to suggest that the bursting that occurred in conjunction with the word "sweat" on this particular trial was necessarily triggered by the presentation of that word.

1. Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435(7045):1102–1107.

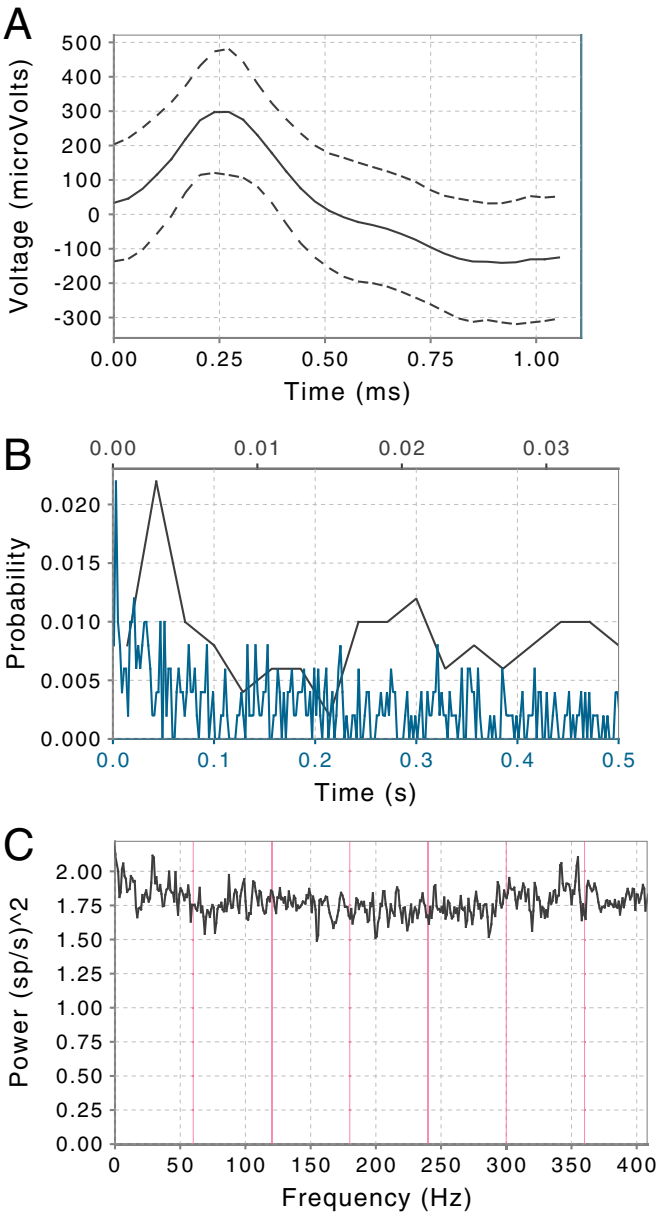

Fig. S3. Waveforms in a cluster identified as single-unit activity after sorting. This cluster was recorded from the left hippocampus. (*A*) The *y* axis shows the waveform shape on a scale of −300 to 500 μV. Dashed lines indicate ±1 SD at each sample point. (*B*) Distribution of interspike intervals on two time scales. The *y* axis shows the probability of interval. On the *x* axis the lower (blue) trace shows the duration of the interval for a range 0.00–0.5 s. The upper (black) trace shows the duration of interval for a range 0.00–0.035 s. Less than 2% of the interstimulus intervals are shorter than 3 ms. (*C*) Power spectral density of event times. The *y* axis shows power spectral density in events^2/Hz. The *x* axis shows the frequency in Hertz. The vertical magenta lines indicate primary and harmonics of the powerline frequency (60 Hz).

Table S1. Behavioral performance measures and number of clusters recorded from each patient

Patient	No. of sessions	Hit rate	False-alarm rate	% correct	d'	Multiunits	Single units
P1	3	0.55	0.17	0.69	1.23	29	13
P2	2	0.45	0.28	0.59	0.53	57	2
P3	1	0.69	0.44	0.63	0.65	11	6
P4	3	0.45	0.21	0.62	0.69	40	4
P5	3	0.58	0.06	0.76	1.79	20	5
P6	1	0.09	0.03	0.53	0.54	5	1
P7	3	0.54	0.08	0.73	1.58	9	2
P8	1	0.06	0.02	0.52	0.01	7	0
P9	1	0.19	0.03	0.58	0.98	8	1
Average	2.0	0.40	0.15	0.63	0.89	20.7	3.8

For patients who completed more than one session, the behavioral measures were computed separately for each session and then averaged across sessions.