Single-Neuron & LFP Representations of Fear Processing in the Human Brain

George Jacob (george.jacob@ruhr-uni-bochum.de)[1], Daniel Pacheco-Estefan[2], Antoine Bouyeure[1], Katia Lehongre[3], Valerio Frazzini[3], Vincent Navarro[3], & Nikolai Axmacher[1]

[1]Department of Neuropsychology, Institute for Cognitive Neuroscience, Ruhr University Bochum

[2]Department of Basic, Developmental and Educational Psychology, Faculty of Psychology, Autonomous University of Barcelona

[3]Institut du Cerveau (ICM), Sorbonne Université

Abstract

Fear extinction involves forming inhibitory safety memories that compete with the original fear traces. We recorded intracranial microelectrode data from 24 epilepsy patients to investigate how neuronal activity supports these representational shifts. During acquisition, amygdala neurons showed increased firing rates (475–525 ms) for aversive outcomes (n = 19). Oscillatory analyses (n = 24) revealed that the entorhinal cortex exhibited high-gamma enhancement (52–70 Hz and 74–88 Hz; both 0.80–1.10 s) following aversive outcome delivery during acquisition, while extinction was marked by low-frequency suppression (2–10 Hz, 1.15–1.50 s) following reinforcement. When combining acquisition and extinction phases, the hippocampus showed early high-gamma responses (76–98 Hz, 0.25–0.40 s) to aversive outcomes. Critically, hippocampal gamma power (38–62 Hz, 0.30–0.40 s) during extinction distinguished stimuli retaining threat value (CS++) from those undergoing extinction (CS+-), with the latter becoming indistinguishable from always-safe cues (CS--). Our results provide a cellular foundation for how the human mind adjudicates between competing memories and updates mental models through sensory evidence and threat evaluation.

Keywords: fear extinction; safety memory; amygdala; hippocampus; entorhinal cortex; single-unit recordings; local-field potentials; intracranial EEG; oscillations

Introduction

Dysfunctional fear and extinction learning are key features of posttraumatic stress disorder and anxiety-related disorders (Ressler, 2020). Fear extinction does not represent the erasure of an original memory, but rather the formation of a new inhibitory "safety memory" that competes with the fear trace. Neuroimaging studies have delineated a network involving the amygdala, hippocampus, and ventromedial prefrontal cortex (vmPFC) in the encoding of these responses (LaBar et al., 1998; Milad et al., 2007). To our knowledge, no previous human study has investigated the contributions of single neurons or local-field potentials (LFPs) recorded from microelectrodes toward fear acquisition and extinction. This represents a critical gap, as LFPs reflect the summed synaptic activity of local neuronal populations—vital for understanding the synchronization of local assemblies—while single-unit recordings provide the most direct measure of neural computation. Recent macrowire iEEG data using the same paradigm have provided deeper insights into the representational dynamics of fear extinction, revealing how safety and threat cues and contexts are differentiated across the network (Pacheco-Estefan et al., 2025). Here, we aim to ground these macrowire iEEG observations in cellular activity, utilizing microwire recordings to investigate how local neuronal populations and micro-LFPs support these representational shifts.

We collected intracranial microelectrode data from 24 patients with drug-resistant epilepsy while they performed a differential fear conditioning and extinction paradigm. We tested three primary hypotheses regarding cellular activity in the amygdala: 1) Neurons exhibit higher firing rates for the unconditioned stimulus (US) compared to no-US; 2) Firing rates will discriminate between the conditioned stimulus (CS+) and the safe stimulus (CS-) during acquisition; and 3) Firing rate differences will diminish during extinction. Additionally, we conducted preregistered exploratory analyses of microwire LFP data across the amygdala, hippocampus, and entorhinal cortex to investigate how oscillatory dynamics coordinate with these changes in single-neuron activity.

Methods

Participants

Twenty-four patients with medically intractable epilepsy were included as participants in this study. Data collection took place at the Pitié Salpêtrière Hospital, Paris. The study was conducted according to the latest version of the Declaration of Helsinki and approved by the Institutional Review Board at the Pitié Salpêtrière Hospital. All patients provided written informed consent before participating.

Experimental Design

Stimuli consisted of three electrical device images (a hair dryer, a ventilator, and a toaster) presented across acquisition, extinction, and test phases. The paradigm is called the “unlucky backpacker” paradigm (see Fig. 1), and the participant is asked to imagine that they are a backpacker who visits various landscapes on holiday. Each of these landscapes represents a specific context (e.g., beach,

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

mountain, etc.), and in each trial within the context, they encounter an electrical device. Some of these devices are faulty and are followed by an aversive stimulus (CS+), whereas others are neutral, and no aversive stimulus is administered (CS-). Unconditioned Stimulus (US): The US combined a negative facial expression with an unpleasant scream (1 s). This intensity was calibrated to be highly aversive but non-painful. The experiment was administered using Presentation software (Neurobehavioral Systems) on a Windows laptop. Behavioral differences were assessed via two-way repeated measures ANOVA (factors: trial type, phase) on average ratings, utilizing Bonferroni-corrected post hoc tests. To track progressive learning, we compared ordinal trial positions across conditions (CS++, CS+-, CS--) using non-parametric Wilcoxon signed-rank tests, applying cluster-based permutation statistics to correct for multiple comparisons.

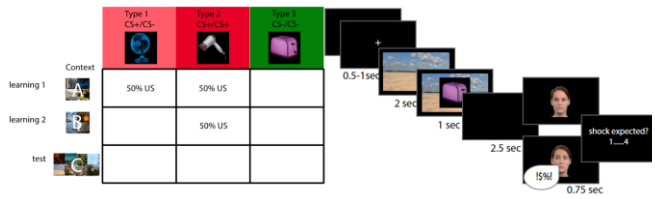

Figure 1: Schema of the "unlucky backpacker" paradigm.
Left: Stimulus types across phases. CS+ paired with aversive outcome (50% reinforcement); CS- never reinforced.
Right: Trial timeline showing context, CS, expectancy rating, and US delivery.

iEEG Recordings and ROIs

All microwire data were recorded via a Neuralynx amplifier at a sampling rate of 32,764 Hz. Each patient was implanted with 1-4 AdTech microwire bundles, with each bundle containing eight channels. Macroelectrode implantation sites were dictated solely by clinical needs. For LFP analysis, all data were downsampled to 1,000 Hz. Electrode localization was done via visual inspection, which is the standard method used to validate MNI normalized data when available. Bipolar referencing was not possible due to a lack of MNI coordinates of microwires, so we prioritized using the hardware reference (dedicated 9th reference microwire), and only re-referenced using the quietest channel in cases where the acquisition reference resulted in poor data quality. We used the acquisition reference for 28 out of 41 total microwires, and a manually selected quiet channel (from one of the eight analysis channels) for 13 microwires. To ensure consistency, in such cases, we always selected the quietest channel available. Region-of-interest (ROI) selection was conducted based on feasibility, since most microwires were distributed across the ROIs we selected, namely, the amygdala, hippocampus (subiculum + CA3-CA4), and entorhinal cortex. Moreover, these are key regions within the fear processing network, which further justified their selection for analysis.

Univariate Analysis

We report here the univariate results obtained from 19 patients, since combined analyses including the newest 5 patients are currently underway. All raw data were preprocessed and spike sorted using wave_clus, using the default parameters. Spike timestamps from the EEG system were temporally aligned with the trigger timestamps, also from the EEG system. Following this, the trial timestamps from the Presentation software (which was used to administer the experiment) were aligned with the trigger timestamps, ensuring that the alignment was correctly done before defining the trial epochs for all analyses. Firing rate analyses were conducted using cluster-based permutation analyses, across the different phases (e.g., acquisition, extinction) and stimuli contrasts (CS+ vs CS-; US vs no-US; CS++ vs. CS+- vs. CS --).

LFP Analysis

Following referencing, we obtained cleanly referenced LFP data, with common-source noise removed and bad channels excluded, ready for signal cleaning. We applied a standardized artefact rejection pipeline, based on previous work by Pacheco et al. (2025). High-amplitude, non-physiological artifacts were automatically detected using a literature-verified standard deviation threshold. Artifact periods were temporarily replaced with NaNs, and all trials with NaNs were removed before time-frequency decomposition. The data were band-pass filtered to keep frequencies between 1-120 Hz. A notch filter was used to remove 50 Hz electrical line noise and its harmonics. Finally, the data were downsampled to 1,000 Hz. The result was one data file per bundle containing the cleaned, preprocessed LFP data in a standard FieldTrip format, ready for direct input for time-frequency power analysis. We used variable wavelet cycles across the 1-100 Hz range (3-25 cycles, with a minimum of 3, and otherwise n/4 where n represents the frequency of interest) for improved time-frequency resolution at lower and higher frequencies.

Results

In the amygdala, univariate firing rate analysis of neuronal spiking during fear acquisition revealed a significant cluster of increased firing rates for US relative to no-US trials, localized to the 475–525 ms post-stimulus window. A total of 24 neurons were included in the amygdala ROI.

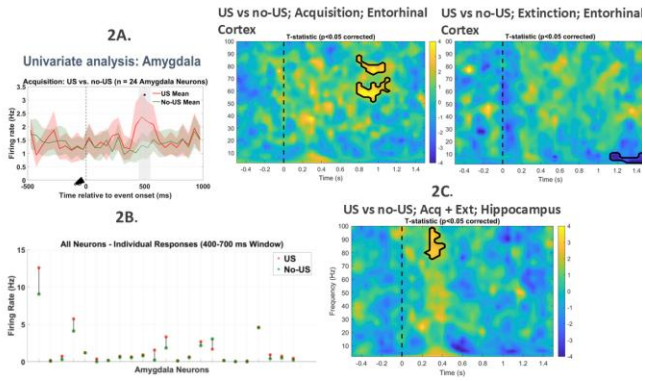

Figure 2: Amygdala and entorhinal-hippocampal responses to US delivery versus omission

2A: Group-level peri-stimulus time histogram of amygdala neuronal firing rate (N = 24 neurons) during acquisition. Shaded regions indicate SEM. Cluster-based permutation testing revealed significantly elevated firing rate for US compared to no-US trials between 475–525 ms post-event (Max cluster stat = 4.48, p = 0.03).

2B: Individual neuron firing rates during the 400–700 ms window. Each point represents a single neuron's mean firing rate for US (red) and no-US (green) trials.

*2C: Time-frequency representations of t-statistics comparing US vs. no-US trials. Black contours indicate significant clusters (p <0.05, cluster-corrected). Top left: Entorhinal cortex during acquisition showed high-gamma enhancement for US trials (*52–70 Hz and 74–88 Hz; both 0.80–1.10 s*). Top right: Entorhinal cortex during extinction showed low-frequency suppression following US delivery (~2–10 Hz, 1.15–1.50 s). Bottom: Hippocampus across acquisition and extinction combined revealed early high-gamma enhancement for US trials (~76–98 Hz, 0.25–0.40 s).*

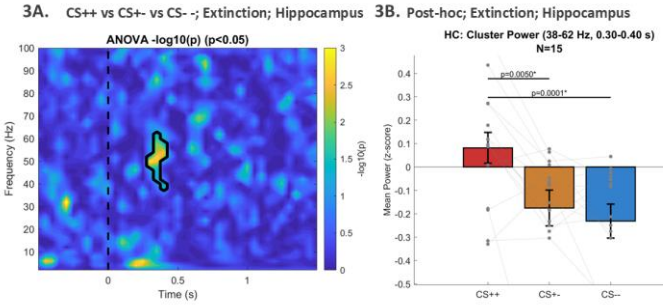

Figure 3: Hippocampal LFP differentiates CS types during extinction.

3A: Time-frequency representation of repeated-measures ANOVA comparing CS++, CS+-, and CS-- conditions during extinction in the hippocampus (N = 15 microwires). Color scale indicates -log10(p), with warmer colors reflecting stronger effects. Black contour outlines a significant cluster (38–62 Hz, 0.30–0.40 s; p <0.05, cluster-corrected).

3B: Post-hoc comparisons of mean z-scored power extracted from the significant cluster. The effect was driven by elevated gamma power for CS++ trials (maintained threat) relative to both CS+- (previously reinforced, now extinguished; p = 0.005) and CS-- (never reinforced; p = 0.0001), which were statistically indistinguishable (p = 0.37). Gray dots represent individual subjects; connecting lines illustrate within-subject patterns. Error bars indicate SEM.

Within the entorhinal cortex ROI, we observed differential spectral responses to aversive outcome delivery across acquisition and extinction. During acquisition, two significant high-frequency clusters were observed during US delivery compared to its omission (N = 10 microwires). The first spanned ~74–88 Hz at 0.80–1.10 s (p <0.001, cluster stat = 93.87), while the second covered ~52–70 Hz at 0.80–1.05 s (p = 0.041, cluster stat = 56.77). During extinction, the entorhinal cortex showed the opposite pattern: low-frequency power (~2–10 Hz) was suppressed following US delivery compared to trials without reinforcement, emerging at 1.15–1.50 s post-onset (p = 0.046).

Combining acquisition and extinction phases, the hippocampus exhibited enhanced high-gamma power (76–98 Hz) in response to US delivery compared to its omission (N = 15 microwires, p = 0.047, cluster stat = 61.04). This effect emerged early, within 0.25–0.40 s post-onset, suggesting rapid hippocampal discrimination of aversive outcomes.

To disentangle threat persistence from reinforcement updating, we compared LFP power across conditioned stimuli with varying predictability (CS++, CS+-, CS--) during extinction. A significant main effect of CS type was identified in the hippocampus (N = 15 microwires, p = 0.012, cluster stat = 107.63), characterized by a gamma-band cluster (38–62 Hz) between 0.30–0.40 s post-stimulus. Post-hoc comparisons revealed that CS++ trials—which maintained threat value—exhibited elevated gamma power (0.082 ± 0.066) relative to both CS+- (−0.176 ± 0.076; p = 0.005) and CS-- (−0.231 ± 0.073; p <0.001). Critically, CS+- and CS-- were statistically indistinguishable (p = 0.37), indicating that extinguished stimuli had shifted to resemble always-safe cues.

Discussion

We identified neural signatures of aversive outcome detection and threat evaluation across the human amygdala, hippocampus, and entorhinal cortex during fear conditioning and extinction.

During fear acquisition, amygdala neurons rapidly differentiated aversive outcomes from their absence, showing

a significant increase in firing around 475-525 ms post onset. Within the entorhinal cortex, we observed distinct spectral signatures of outcome processing across acquisition and extinction. During acquisition, US delivery elicited enhanced high-gamma activity (52–70 Hz and 74–88 Hz; both 0.80–1.10 s) compared to its omission, suggesting that the entorhinal cortex rapidly encodes aversive sensory information. During extinction, this pattern reversed: low-frequency power (2–10 Hz) was suppressed following US delivery compared to trials without reinforcement (1.15–1.50 s). Given that the no-US condition during extinction includes both extinguished (CS+-) and never-reinforced (CS--) stimuli, this suppression may reflect differential processing of reinforced versus non-reinforced outcomes rather than a prediction error signal per se. In the hippocampus, combining acquisition and extinction phases revealed early high-gamma enhancement (76–98 Hz, 0.25–0.40 s) following US delivery compared to its omission. This rapid response suggests that the hippocampus quickly discriminates aversive outcomes, potentially supporting the encoding of outcome-related information into episodic memory. Critically, the hippocampal ANOVA comparing CS types during extinction revealed that gamma power (38–62 Hz, 0.30–0.40 s) specifically tracked current threat status. Stimuli that remained threatening (CS++) elicited increased gamma power relative to both extinguished stimuli (CS+-) and always-safe cues (CS--), which were statistically indistinguishable from each other. This indicates that hippocampal gamma does not simply reflect stimulus history, but rather the updated threat relevance following extinction learning. Future studies should investigate directional information flow between entorhinal cortex and hippocampus during extinction, potentially using Granger causality or phase-amplitude coupling analyses. Additionally, spike-field coherence analyses may reveal whether single-unit activity in the amygdala coordinates with hippocampal gamma oscillations during threat evaluation.

In summary, our study characterizes the human brain's ability to navigate shifting environmental contingencies. During acquisition, amygdala neurons signal aversive outcomes through increased firing rates (475–525 ms), while the entorhinal cortex exhibits high-gamma enhancement to US delivery. As we transition to extinction, entorhinal low-frequency power differentiates aversive from non-aversive outcomes, and hippocampal gamma oscillations track updated threat relevance by distinguishing stimuli that retain danger from safe cues. This cellular and oscillatory evidence suggests that fear extinction is a dynamic process of adjudicating between competing memories, where single-neuron and LFP dynamics provide the neural substrate for updating mental models based on shifting sensory evidence and threat evaluation.

References

Ressler, K. J. (2020). Translating across circuits and genetics toward progress in fear- and anxiety-related disorders. *American Journal of Psychiatry*, *177*(3), 214–222. https://doi.org/10.1176/appi.ajp.2020.20010055

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron*, *20*(5), 937–945. https://doi.org/10.1016/s0896-6273(00)80475-4

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological Psychiatry*, *62*(5), 446–454. https://doi.org/10.1016/j.biopsych.2006.10.011

Pacheco-Estefan, D., Bouyeure, A., Jacob, G., et al. (2025). Representational dynamics during extinction of fear memories in the human brain. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-025-02268-5