COGNITIVE NEUROSCIENCE

Theta oscillations synchronize human medial prefrontal cortex and amygdala during fear learning

Si Chen[1,2], Zheng Tan[1,2], Wenran Xia[1,2], Carlos Alexandre Gomes[3], Xilei Zhang[1], Wenjing Zhou[4], Shuli Liang[5], Nikolai Axmacher[3,6], Liang Wang[1,2]*

Numerous animal studies have demonstrated that fear acquisition and expression rely on the coordinated activity of medial prefrontal cortex (mPFC) and amygdala and that theta oscillations support interregional communication within the fear network. However, it remains unclear whether these results can be generalized to fear learning in humans. We addressed this question using intracranial electroencephalography recordings in 13 patients with epilepsy during a fear conditioning paradigm. We observed increased power and interregional synchronization of amygdala and mPFC in theta (4 to 8 hertz) oscillations for conditioned stimulus (CS+) versus CS−. Analysis of information flow revealed that the dorsal mPFC (dmPFC) led amygdala activity in theta oscillations. Last, a computational model showed that trial-by-trial changes in amygdala theta oscillations predicted the model-based associability (i.e., learning rate). This study provides compelling evidence that theta oscillations within and between amygdala, ventral mPFC, and dmPFC constitute a general mechanism of fear learning across species.

INTRODUCTION

Neural oscillations have been suggested as a mechanism for the coordination of cell assemblies that synchronizes spiking activity and gates interregional information transfer (*1*). Numerous animal studies have demonstrated that fear learning and expression of fear rely on the coordinated activity of amygdala and medial prefrontal cortex (mPFC) (*2*–*5*) and that theta (4 to 8 Hz) oscillations support communication within this network (*2*–*4*, *6*). A recent nonhuman primate study showed that amygdala spikes are synchronized with mPFC activity and transfer error signals to support aversive learning (*3*). An involvement of mPFC and amygdala in fear expression has also been observed in human neuroimaging (*7*–*10*) and has been indirectly inferred from scalp electroencephalography (EEG) studies (*11*, *12*). Within the mPFC, two subregions have been proposed to play opposing functions in fear memory: The rodent prelimbic cortex, which is homologous to human dorsal mPFC (dmPFC) (*13*), participates in fear expression (*14*). By contrast, the rodent infralimbic cortex, homologous to human ventral mPFC (vmPFC), is responsible for fear extinction (*15*).

Because of the location of the fear network in deep brain regions, direct neurophysiological evidence on fear learning is scarce in humans. Intracranial EEG (iEEG) recordings in patients with epilepsy are pivotal for gaining direct access to these areas. However, while several recent iEEG studies investigated neural mechanisms underlying the role of emotions for declarative memory formation (*16*–*18*), the basic mechanisms of human fear learning have remained largely unexplored. Specifically, the neurophysiological signals that may govern communications between amygdala, vmPFC, and dmPFC during human fear learning are still unknown. In addition, some studies suggest that when organisms learn cue-reinforcer associations, they apply a rule known as associability that is dynamically determined by trial-wise prediction errors (*19*). Recent neuroimaging studies using associability-based computational models demonstrated that the human mPFC is implicated in fear extinction (*20*), but the role of amygdala is less computationally characterized in fear learning, especially for iEEG data.

To address these questions, we used a classic fear conditioning paradigm (Fig. 1A). On the basis of iEEG data simultaneously recorded from dmPFC, vmPFC, and amygdala (Fig. 1B and fig. S1), we found that human fear learning relies on information transfer via theta oscillations between the amygdala and two areas in the mPFC, providing strong evidence for the employment of theta synchronization during fear learning in the human brain.

RESULTS

Skin conductance responses during fear learning

Thirteen patients with epilepsy participated in the fear conditioning task (Fig. 1A), and in 11 of them, skin conduct responses (SCRs) could be recorded (table S1 and fig. S2). To avoid the confounding effect of electric stimulation on SCR, we excluded the trials paired with stimulation from SCR analysis. The grand average of the SCRs showed successful fear acquisition in 11 patients (Wilcoxon signed-rank test, $P = 0.04$; Fig. 1C). Among them, seven patients showed successful fear acquisition as indicated by a differential SCR to the CS+ relative to the CS− (difference > 0.05 μs) during the second half of acquisition, while four patients failed to show conditioned responses (CRs).

To explore the trial-by-trial learning dynamics of SCRs, we implemented SCR analyses via psychophysiological modeling (PsPM), which has been shown to be more sensitive and objective than conventional manual peak scoring methods of SCR data (*21*, *22*). We analyzed these data in the entire group of individuals (i.e., including both learners and nonlearners) using a linear mixed-effects (LME)

[1]CAS Key Laboratory of Mental Health, Institute of Psychology, Beijing, China. [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. [3]Department of Neuropsychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, Bochum, Germany. [4]Epilepsy Center, Tsinghua University Yuquan Hospital, Beijing, China. [5]Functional Neurosurgery Department, Beijing Children's Hospital, Capital Medical University, Beijing, China. [6]State Key Laboratory of Cognitive Neuroscience and Learning and IDG/McGovern Institute for Brain Research, Beijing Normal University, Xinjiekouwai Street 19, Beijing 100875, China.
*Corresponding author. Email: lwang@psych.ac.cn

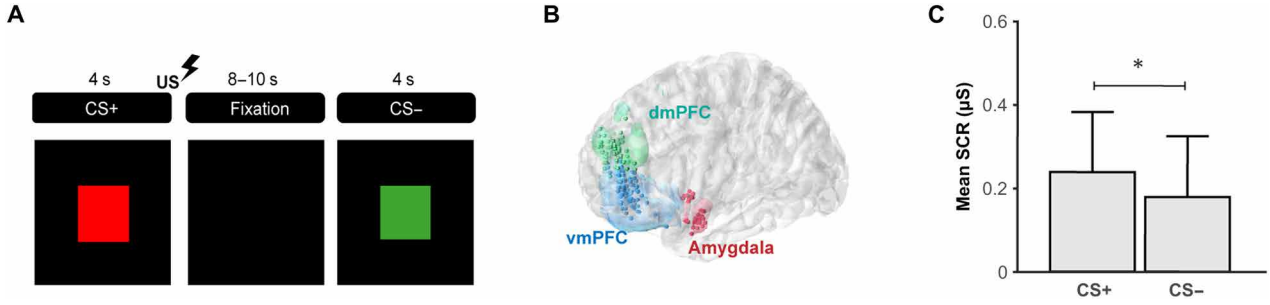

Fig. 1. Schematic depiction of the experimental paradigm and SCR results. (**A**) Trial structure and timeline for a single CS trial. The visual stimulus was presented for 4000 ms followed by an 8000- to 10,000-ms intertrial interval. A 15-ms electric stimulation (US) occurred at the end of 50% of all CS+ trials. (**B**) Depiction of electrode contact localizations in amygdala, vmPFC, and dmPFC for all 13 patients superimposed on a semitransparent MNI152 brain (as viewed from the left side). (**C**) Means (±SEM) CS-evoked SCRs during late acquisition over all patients (Wilcoxon signed-rank tests, $P = 0.04$). *$P < 0.05$.

model with trial number and condition (CS+ versus CS−) as fixed effects and "subject" as random effect (fig. S3A). This analysis revealed a significant interaction between condition and trial number [$F(1,2250.1) = 9.11$, $P = 0.002$], indicating significant learning effects on the condition differences. We also conducted a linear trend analysis and found a significant linear trend for CS+ items [$t(2248) = 2.69$, $P = 0.007$] but not for CS− items [$t(2248) = -1.41$, $P = 0.16$].

Next, we separately tested learning effects in the first and the second half of the data. Again, we applied LME model with trial number and condition (CS+ versus CS−) as fixed effects and subject as random effect. For both the first and the second half, we found a significant condition effect [first half: $t(2248) = 2.69$, $P = 0.007$; second half: $t(2250) = -6.74$, $P < 0.001$; fig. S3B].

Last, we directly compared learning effects in the first and the second half of the data. We applied LME model with trial number, condition (CS+ versus CS−), and "learning phase" (first versus second half) as fixed effects and subject as random effect. This analysis showed that the condition difference was significantly more pronounced in the second than the first half of the data [$t(2250) = 2.45$, $P = 0.014$], which suggests that learning effects become more stable in the second half. Therefore, we focused on the second half of trials in the following analysis.

Theta oscillations in amygdala, vmPFC, and dmPFC during fear learning

We then examined iEEG responses in amygdala, dmPFC, and vmPFC. We built an LME model with condition (CS+ versus CS−) as fixed effect and individual and electrode as random effects. When LME model was applied to the time-frequency contrast map (CS+ versus CS−) across the second half of trials, clusters with significantly increased power in the CS+ condition were found in all regions: In the amygdala, a significant cluster was found between 700 and 1700 ms and between 3 and 12 Hz ($P = 0.042$; Fig. 2A, top row); in the vmPFC, between 600 and 2000 ms and between 2 and 22 Hz ($P = 0.033$; Fig. 2A, middle row); and in the dmPFC, between 300 and 2000 ms and between 2 and 7 Hz ($P = 0.026$; Fig. 2A, bottom row). These clusters were dominated by theta activity, as indicated by the percentage area of the respective clusters that fell in the theta frequency range: Using a definition of theta as 4 to 8 Hz, we found 48, 42, and 74% for amygdala, vmPFC, and dmPFC, respectively; when a broader definition of theta was applied that is more in line with previous results from animals (2 to 12 Hz) (*23*), the area percentages increased to 91, 82, and 100% for these three regions. When theta was quantitatively averaged across power between 4 and 8 Hz and between 0 and 2000 ms after cue onset, this analysis revealed increased theta power in amygdala [LME, $t(44) = 6.92$, $P = 1^{-8}$; Fig. 2B, left], vmPFC [LME, $t(96) = 5.1$, $P = 1^{-6}$; Fig. 2B, middle], and dmPFC [LME, $t(62) = 7.1$, $P = 1^{-9}$; Fig. 2B, right] for CS+ versus CS− across the second half of trials in which learning effects were more pronounced. In contrast, no significant clusters were found when the first half of trials was considered (fig. S5), possibly because condition differences only started to emerge during this period. When we calculated learning effects on theta power across all trials, including trials from the early conditioning phase, we again observed an increased power of theta oscillations in amygdala, dmPFC, and vmPFC [amygdala: $t(44) = -2.83$, $P = 0.007$; vmPFC: $t(62) = -4.7$, $P = 1^{-5}$; dmPFC: $t(96) = -2.7$, $P = 0.008$; fig. S4]. While these effects were obtained using a white matter referencing scheme, we obtained the same effects when using a bipolar montage (fig. S6).

To analyze whether oscillatory activity in the theta band can actually be used to predict the patients' performance in fear learning, we also performed a receiver operating characteristic (ROC) analysis. The area under the curve (AUC) was computed using a leave-one-subject-out cross-validation scheme (chance = 0.5). We found that theta power in mPFC regions, but only by trend in the amygdala, distinguished learners from nonlearners (amygdala: AUC = 0.65, $P = 0.09$; dmPFC: AUC = 0.87, $P < 0.001$; vmPFC: AUC = 0.82, $P < 0.001$; fig. S7A).

Changes in theta power may be either due to narrow-band oscillations or broad-band shifts in the 1/f spectrum, which putatively reflect different neural mechanisms and cognitive functions (*24*). To characterize theta responses during fear learning, we applied a previously established algorithm that disentangles theta oscillations from the background 1/f spectrum ["better oscillation detection method" (BOSC); (*25*)] (Fig. 3A). The amygdala showed a peak at 9 Hz in 17.7 and 16.34% of the detection time for CS+ and CS−, respectively; dmPFC showed a peak at 5.5 Hz in 18.9% (CS+) and 17.7% (CS−) of the detection time; and vmPFC showed a peak at 6.5 Hz in 11.77% (CS+) and 11.61% (CS−) of the detection time (Fig. 3B). An ROC analysis revealed a higher frequency of those theta oscillations that showed the highest percentage of oscillation times in amygdala as compared to either dmPFC (AUC = 0.91, $P < 0.001$, permutation test; fig. S7B) or vmPFC (AUC = 0.8893, $P < 0.001$), with no difference between dmPFC and vmPFC (AUC = 0.43, $P = 0.95$).

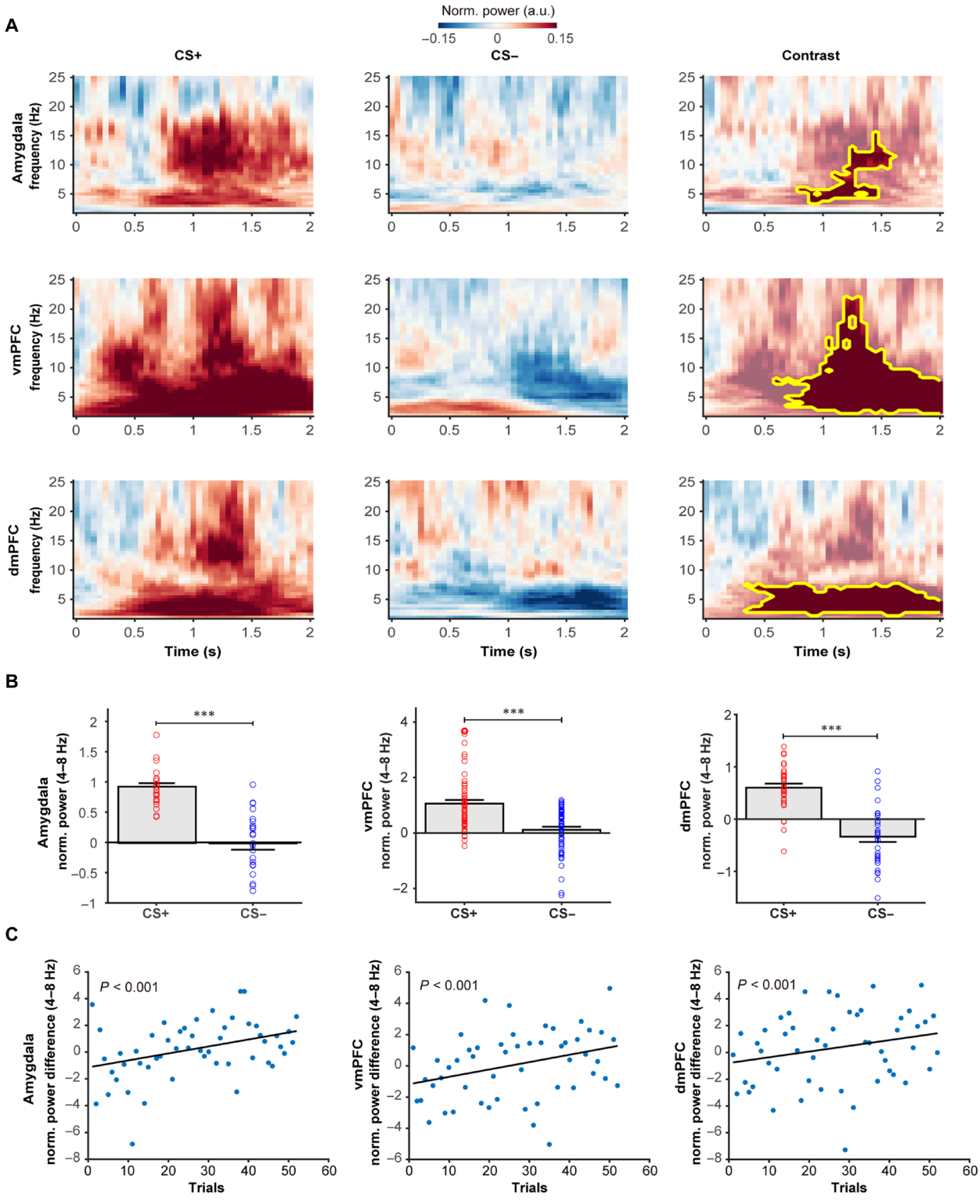

Fig. 2. Theta power in amygdala, dmPFC, and vmPFC during fear learning. (**A**) Group time-frequency spectrograms for CS+ condition (left), for CS− condition (middle), and contrast (CS+ versus CS−) with clusters after cluster-based permutation test for multiple comparison in amygdala (top), vmPFC (middle), and dmPFC (bottom) following CS onset. For all three regions, only the largest cluster passed the permutation test ($P < 0.05$). (**B**) Averaged (±SEM) theta power (4 to 8 Hz) across electrode contacts during the 2000 ms following CS onset for CS+ and CS− conditions for the second half of trials in amygdala [left; LME, CS+ versus CS−: $t(44) = 6.92$, $P = 1^{-8}$], vmPFC [middle; LME, CS+ versus CS−: $t(96) = 5.1$, $P = 1^{-6}$], and dmPFC [right; LME, CS+ versus CS−: $t(62) = 7.1$, $P = 1^{-9}$]. Red and blue open circles indicate electrode contact–wise data. (**C**) Averaged trial-by-trial theta power difference across electrode contacts in amygdala [left; LME, $t(903) = 5.16$, $P = 2^{-7}$], vmPFC [middle; LME, $t(1892) = 5.77$, $P = 9^{-9}$], and dmPFC [right; LME, $t(1407) = 4.48$, $P = 8^{-6}$]. Black line is the best fitting line using the robust fit. ***$P < 0.001$. a.u., arbitrary units.

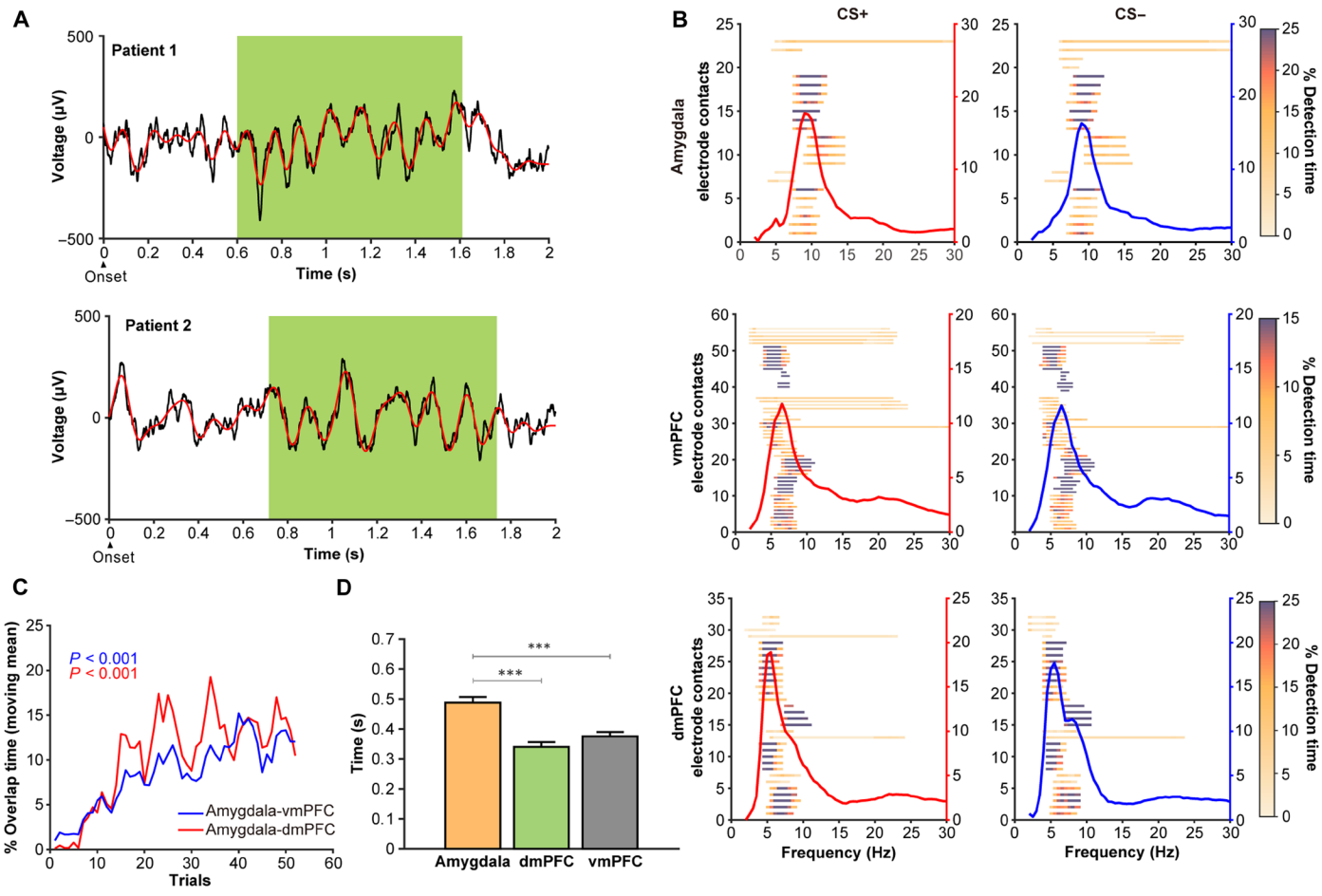

Fig. 3. Time characteristics of theta oscillations detected by BOSC. (**A**) Two example trials with theta oscillations after CS onset from different patients. Black, raw signal; red, low-frequency component of the raw signal (<10 Hz); green shading, time periods with theta oscillations detected by BOSC. (**B**) Percentage of time during which BOSC detected a given instantaneous frequency on all electrode contacts in amygdala (top), vmPFC (middle), and dmPFC (bottom). Red (CS+) or blue (CS–) line shows a summary plot across electrode contacts. (**C**) Learning-dependent increase in the overlap between time periods showing simultaneous theta oscillations in amygdala and dmPFC [LME, $t(2494) = 9.49$, $P < 0.001$] and amygdala-vmPFC [LME, $t(5978) = 15.06$, $P < 0.001$]. (**D**) Averaged (±SEM) onset (across electrode contacts) of theta oscillations detected by BOSC in three brain regions [LME, amygdala versus dmPFC: $t(53) = 5.06$, $P = 5^{-6}$; amygdala versus vmPFC: $t(70) = 4.69$, $P = 1^{-5}$; dmPFC versus vmPFC: $t(86) = -1.32$, $P = 0.19$]. ***$P < 0.001$.

Temporal dynamics of theta oscillations across learning

Next, we explored the dynamic profiles of theta oscillations in the three brain regions. All regions showed a significant increase in theta power throughout the learning phase [LME, amygdala: $t(903) = 5.16$, $P = 2^{-7}$; dmPFC: $t(1892) = 5.77$, $P = 9^{-9}$; vmPFC: $t(1407) = 4.48$, $P = 8^{-6}$; Fig. 2C]. As expected, patients who failed to discriminate CS+ from CS– did not show the increase in theta power for CS+ versus CS– in any of the three regions [LME and amygdala: $t(28) = -1.64$, $P = 0.11$; dmPFC: $t(42) = -1.56$, $P = 0.13$; vmPFC: $t(42) = -4.1$, $P = 1^{-4}$; fig. S8A], nor did they show a gradual increase in the trial-by-trial analysis [LME, amygdala: $t(658) = 1.43$, $P = 0.15$; vmPFC: $t(1016) = -2.2$, $P = 0.02$; dmPFC: $t(1057) = 1.66$, $P = 0.1$; fig. S8B].

To investigate whether the theta onset covaries with the progression of learning, we divided the entire acquisition phase into six blocks, using a sliding window of 12 trials with 4 trials overlap. We then extracted the onsets of theta oscillations detected by BOSC and correlated them with the block number using LME model, with individuals and electrodes as random factors and "onset latencies" in the amygdala, dmPFC, and vmPFC as dependent variables, respectively. We found that the latency of theta onsets covaried with the progression of learning in amygdala and vmPFC, specifically for CS+ trials but not for CS– trials [amygdala: CS+, $t(107) = -4.67$, $P = 8^{-6}$; CS–, $t(108) = -1.44$, $P = 0.15$; vmPFC: CS+, $t(286) = -4.82$, $P = 2^{-6}$; CS–, $t(289) = -1.35$, $P = 0.18$; dmPFC: CS+, $t(186) = -1.39$, $P = 0.17$; CS–, $t(187) = 1.79$, $P = 0.08$; fig. S9A]. As learning proceeded, theta oscillations occurred at earlier latencies. Specifically, amygdala theta oscillations had a latency of 670 ± 60 ms in the first block and 360 ± 50 ms in the sixth block; vmPFC theta oscillations initially had a latency of 470 ± 40 ms in the first block and 280 ± 30 ms in the sixth block.

To investigate whether theta frequencies also covary with the progression of learning, we extracted the peak frequency in the 2- to 12-Hz band of each electrode and correlated it with the block number using LME model, with individuals and electrodes as random factors and "theta peak" in the amygdala, dmPFC, and vmPFC, respectively,

as dependent variable. Although absolute differences were relatively small, we found significant linear changes of peak frequencies across CS+ trials in the dmPFC and vmPFC but not in the amygdala [amygdala: CS+, $t(112) = -1.59$, $P = 0.11$; CS−, $t(112) = -1.1$, $P = 0.27$; vmPFC: CS+, $t(292) = -2.16$, $P = 0.03$; CS−, $t(292) = -0.09$, $P = 0.93$; dmPFC: CS+, $t(190) = -2.84$, $P = 0.005$; CS−, $t(190) = 1.55$, $P = 0.12$; fig. S9B]. Specifically, dmPFC theta oscillations had a peak frequency at 6.93 ± 0.2 Hz in the first block and 5.97 ± 0.16 Hz in the sixth block; vmPFC theta oscillations initially had a peak frequency at 6.02 ± 0.17 Hz in the first block and 5.57 ± 0.1 Hz in the sixth block. No differences were found for CS− trials (all $P > 0.05$).

We also tested whether time periods showing simultaneous narrow-band theta oscillations between all pairs of amygdala-dmPFC and amygdala-vmPFC contacts changed with learning. We found gradually increasing overlapping oscillatory time periods across trials for both amygdala-dmPFC contacts [LME, $t(2494) = 9.49$, $P < 0.001$; Fig. 3C] and amygdala-vmPFC contacts [LME, $t(5978) = 15.06$, $P < 0.001$; Fig. 3C].

Together, these results show clear evidence that human fear learning is associated with increased theta power in amygdala, vmPFC, and dmPFC. Theta effects exhibit higher frequencies in the amygdala than in the mPFC regions. Moreover, we found an increasing overlap between regional theta oscillations when learning proceeded, providing a neurophysiological basis for interregional communication.

Interregional theta synchronization and information transfer

We extracted the onset of detected theta oscillations and averaged over all contacts in each region. We found a shorter latency in the mPFC (dmPFC, 339.7 ms; vmPFC, 368.2 ms) and a longer latency in the amygdala (488.3 ms) [LME, amygdala versus dmPFC: $t(53) = 5.06$, $P = 5^{-6}$; amygdala versus vmPFC: $t(70) = 4.69$, $P = 1^{-5}$; dmPFC versus vmPFC: $t(86) = -1.32$, $P = 0.19$; Fig. 3D]. These results indicate that theta oscillations occur earlier in mPFC regions than in the amygdala, suggesting that amygdala responses are driven by oscillations in the mPFC. To directly test this, we analyzed functional interactions between amygdala and the two mPFC regions. We measured imaginary coherence for CS+ versus CS− trials during the second half of trials and found significantly increased theta coherence between amygdala and vmPFC [$t(228) = 5.03$, $P = 9^{-7}$; Fig. 4A, top] and between amygdala and dmPFC [LME, $t(94) = 5.58$, $P = 2^{-7}$; Fig. 4A, bottom]. In addition, when we assessed the learning effects on theta coherence across all trials, we found significantly higher theta coherence between amygdala and dmPFC for CS+ versus CS− stimuli [$t(94) = -2.73$, $P = 0.008$; fig. S4, right]. However, theta coherence between amygdala and vmPFC was not significantly higher for CS+ versus CS− trials when all trials were taken into account [$t(228) = 1.32$, $P = 0.19$; fig. S4, left]. The theta-specific interregional synchronization was visible in the imaginary coherence spectrum between amygdala and vmPFC ($P = 0.027$ for the cluster in 4.6 to 5.0 Hz; $P = 0.014$ for the cluster in 6.1 to 7.0 Hz; Fig. 4B, top) and between amygdala and dmPFC (cluster-based permutation test, $P = 0.005$ for the cluster in 4.6 to 6.5 Hz; Fig. 4B, bottom). Furthermore, theta coherence between amygdala and two mPFC regions increased significantly across the entire learning phase [LME, amygdala and dmPFC: $t(1896) = 1.96$, $P = 0.05$; amygdala and vmPFC: $t(4606) = 7.26$, $P = 4^{-13}$; Fig. 4C].

To further examine the directionality of information transfer between the amygdala and the two mPFC regions, we calculated the phase slope index (PSI). Comparing the CS+ with the CS− condition, we observed significantly more negative phase slopes from the amygdala to the dmPFC, indicating fear-related transfer of theta activity from dmPFC to amygdala (significant period, 20 to 1340 ms after CS+ onset; LME model with permutation test, all $P < 0.05$; Fig. 4D). The PSI was negatively correlated with imaginary coherence between amygdala and dmPFC [LME, $t(46) = -2.43$, $P = 0.02$; Fig. 4E], indicating that the PSI from dmPFC to amygdala was positively correlated with dmPFC-amygdala imaginary coherence. For the amygdala-vmPFC circuit, the PSI did not show a significant information transfer in any direction. These results indicate that synchronized theta oscillations constitute an information channel for interregional communication between amygdala and both mPFC regions during fear learning and that directed transfer of information from dmPFC to amygdala facilitates the processing of fear information.

Model-based analysis of associability-related activity in the amygdala

Computational learning models supposed that instructive signals are activated only when there is a discrepancy between expectations based on sensory cues (here, CS) and outcomes [here, the aversive unconditioned stimulus (US)] (*26*, *27*). According to the Pearce-Hall model with time-varying learning rates (*26*, *28*, *29*), surprise generated by the unexpected occurrence of the US controls the acquisition rate by modulating the internal processing of the CS, an effect referred to as "associability." On the basis of this, we fit an associability model using the SCR data of individual patients that is consistent with previous associability model fitting approach (see below for details) (*20*). The core parameter of the model is the associability that is dynamically determined by the (unsigned, i.e., absolute) prediction error of the previous trial. Using this computational model, we can identify the neural processes that govern learning in individual trials.

We first validated the associability model behaviorally using SCRs. We conducted an LME model with SCR as dependent variable, associability as fixed effect and subject and (nested) trial as random effects for assessing correlations between associability and SCR. These analyses revealed a significant correlation between associability and SCR [LME, $t(388) = 4.69$, $P = 3^{-6}$; Fig. 5B]. Second, to quantitatively identify the neural correlates of associability, we conducted an LME model with theta power in the three regions as dependent variable, associability as fixed effect and subject and (nested) trial as random effects for assessing correlations between associability and theta power. These analyses revealed a significant correlation between associability and amygdala theta power [LME, $t(388) = 2.12$, $P = 0.035$; Fig. 5C]. By contrast, applying the same analysis to the data from vmPFC or dmPFC did not reveal any correlations between associability and theta power [LME, vmPFC: $t(388) = 1.31$, $P = 0.191$; dmPFC: $t(388) = 0.99$, $P = 0.323$].

DISCUSSION

The formation of fear memory involves interactions of the amygdala with other areas of the fear network, particularly within the mPFC. This study demonstrates, in humans, that theta oscillations within an amygdala-mPFC circuit play an important role during fear learning. We found increased theta oscillations in the amygdala, vmPFC, and dmPFC of patients who successfully learned the association but did not find these effects in patients who failed to acquire the fear CR. We then showed that oscillatory activity was largely governed

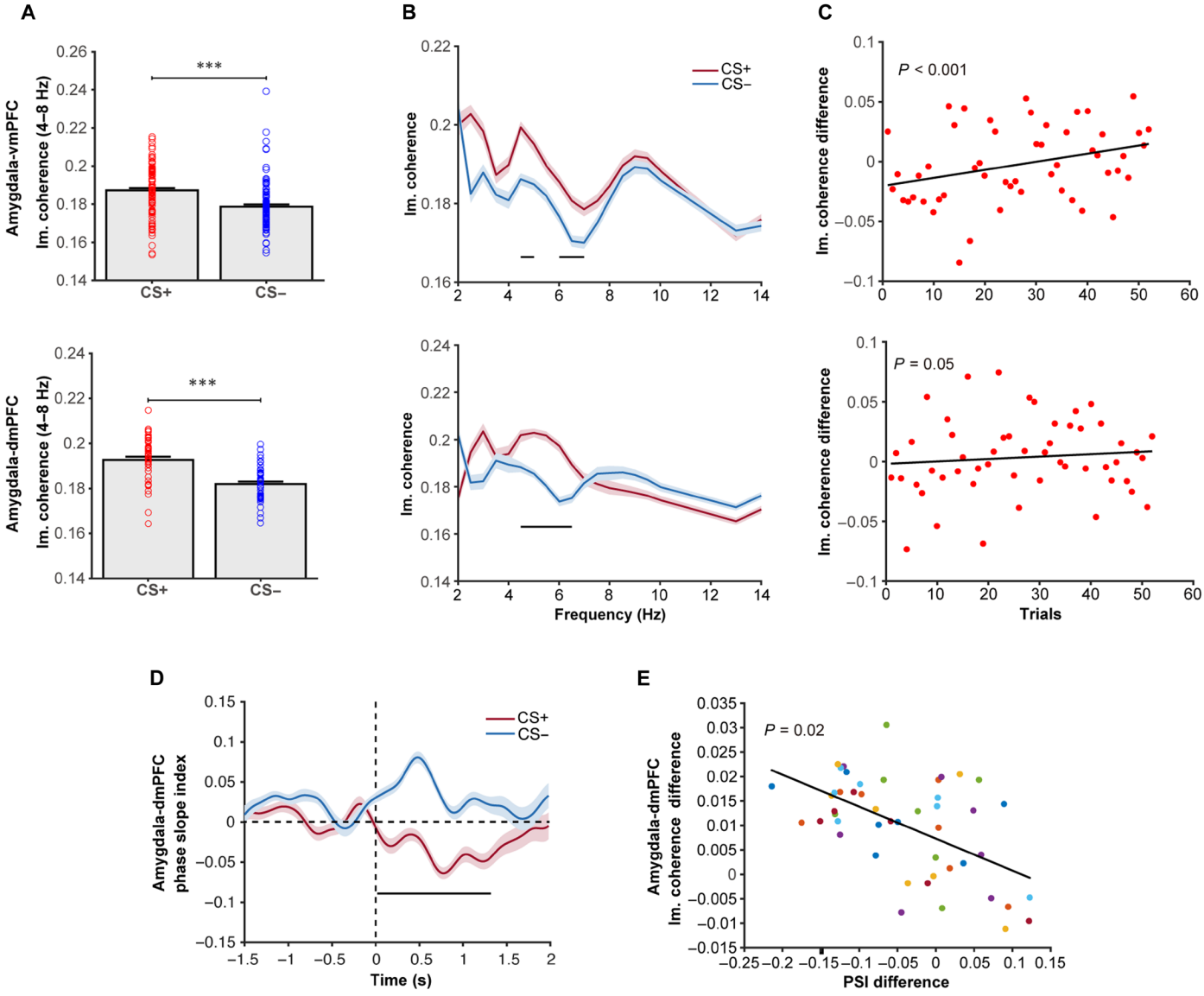

Fig. 4. Theta coherence between mPFC and amygdala during fear learning. (**A**) Averaged (±SEM) imaginary coherence across electrode contacts in the theta band for CS+ and CS– between amygdala and dmPFC [bottom; LME, CS+ versus CS–: $t(94) = 5.58$, $P = 2^{-7}$] and between amygdala and vmPFC [top; LME, CS+ versus CS–: $t(228) = 5.03$, $P = 9^{-7}$]. Red and blue open circles represent electrode contact–wise data. (**B**) Imaginary coherence between amygdala and dmPFC (top) and between amygdala and vmPFC (bottom) calculated across frequencies. Horizontal black lines below the curves denote significant differences between CS+ and CS– (LME model with permutation test, all $P < 0.05$). (**C**) Averaged trial-by-trial imaginary coherence across electrode contacts in the theta band for CS+ minus CS– between amygdala and dmPFC [top; LME, $t(1896) = 1.96$, $P = 0.05$] and between amygdala and vmPFC [bottom; $t(4606) = 7.26$, $P = 4^{-13}$]. Black line is the best fitting line using robust fit. (**D**) PSI calculated point by point across time using the theta signal from the amygdala as modulating channel and dmPFC as modulated channel. Black lines below the graph denote significant differences between a period of 20 to 1340 ms after CS onset (LME model with permutation test, all $P < 0.05$), showing that theta activity from dmPFC precedes amygdala across most of the stimulus presentation period. (**E**) Relationship between averaged imaginary coherence and PSI effects in the theta band for CS+ minus CS– between amygdala and dmPFC [LME, $t(46) = -2.43$, $P = 0.02$]. Black line is the best fitting line using robust fit. For (A) to (E), results were based on five patients with electrodes implanted in both amygdala and vmPFC/dmPFC. ***$P < 0.001$.

by theta oscillations and that the peak frequency was higher in the amygdala than in mPFC regions. Theta oscillations were synchronized between the amygdala and both mPFC areas, although these long-range neural synchronizations showed dissociable dynamic patterns: Analysis of information transfer revealed that amygdala activity was driven by the dmPFC but not by vmPFC. Last, computational modeling showed that activity in the amygdala, but not in mPFC regions, reflected associability as a measure of learning.

These findings provide direct evidence that human fear learning involves the amygdala-mPFC circuit and that theta oscillations serve as a neural mechanism for interregional communication and directional information transfer between amygdala and dmPFC. These findings expand our understanding of the amygdala-prefrontal circuit (*2*–*4*, *6*, *10*, *11*, *18*, *30*, *31*). These effects may be instrumental for the development of therapeutic strategies for anxiety disorders such as post-traumatic stress disorder.

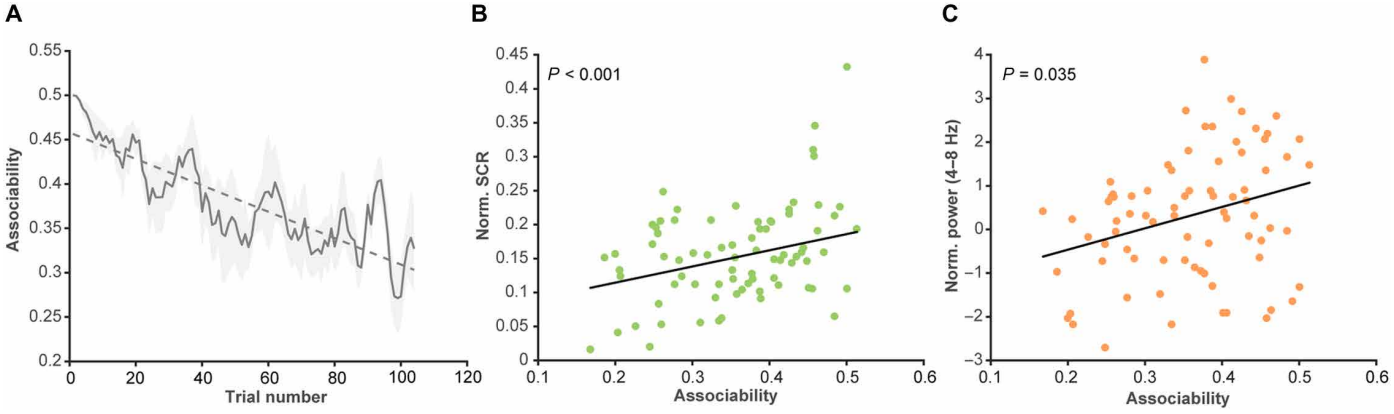

Fig. 5. Relationship between associability and amygdala theta power. (**A**) Averaged associability from computational modeling across individuals. (**B**) Relationship between associability of the computational model and normalized SCR data [LME, $t(388) = 4.69$, $P = 3^{-6}$]. Dots indicate estimates for individual trials across patients, calculated from the LME model. The black line is the best fitting line using the robust linear regression. (**C**) Relationship between associability of computational modeling and theta power in amygdala [LME, $t(388) = 2.12$, $P = 0.035$].

Increased theta power in the human amygdala and mPFC

Theta frequency oscillations have long been implicated in human learning and memory. Human studies of memory have provided a wide range of evidence for theta's role in successful learning, memory formation, and retrieval [for a recent review, see (*32*)]. Moreover, abundant work revealed theta oscillation–mediated communication between distant brain regions as a neural mechanism for information transfer in the human brain (*33*). In addition, various animal studies have demonstrated that theta oscillations support the interaction between amygdala and mPFC during the acquisition (*3*), expression (*30*), discrimination (*6*), and consolidation (*31*, *34*) of fear memory. Moreover, artificial induction of 4- and 8-Hz oscillations in the basolateral amygdala exerts a bidirectional control over conditioned freezing behavior in an experience- and context-specific manner (*35*), demonstrating the causal relevance of theta oscillations in rodents. iEEG studies in patients with epilepsy demonstrated that processing emotionally salient events recruited an amygdala-related network and that successful pattern separation of emotional stimuli was associated with theta oscillations (*16*, *18*). Consistent with these studies, we found that theta oscillations in the human amygdala and mPFC were significantly enhanced when a CS+ was presented as compared to a CS− in patients who successfully acquired fear. Moreover, using the modified BOSC, we found that theta oscillations indeed dominated the largest proportion of oscillation time in the entire time period, and the predominant peak frequency with the highest percentage of theta oscillation time was significantly different between the amygdala and both mPFC regions. Overall, the cross-species enhancement of theta oscillations suggests a common neural mechanism that effectively promotes learning of fear memory.

Amygdala-dmPFC interactions during fear learning

Extensive research confirms that aversive conditioning requires coordinated activity among different brain regions. In line with this, animal studies have demonstrated that codischarge of neurons in the dmPFC and amygdala is enhanced during resistance to extinction behavior (*36*). During fear discrimination, increased theta synchronization between mPFC and amygdala was found only in rodents that successfully differentiated between fear and safety cues (*6*). Primate research found increased power and phase synchrony in the theta range in amygdala and dorsal anterior cingulate cortex during aversive conditioning and that the synchrony was linked to single-unit spiking (*3*). Consistent with animal studies, we found enhanced coordination of theta activity between the dmPFC and amygdala during CS+ presentation.

Moreover, we found a dmPFC-to-amygdala directionality in theta oscillatory activity in response to the CS+, which is consistent with previous findings at the level of local field potentials (*2*) and firing rates in rodents (*6*). Although the precise role of dmPFC projections to amygdala is still unknown in humans, it has been proposed that activity in dmPFC may support fear conditioning both by acting as a bridging signal between representations of the CS and US and by providing information regarding the predictive value of the CS (*37*). Attention regulation may also underlie an indispensable function of the dmPFC during aversive learning (*38*, *39*). Our findings indicate that the dmPFC may use the anatomical projection to the amygdala to transfer emotional information and promote the formation of fear memory.

Anatomically, electron microscopic examinations revealed that a great majority of mPFC axon terminals terminate on dendritic spines of amygdala projection neurons and only a few on putative interneurons (*40*). This excitatory-to-excitatory projection results in fear enhancement (*41*). Using extracellular stimulation of amygdala in mice revealed that putative dmPFC principal neurons exhibit antidromic responses to amygdala stimulations, suggesting that the dmPFC preferentially targets the amygdala to drive fear responses (*42*). Furthermore, the dmPFC-to-amygdala directionality suggests that fear conditioning requires cortical inputs to the amygdala, rather than only relying on a faster subcortical pathway. Together, the increased coherence can serve as a means for facilitating interregional information communication between these core areas within the fear network.

Amygdala-vmPFC interactions during fear learning

The amygdala showed theta-based interactions not only with the dmPFC but also with the vmPFC during fear learning. This finding seems to conflict with previous results that the vmPFC is mainly involved in fear extinction (*37*). Previous animal studies explored the functional and neural mechanisms of vmPFC through cerebral

lesions, electric stimulation, and tracking techniques (*14*, *15*, *43*). However, most of the animal experimental manipulations were carried out after fear conditioning, and thus, the role of vmPFC in fear learning itself remains to be elucidated, in particular regarding neural activity. In humans, several neuroimaging studies suggest that the vmPFC is particularly recruited during fear extinction (*8*, *11*, *12*). Anatomically, a robust connection has been described between vmPFC and amygdala. An imaging tractography analysis of infralimbic and prelimbic connectivity in mice found both dmPFC and vmPFC connectivity with the amygdala (*44*). A targeted optogenetic stimulation study found that infralimbic inputs innervated principal neurons in the amygdala but not medial intercalated neurons (*45*), indicating that there is an anatomical connection between vmPFC and amygdala. However, it has remained unclear whether this connection plays an important role in fear learning.

The partial reinforcement rate used in the current study was lower compared with most animal studies, which may introduce uncertainty in the processing of fear-related stimuli. The vmPFC has been proven to be important in human decision-making, particularly in judgements under conditions of uncertainty (*46*). Patients with bilateral lesions of the vmPFC have difficulties in choosing between options with uncertain outcomes, whether the uncertainty is in the form of a risk or of an ambiguity (*47*–*49*). For fear conditioning with partial reinforcement rates, the probability of an aversive stimulus appearing is clearly uncertain. Functional magnetic resonance imaging (fMRI) data showed that the level of ambiguity in judgements correlates positively with activation in the amygdala and vmPFC, suggesting a neural system for evaluating general uncertainty (*50*). Another fMRI study revealed a clear learning curve that indexed the formation of associative fear in the vmPFC (*51*). These findings describe the process by which organisms explore their surroundings: In the case of uncertainty, the available information decreases, and the risk of decision-making based on available information increases. In this situation, the brain must mobilize more cognitive resources to find additional information from the environment. This may explain the different findings in the vmPFC across studies. Instead of being limited to fear extinction, the vmPFC may be widely involved in other aspects of emotional processing and regulation (*52*).

Associability in computational modeling predicts amygdala theta activity

Theories of associative learning suppose that, when animals learn cue-reinforcer associations, they track a quantity known as associability that reflects the extent to which each cue has previously been accompanied by surprise (*26*). This surprise can be generated either by the unexpected appearance of the US (fear learning) or the unexpected omission of the US (reversal learning and extinction learning). Previous studies (*28*, *29*) described the progression of both associative learning and reversal learning using the associability model: They found that the associability of each CS increased during early acquisition (unexpected appearance of the US); declined during late acquisition, when outcomes became more expected; and then increased rapidly during the reversal phase, when the outcomes were again unexpected. Another study (*20*) used the associability model to show that replacing threat with tone outcomes improved extinction retention, demonstrating the general role played by surprise in associative extinction (unexpected omission of the US). Hence, the associability-based model is suitable for both initial fear learning and extinction learning.

Using this computational modeling approach, we found that the associability was positively correlated with SCR and amygdala theta activity. This suggests that the amygdala has a specific functional role in controlling associability during learning, which is consistent with previous studies (*28*, *53*). Our findings extend the computational characterization of learning signals in the human amygdala, suggesting that it is complementary to representations of prediction error in the striatum. The same effect was not found in the mPFC regions, although amygdala activity was significantly driven by the dmPFC. In the future, it will be important to investigate how the dmPFC influences the amygdala-dependent associability signal.

There are several limitations in the current study. First, since the position of the electrode is determined clinically, we can only roughly explore the activity of the entire amygdala; the function of different subregions of the amygdala needs further exploration. Second, fear conditioning studies vary in the trial number and order, type of CS/US stimuli, reinforcement rate, and CS-US delay, which seems to affect the pattern of amygdala and mPFC activity. Despite these limitations, this study demonstrated that theta oscillations built up long-range communication and directional information transfer within amygdala-mPFC circuits during fear learning. The present work thereby supports that theta oscillations within and between amygdala, vmPFC, and dmPFC constitute a general mechanism of fear learning across species.

MATERIALS AND METHODS

Experimental design

A fear conditioning paradigm was adapted for simultaneous SCR-iEEG recordings (Fig. 1A). During the task, patients were instructed to view the stimuli displayed on the screen. Two different color squares (red and green) constituted the CS, which were randomly assigned to CS+ or CS− across patients. CS+ was reinforced with a mild electric stimulation (US; contingency of 50%), while CS− stimuli were never paired with a US. CS+ and CS− were presented 52 times each and were presented on a desktop screen (visual angle, 16°) in random order (patient 7 completed only 26 CS+ and 26 CS− trials due to time constraints). Each stimulus lasted for 4 s with a jittered intertrial interval of 8 to 10 s (CS offset to next CS onset). The US delivery occurred upon the termination of CS+ (i.e., delay conditioning). During the intertrial interval, a white fixation cross was shown on a black background.

The US consisted of a 15-ms single-pulse electric stimulation that was delivered from a transcutaneous current stimulator (STM200) using two steel disk electrodes attached to the right wrist (radius, 5.2 cm). During a workup procedure, the intensity of the electric stimulation was set individually to an extent that was subjectively perceived as "uncomfortable but not painful." In addition, participants had to rate the negative valence of the US higher than 7 on a 10-point Likert scale (0, not unpleasant at all; 10, extremely unpleasant).

Subjects

A total of 13 patients with medically refractory epilepsy (six females, average age: mean = 26 years, SD = 6.54; table S1) were recruited from Yuquan Hospital, Tsinghua University, and the First Affiliated Hospital of General Hospital of People's Liberation Army. The patients underwent intracranial electrode implantation as part of their clinical evaluation for epilepsy surgery when noninvasive studies could not adequately localize the origin of their seizures. All patients had

normal or corrected-to-normal vision and had no history of head trauma or encephalitis. Preoperative MRI did not reveal structural abnormalities in the amygdala and mPFC for all patients. All patients provided written informed consent. This research protocol was approved by the institutional review boards at the hospital sites.

All patients apart from one completed the same number of trials. Patient 7 only conducted 26 CS+ and 26 CS− trials due to time constraints. For this reason, we excluded the data from this patient from all trial-level analyses. Since this patient was only implanted in the amygdala and not in any of the mPFC subregions, data from this patient were also not used for connectivity analyses. The data from this patient were thus only included in the analysis of the time-frequency (power) effects in the amygdala and the BOSC-based analyses of averaged detection times of theta oscillations and theta onsets in the amygdala.

Electrode reconstruction

Postoperative computed tomography (CT) images were coregistered to preoperative T1-weighted MR images using FreeSurfer (v6.0.0, http://surfer.nmr.mgh.harvard.edu/). The registration was visually verified and manually adjusted if necessary. The implanted electrodes were reconstructed using the stereotactic localization software (*54*). Then, all electrodes' coordinates were mapped onto a standard MNI space (Fig. 1B).

Electrode localizations in regions of interest

For determining the exact electrode locations in the amygdala, we used an automatic parcellation via FreeSurfer (*55*) and visually verified the accuracy of the procedure. The dmPFC was delineated to include bilateral areas 32d and 9, while the vmPFC included bilateral areas 11m and 14m according to the medial prefrontal template (*56*). Given that the medial prefrontal templates were depicted in the standard MNI space, the electrodes were arranged to each area of the template based on their MNI coordinates (table S2).

SCR data acquisition and analysis

SCR was recorded using a BIOPAC system and measured with two steel disk electrodes attached to the middle phalanges of the index and middle fingers of the left hand. Data were recorded at a sampling rate of 200 Hz and low pass–filtered to 10 Hz online, followed with a 1-Hz low-pass filtering. After visual inspection for artifacts, the amplitude of SCR was calculated in each trial as the base-to-peak difference in skin conductance of the largest deflection (in microsiemens) within a window of 500 to 6000 ms after stimulus onset, using an open-source toolbox (https://github.com/mateusjoffily/EDA). The minimal response criterion was 0.02 μS (*57*). Responses below this threshold were encoded as zero. The raw SCR scores were scaled according to each participant's maximum SCR value across all CS trials (*11*, *58*). Since we were interested in whether patients successfully acquired fear responses to the CS, only SCRs during the second half of all trials were averaged to assess fear acquisition. Four patients who failed to discriminate between CS+ and CS− in SCR (difference < 0.05 μS) were used for control analysis (Fig. 1C and table S1). This resulted in a total of nine patients (five females, average age: mean = 26.44 years, SD = 5.9) who were included into all electrophysiological data analysis. Two participants in whom SCR recordings failed because of equipment problems but orally reported successful fear acquisition were also included in electrophysiological data analysis. Group results remained the same when these two patients were excluded (fig. S10). Unless otherwise specified, only participants with successful fear acquisition were included in iEEG data analysis.

For the trial-by-trial SCR analysis, we used a PsPM tool (*59*). The data were modeled using dynamic causal modeling (nonlinear model), which allows estimating the amplitudes of anticipatory and evoked SCR under a canonical response function (*60*). After trimming the time series, the SCR data were filtered with a first-order bidirectional bandpass Butterworth filter (cutoff frequencies, 0.016 to 5 Hz) and further down-sampled to 10 Hz. For each trial, we estimated (i) a fixed-dispersion response at CS onset (i.e., CS trials with and without US), (ii) a fixed-latency response at US onset, and (iii) a flexible latency that comprised the time window between CS offset and US onset (minus 1 s, to allow disambiguation of the US response from the CS response). The data for all trials and conditions were interpolated with last observation carried backward. This was performed because after the cue is presented and the outcome received, the associative strength for the cue will immediately be updated, but it will only be expressed behaviorally when the subject is presented with the cue again (*61*, *62*). To prevent extreme responses from biasing calculations, extreme values with at least 5 SD above or below the mean were excluded. In addition, only nonreinforced trials (CS + US−) were included in the analysis to avoid contamination by the response to the US.

We applied PsPM to SCR data to detect learning effects because of its higher sensitivity for detecting learning-related condition differences (*21*, *22*). PsPM is modeling the data using a canonical response function of sudomotor activity and uses a bidirectional filter to mitigate the effects of unspecific noise (e.g., slow drift components). The values (parameter estimates) that are shown in the PsPM plot represent the goodness of fit to this response function. Because of the more indirect analysis approach, apparent linear increases in PsPM-based parameter estimates should not be interpreted as reflecting increases in SCR but only of predictability by the model. Thus, we fit and validated the computational model of cue-reinforcer associability based on the raw SCR values derived from a conventional peak scoring method rather than based on the PsPM parameter estimates, which also made the model input more consistent with previous studies (*20*, *28*, *29*).

Statistical analysis was performed using LME model in which learning was predicted from trial number and condition (fixed effects), with patients as random effect. In addition, a separate model was conducted by adding two polynomials (first and second degree) in the regression equation to test for linear and quadratic trends.

iEEG: Data collection and preprocessing

iEEG data were acquired using a Nihon-Kohden recording system (Yuquan Hospital) and a Blackrock Neuroport system (First Affiliated Hospital of General Hospital of People's Liberation Army) at a sampling rate of 2000 Hz. All the data were analyzed in MATLAB combined with open-source toolboxes. After acquisition, iEEG data were band pass–filtered from 0.6 to 200 Hz using a zero–phase delay finite impulse response filter and down-sampled to 500 Hz. Noise (50 Hz) and its harmonics were removed using a notch filter. The resulting data were rereferenced per contact to one contact located in white matter that was close to gray matter contacts (*63*, *64*). Specifically, for each electrode targeting the regions of interests, candidate reference contacts were chosen from contacts located in white matter by visual inspection of the coregistered CT. The filtered iEEG data from candidate contacts were then visually inspected one by one, and a

single contact with little or no apparent EEG activity was chosen as the reference for other contacts.

iEEG data analyses

Epochs from −4500 ms before to 5000 ms after CS onset (plus additional "buffer" window of 3000 ms before and after the window of interest) were extracted from the preprocessed iEEG data. To avoid interictal activity or even subclinical or clinical seizures that may confound the results, we carefully verified that no subclinical or even clinical seizures occurred during or immediately before the experiment. In addition, we conducted a multistep procedure to remove all data points that were affected by interictal events or other artifacts. First, epochs contaminated by epileptic activities were identified by an automated detection algorithm (*65*). The detection algorithm discriminates signals containing interictal discharges from background activity based on adaptive modeling of signal envelopes within the range from 10 to 60 Hz that is primarily affected by interictal discharges. Second, after removal of epochs contaminated by epileptic activity, outliers in the remaining epochs were labeled and removed if their voltage value surpassed 5 SD of all epochs at any time point. Third, channels with more than 20% of epochs containing either epileptic discharges or labeled as bad trials were excluded from further analysis. Fourth, we screened all data visually and removed all trials with remaining artifacts (this was performed in a manner that was blind to the experimental conditions).

Time-frequency transformations were performed using the multitaper method implemented in FieldTrip. Briefly, the time-frequency analysis was performed for each epoch using a multitaper power spectral density estimation using a frequency-dependent time window of five cycles (i.e., window size decreasing with frequency increasing) and a time-bandwidth product of 2 (three tapers). In this study, we mainly focused on electrophysiological activity from CS onset to 2000 ms afterward (*11*). Each epoch was baseline corrected (−1500 to −500 ms before CS onset, to avoid smearing of changes in theta power between pre- and post-CS time bins). Then, theta power was calculated by averaging the normalized power over 4 to 8 Hz. For the analysis of the distribution of peak frequencies, we extracted the peak frequency from the power spectrum of each electrode contact to generate the electrode-wise distribution.

We used the modified BOSC method to identify rhythmic activity at specific frequencies within the background of nonrhythmic EEG components (*25*). The BOSC analysis was applied on the signals recorded from each electrode contact and kept the frequency band that fell within a frequency range of 1 to 30 Hz. For each contact, we also focused on the 2000-ms period after CS onset. For a given frequency f, a time period was defined as containing rhythmic activity if the wavelet power at f exceeded a power threshold PT(f) and lasted for more than a duration threshold (set to four cycles, DT = 4/f). The background spectrum was estimated by first removing points greater than 1 SD of the spectrum's linear fitting residual in log-log coordinates (modified based on eBOSC) (*66*) then refitting the remaining spectrum with linear regression. PT(f) was set to the 95th percentile of this theoretical probability distribution. For each iEEG contact, the proportion of time in which oscillatory signals at a given frequency f was detected. We created a summary plot of the proportion of time in which theta oscillations were detected and extracted the onset of detected theta oscillation across all contacts. We also calculated the proportion of time showing simultaneous theta oscillations between the amygdala and dmPFC and between the amygdala and vmPFC.

For coherence analyses, a method based on imaginary coherence was used to avoid the effect of a common input on the interregional neural synchronization (*2*). Imaginary coherence was calculated as

$$iCoh_{xy} = \frac{|\,Im(n^{-1}\sum_{n=1}^{n=N} S_{xy})\,|}{\sqrt{(n^{-1}\sum_{n=1}^{n=N} S_{xx})\,.\,(n^{-1}\sum_{n=1}^{n=N} S_{yy})}}$$

where S_{xy} denotes the cross-spectral density between activities at electrodes x and y, S_{xx} and S_{yy} are the autospectral densities for electrodes x and y, N denotes the time bins, and $iCoh$ denotes the absolute imaginary coherence. Cross- and autospectral densities were calculated with Morlet wavelets (five cycles) over 2000 ms after CS onset.

As an index of dominant unidirectional interaction, PSI indicates the direction of coupling between two signals. Given a prespecified bandwidth parameter, it reflects the change of phase difference between adjacent frequency bins and is weighted by the magnitude of the coherence. The PSI is defined as

$$\widetilde{\Psi}_{x,y} = Im\Big(\sum_{f\in F} C_{x,y}^{*}(f)\, C_{x,y}(f+\delta f)\Big), F \in (f-w, f+w)$$

where $C_{x,y}$ is the complex coherence, δf is the frequency resolution, f is the center frequency (6 Hz) of the targeted frequency range (4 to 8 Hz), and Im denotes the imaginary part. We used $2w$ as the bandwidth (4 Hz) for which the phase slope was calculated and chose it to be four times of the frequency resolution δf (1 Hz). To understand the directionality between two signals involved in fear learning, we segmented the phase of the modulating signal (amygdala) and of the modulated signal (dmPFC and vmPFC) into N epochs and used them as inputs to calculate the PSI for both CS+ and CS− conditions. The sign of the PSI informs which signal is temporally leading the other one. For example, when the phase differences between the modulating signal A and the modulated signal B increase with frequencies, a positive slope of the phase spectrum is expected (A➔B). By performing the PSI analysis with a sliding window of 1000 ms, spaced at 20 ms (98% overlap), we were able to track the time course of directionality between two signals.

ROC analyses were conducted using the frequency with the highest percentage of theta oscillation time detected by BOSC and power in the theta band. For the first variable, we extracted the frequency with the highest percentage of theta oscillations of each electrode contact and linked each value to the diagnosis—brain region 1 or brain region 2 (amygdala or dmPFC, amygdala or vmPFC, and vmPFC or dmPFC). We then conducted an ROC analysis using the ROC function in MATLAB to explore whether an ideal observer could predict which brain region it belongs to when relying only on this variable (the frequency with the highest percentage of theta oscillation times). With regard to theta power, we extracted the theta power difference (CS+ minus CS−) of each electrode contact and linked each value to the diagnosis "learner or nonlearner." Then, we conducted an ROC analysis to explore whether theta power could predict whether a given subject is a learner or a nonlearner. Accuracy was measured as the area under the ROC curve. We then conducted a permutation test in which we generated a null distribution by randomly shuffling labels of brain regions (for variable 1) or learner/nonlearner (for variable 2), computed the AUC across all patients, and repeated this procedure 1000 times. The 95th percentile threshold was extracted, and AUCs above this threshold were considered significant. Unless

otherwise specified, iEEG data analyses were based on the second half of trials.

Computational modeling

We fitted and validated an associability model behaviorally using SCR data. In this model, x_n indicates the CS on trial n (CS+ or CS−), and r_n is the US (1 for US, 0 for no US). Predictions of value (i.e., occurrence of electric stimulation) $V_n(x_n)$ were defined for each stimulus type (x_n) on trial n. The prediction error $\delta_n = r_n - V_n(x_n)$ measures the difference between the actual and predicted electric stimulation on trial n. We used a dynamic learning rate that gated the speed of learning based on the Pearce-Hall associability rule (*20*, *28*, *29*, *67*). The resulting model for fear conditioning was as follows

$$V_{n+1}(x_n) = V_n(x_n) + k\,\alpha_n(x_n)\,\delta_n$$
$$\alpha_{n+1}(x_n) = \eta\,|\,\delta_n\,| + (1-\eta)\,\alpha_n(x_n)$$

where η indicates the weight assigned to the most recent absolute value of the prediction error (indicating the accuracy of value prediction) in the conditioning phase, and k indicates a normalization factor. The value V_0 and the associability α_0 were initially set to 0.5.

We tested the fit of the above model by minimizing the difference between model-predicted associability and observed SCR. We optimized the free parameters of the model (η and k) by maximizing the probability of observing the measured sequence of SCRs following each CS. This maximization was achieved via the maximum likelihood estimation. Likelihoods were performed over all trials but omitting trials paired with electric stimulation to avoid possible contamination due to stimulation artifacts. Then, we identified associability-related brain activity by conducting an LME model with theta power in the three regions as dependent variables, associability as fixed effect, and subject and (nested) trial as random effects for assessing correlations between associability and theta power.

To further test whether learning dynamics in our study can be better explained by the associability model than by alternative models, we compared the fit of the Rescorla-Wagner (RW) model of learning by prediction errors (in which learning rate is a constant) to the trial-by-trial skin conductance responses with the fit to the augmented hybrid models that gated prediction error–driven learning by associability. We fit these models separately to each individual subject's SCRs and performed likelihood ratio tests on the data aggregated across subjects. The classical likelihood ratio tests were based on the verification of the null hypothesis that the improvement in fit of the more complicated model relative to the simpler one was only due to the higher number of predictors. We found that the SCR scores were best explained by the inclusion of associability α and the value V as additional predictors to explain variance in SCRs [Hybrid (α + V) versus RW (V), $\chi_3^2 = 73.73$, $P = 6^{-16}$; versus Hybrid(V), $\chi_1^2 = 35.86$, $P = 2^{-9}$; versus Hybrid(α), $\chi_1^2 = 36.55$, $P = 1^{-9}$]. Thus, this result indicates that the associability-based model best accounts for fear learning in our study.

Statistical analysis

An LME model was used for statistical analysis and implemented using the *fitlme* function in MATLAB. Given that the LME model can handle the case where the predictor variables are not independent, the information of all contacts can be effectively used (*68*).

We implemented LME model with patient and electrode as two random effects and used the restricted maximum likelihood method to estimate LME parameters. Post hoc tests of P values were performed using Bonferroni correction to correct for multiple comparisons. To assess the effects of condition (CS+ versus CS−) on theta power, coherence, and PSI, we compared CS+ with CS−

$$Y \sim \text{condition} + (1\,|\,\text{subject}) + (1\,|\,\text{subject:electrode})$$

where the dependent variable Y is either theta power, imaginary coherence, or PSI. The condition variable is coded as 1 for CS− and 2 for CS+.

To assess the effects of learning dynamics on theta power, coherence, theta onsets detected by BOSC, theta peak frequency detected by BOSC, and overlap between time periods showing simultaneous theta oscillations in amygdala and vmPFC/dmPFC

$$Y \sim \text{trial/block} + (1\,|\,\text{subject}) + (1\,|\,\text{subject:electrode})$$

where the dependent variable Y is either theta power, imaginary coherence, or overlapping time periods. The "trial/block" variable is coded as the trial/block number.

To assess the relationship between associability of computational modeling and SCR data or theta power in the three brain regions

$$Y \sim \text{associability} + (1\,|\,\text{subject}) + (1\,|\,\text{subject:trial})$$

where the dependent variable Y is either theta power or SCR data.

To compare temporal dynamics of theta oscillation between different brain regions

$$Y \sim \text{area} + (1\,|\,\text{subject}) + (1\,|\,\text{subject:electrode})$$

where the dependent variable Y is the onset time of theta oscillation detected by BOSC.

A Wilcoxon signed-rank test was used to access successful fear acquisition in SCR. Spearman's correlation was used to calculate the relationship between model-based associability and theta power across trials as well as dynamic changes in both theta power and imaginary coherence across trials.

Cluster-based permutation tests were performed by shuffling condition labels for data epochs to create distributions under the null hypothesis. Specifically, in each permutation, the CS+/CS− condition labels were shuffled between data epochs while keeping the labels of electrode contacts and subjects unchanged. LME model was then applied to the time-frequency spectrograms or imaginary coherograms derived from shuffled data. A P value of 0.05 was chosen to find out significant points. Clusters were then extracted, and the largest cluster's size was passed to the null distribution. This procedure was repeated 1000 times. Cluster sizes derived from the unshuffled data epochs were compared to the 95th percentile of the null distribution, and those with a size larger than this threshold were labeled as significant clusters.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/34/eabf4198/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. G. Buzsáki, C. A. Anastassiou, C. Koch, The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* **13**, 407–420 (2012).

2. N. Karalis, C. Dejean, F. Chaudun, S. Khoder, R. R. Rozeske, H. Wurtz, S. Bagur, K. Benchenane, A. Sirota, J. Courtin, C. Herry, 4-Hz oscillations synchronize prefrontal–amygdala circuits during fear behavior. *Nat. Neurosci.* **19**, 605–612 (2016).
3. A. H. Taub, R. Perets, E. Kahana, R. Paz, Oscillations synchronize amygdala-to-prefrontal primate circuits during aversive learning. *Neuron* **97**, 291–298.e3 (2018).
4. J. M. Stujenske, E. Likhtik, M. A. Topiwala, J. A. Gordon, Fear and safety engage competing patterns of theta-gamma coupling in the basolateral amygdala. *Neuron* **83**, 919–933 (2014).
5. K. A. Corcoran, G. J. Quirk, Activity in prelimbic cortex is necessary for the expression of learned, but not innate, fears. *J. Neurosci.* **27**, 840–844 (2007).
6. E. Likhtik, J. M. Stujenske, M. A. Topiwala, A. Z. Harris, J. A. Gordon, Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. *Nat. Neurosci.* **17**, 106–113 (2014).
7. M. R. Milad, G. J. Quirk, R. K. Pitman, S. P. Orr, B. Fischl, S. L. Rauch, A role for the human dorsal anterior cingulate cortex in fear expression. *Biol. Psychiatry* **62**, 1191–1194 (2007).
8. M. R. Milad, C. I. Wright, S. P. Orr, R. K. Pitman, G. J. Quirk, S. L. Rauch, Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol. Psychiatry* **62**, 446–454 (2007).
9. K. S. LaBar, J. C. Gatenby, J. C. Gore, J. E. LeDoux, E. A. Phelps, Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron* **20**, 937–945 (1998).
10. E. A. Phelps, M. R. Delgado, K. I. Nearing, J. E. LeDoux, Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron* **43**, 897–905 (2004).
11. M. F. J. Sperl, C. Panitz, I. M. Rosso, D. G. Dillon, P. Kumar, A. Hermann, A. E. Whitton, C. Hermann, D. A. Pizzagalli, E. M. Mueller, Fear extinction recall modulates human frontomedial theta and amygdala activity. *Cereb. Cortex* **29**, 701–715 (2018).
12. E. M. Mueller, C. Panitz, C. Hermann, D. A. Pizzagalli, Prefrontal oscillations during recall of conditioned and extinguished fear in humans. *J. Neurosci.* **34**, 7059–7066 (2014).
13. M. R. Milad, G. J. Quirk, Fear extinction as a model for translational neuroscience: Ten years of progress. *Annu. Rev. Psychol.* **63**, 129–151 (2012).
14. I. Vidal-Gonzalez, B. Vidal-Gonzalez, S. L. Rauch, G. J. Quirk, Microstimulation reveals opposing influences of prelimbic and infralimbic cortex on the expression of conditioned fear. *Learn. Mem.* **13**, 728–733 (2006).
15. D. Sierra-Mercado, N. Padilla-Coreano, G. J. Quirk, Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. *Neuropsychopharmacology* **36**, 529–538 (2011).
16. J. Zheng, R. F. Stevenson, B. A. Mander, L. Mnatsakanyan, F. P. K. Hsu, S. Vadera, R. T. Knight, M. A. Yassa, J. J. Lin, Multiplexing of theta and alpha rhythms in the amygdala-hippocampal circuit supports pattern separation of emotional information. *Neuron* **102**, 887–898.e5 (2019).
17. H. Oya, H. Kawasaki, M. A. Howard III, R. Adolphs, Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *J. Neurosci.* **22**, 9502–9512 (2018).
18. J. Zheng, K. L. Anderson, S. L. Leal, A. Shestyuk, G. Gulsen, L. Mnatsakanyan, S. Vadera, F. P. K. Hsu, M. A. Yassa, R. T. Knight, J. J. Lin, Amygdala-hippocampal dynamics during salient information processing. *Nat. Commun.* **8**, 14413 (2017).
19. C. Herry, J. P. Johansen, Encoding of fear learning and memory in distributed neuronal circuits. *Nat. Neurosci.* **17**, 1644–1654 (2014).
20. J. E. Dunsmoor, M. C. W. Kroes, J. Li, N. D. Daw, H. B. Simpson, E. A. Phelps, Role of human ventromedial prefrontal cortex in learning and recall of enhanced extinction. *J. Neurosci.* **39**, 3264–3276 (2019).
21. D. R. Bach, F. Melinscak, Psychophysiological modelling and the measurement of fear conditioning. *Behav. Res. Ther.* **127**, 103576 (2020).
22. M. Staib, G. Castegnetti, D. R. Bach, Optimising a model-based approach to inferring fear learning from skin conductance responses. *J. Neurosci. Methods* **255**, 131–138 (2015).
23. J. Lesting, R. T. Narayanan, C. Kluge, S. Sangha, T. Seidenbecher, H.-C. Pape, Patterns of coupled theta activity in amygdala-hippocampal-prefrontal cortical circuits during fear extinction. *PLOS ONE* **6**, e21714 (2011).
24. N. A. Herweg, A. D. Sharan, M. R. Sperling, A. Brandt, A. Schulze-Bonhage, M. J. Kahana, Reactivated spatial context guides episodic recall. *J. Neurosci.* **40**, 2119–2128 (2020).
25. T. A. Whitten, A. M. Hughes, C. T. Dickson, J. B. Caplan, A better oscillation detection method robustly extracts EEG rhythms across brain state changes: The human alpha rhythm as a test case. *Neuroimage* **54**, 860–874 (2011).
26. J. M. Pearce, G. Hall, A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).
27. R. A. Rescorla, A. R. Wagner, A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. in *Classical Conditioning II: Current Research and Theory*, A. H. Black, W. F. Prokasy, Eds. (Appleton-Century-Crofts, 1972), pp. 64–99.
28. J. Li, D. Schiller, G. Schoenbaum, E. A. Phelps, N. D. Daw, Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.* **14**, 1250–1252 (2011).
29. C. M. Raio, C. A. Hartley, T. A. Orederu, J. Li, E. A. Phelps, Stress attenuates the flexible updating of aversive value. *Proc. Natl. Acad. Sci.* **114**, 11241–11246 (2017).
30. J. Courtin, F. Chaudun, R. R. Rozeske, N. Karalis, J. Baufreton, H. Wurtz, C. Gonzalez-Campo, T. C. M. Bienvenu, C. Herry, A. Abdi, Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* **505**, 92–96 (2014).
31. D. Popa, S. Duvarci, A. T. Popescu, C. Léna, D. Paré, Coherent amygdalocortical theta promotes fear memory consolidation during paradoxical sleep. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6516–6519 (2010).
32. N. A. Herweg, E. A. Solomon, M. J. Kahana, Theta oscillations in human memory. *Trends Cogn. Sci.* **24**, 208–227 (2020).
33. J. Minxha, R. Adolphs, S. Fusi, A. N. Mamelak, U. Rutishauser, Flexible recruitment of memory-based choice representations by the human medial frontal cortex. *Science* **368**, eaba3313 (2020).
34. M. Nishida, J. Pearsall, R. L. Buckner, M. P. Walker, REM sleep, prefrontal theta, and the consolidation of human emotional memory. *Cereb. Cortex* **19**, 1158–1166 (2009).
35. M. Ozawa, P. Davis, J. Ni, J. Maguire, L. Reijmers, T. Papouin, L. Reijmers, Experience-dependent resonance in amygdalo-cortical circuits supports fear memory retrieval following extinction. *Nat. Commun.* **11**, 4358 (2020).
36. U. Livneh, R. Paz, Amygdala-prefrontal synchronization underlies resistance to extinction of aversive memories. *Neuron* **75**, 133–142 (2012).
37. M. R. Gilmartin, N. L. Balderston, F. J. Helmstetter, Prefrontal cortical regulation of fear learning. *Trends Neurosci.* **37**, 455–464 (2014).
38. B. A. Vogt, G. Paxinos, Cytoarchitecture of mouse and rat cingulate cortex with human homologies. *Brain Struct. Funct.* **219**, 185–192 (2014).
39. K. D. Davis, W. D. Hutchison, A. M. Lozano, R. R. Tasker, J. O. Dostrovsky, Human anterior cingulate cortex neurons modulated by attention-demanding tasks. *J. Neurophysiol.* **83**, 3575–3577 (2000).
40. M. Brinley-Reed, F. Mascagni, A. J. McDonald, Synaptology of prefrontal cortical projections to the basolateral amygdala: An electron microscopic study in the rat. *Neurosci. Lett.* **202**, 45–48 (1995).
41. E. Likhtik, J. G. Pelletier, R. Paz, D. Paré, Prefrontal control of the amygdala. *J. Neurosci.* **25**, 7429–7437 (2005).
42. J. Courtin, N. Karalis, C. Gonzalez-Campo, H. Wurtz, C. Herry, Persistence of amygdala gamma oscillations during extinction learning predicts spontaneous fear recovery. *Neurobiol. Learn. Mem.* **113**, 82–89 (2014).
43. E. Knapska, M. Mikosz, A. Nowak, D. Owczarek, M. Sheng, S. Maren, M. Wawrzyniak, T. Werka, M. Macias, I. A. Cymerman, M. Pieprzyk, L. Kaczmarek, J. Jaworski, L. Kaczmarek, Functional anatomy of neural circuits regulating fear and extinction. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17093–17098 (2012).
44. D. A. Gutman, O. P. Keifer, M. E. Magnuson, D. C. Choi, W. Majeed, S. Keilholz, K. J. Ressler, A DTI tractography analysis of infralimbic and prelimbic connectivity in the mouse using high-throughput MRI. *Neuroimage* **63**, 800–811 (2012).
45. C. Strobel, R. Marek, H. M. Gooch, R. K. P. Sullivan, P. Sah, Prefrontal and auditory input to intercalated neurons of the amygdala. *Cell Rep.* **10**, 1435–1442 (2015).
46. P. Domenech, S. Rheims, E. Koechlin, Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. *Science* **369**, eabb0184 (2020).
47. L. K. Fellows, M. J. Farah, The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? *Cereb. Cortex* **17**, 2669–2674 (2007).
48. A. Bechara, D. Tranel, H. Damasio, Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* **123**, 2189–2202 (2000).
49. M. Koenigs, L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, A. Damasio, Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* **446**, 908–911 (2007).
50. M. Hsu, M. Bhatt, R. Adolphs, D. Tranel, C. F. Camerer, Neural systems responding to degrees of uncertainty in human decision-making. *Science* **310**, 1680–1683 (2005).
51. R. M. Visser, H. S. Scholte, T. Beemsterboer, M. Kindt, Neural pattern similarity predicts long-term fear memory. *Nat. Neurosci.* **16**, 388–390 (2013).
52. A. Etkin, C. Büchel, J. J. Gross, The neural bases of emotion regulation. *Nat. Rev. Neurosci.* **16**, 693–700 (2015).
53. M. Davis, P. J. Whalen, The amygdala: Vigilance and emotion. *Mol. Psychiatry* **6**, 13–34 (2001).
54. C. Qin, Z. Tan, Y. Pan, Y. Li, L. Wang, L. Ren, W. Zhou, L. Wang, Automatic and precise localization and cortical labeling of subdural and depth intracranial electrodes. *Front. Neuroinform.* **11**, 10 (2017).
55. R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, R. J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).

56. F. X. Neubert, R. B. Mars, J. Sallet, M. F. S. Rushworth, Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2695–E2704 (2015).
57. D. Schiller, J. W. Kanen, J. E. Ledoux, M. Mon, E. A. Phelps, Extinction during reconsolidation of threat memory diminishes prefrontal cortex involvement. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20040–20045 (2013).
58. D. T. Lykken, P. H. Venables, Direct measurement of skin conductance:a proposal for standardization. *Psychophysiology* **8**, 656–672 (1971).
59. D. R. Bach, K. J. Friston, R. J. Dolan, An improved algorithm for model-based analysis of evoked skin conductance responses. *Biol. Psychol.* **94**, 490–497 (2013).
60. D. R. Bach, J. Daunizeau, K. J. Friston, R. J. Dolan, Dynamic causal modelling of anticipatory skin conductance responses. *Biol. Psychol.* **85**, 163–170 (2010).
61. D. R. Bach, M. Näf, M. Deutschmann, S. K. Tyagarajan, B. B. Quednow, Threat memory reminder under matrix metalloproteinase 9 inhibitor doxycycline globally reduces subsequent memory plasticity. *J. Neurosci.* **39**, 9424–9434 (2019).
62. P. Homan, Q. Lin, J. W. Murrough, L. Soleimani, D. R. Bach, R. L. Clem, D. Schiller, Prazosin during threat discrimination boosts memory of the safe stimulus. *Learn. Mem.* **24**, 597–601 (2017).
63. G. Arnulfo, J. Hirvonen, L. Nobili, S. Palva, J. M. Palva, Phase and amplitude correlations in resting-state activity in human stereotactical EEG recordings. *Neuroimage* **112**, 114–127 (2015).
64. D. Chen, L. Kunz, W. Wang, H. Zhang, W. X. Wang, A. Schulze-Bonhage, P. C. Reinacher, W. Zhou, S. Liang, N. Axmacher, L. Wang, Hexadirectional modulation of theta power in human entorhinal cortex during spatial navigation. *Curr. Biol.* **28**, 3310–3315.e4 (2018).
65. R. Janca, P. Jezdik, R. Cmejla, M. Tomasek, G. A. Worrell, M. Stead, J. Wagenaar, J. G. R. Jefferys, P. Krsek, V. Komarek, P. Jiruska, P. Marusic, Detection of interictal epileptiform discharges using signal envelope distribution modelling: Application to epileptic and non-epileptic intracranial recordings. *Brain Topogr.* **28**, 172–183 (2015).
66. J. Q. Kosciessa, T. H. Grandy, D. D. Garrett, M. Werkle-bergner, Single-trial characterization of neural rhythms: Potential and challenges. *Neuroimage* **206**, 116331 (2020).
67. M. E. Le Pelley, I. P. L. McLaren, Learned associability and associative change in human causal learning. *Q. J. Exp. Psychol. B* **56**, 68–79 (2003).
68. R. Thompson, H. D. Patterson, Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554 (1971).

Acknowledgments: We are grateful to all the participants who volunteered to participate in this study. We thank Y. Xue for help with SCR data acquisition, J. Li for providing computational modeling methods, and S. Qin for helpful discussion. **Funding:** This study was supported by the Strategic Priority Research Program of Chinese Academy of Science (XDB32010300), the National Key Research and Development Program of China (2016YFC1307200), the Beijing Municipal Science and Technology Commission (Z171100000117014), the National Natural Science Foundation of China (31771255 and 32020103009), and CAS Interdisciplinary Innovation Team (JCTD-2018-07). N.A. received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Projektnummer 316803389–SFB 1280 and via Projektnummer 122679504–SFB 874. **Author contributions:** Conceptualization: S.C., Z.T., and L.W. Data collection: S.C., W.X., W.Z., S.L., and X.Z. Data analysis: S.C., Z.T., C.A.G., N.A., and L.W. Writing—original draft: S.C. Writing—review and editing: S.C., N.A., and L.W. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 28 October 2020
Accepted 29 June 2021
Published 18 August 2021
10.1126/sciadv.abf4198

Citation: S. Chen, Z. Tan, W. Xia, C. A. Gomes, X. Zhang, W. Zhou, S. Liang, N. Axmacher, L. Wang, Theta oscillations synchronize human medial prefrontal cortex and amygdala during fear learning. *Sci. Adv.* **7**, eabf4198 (2021).