#### **ORIGINAL PAPER**



# Conflict Dynamics of Post-Retrieval Extinction: A Comparative Analysis of Unconditional and Conditional Reminders Using Skin Conductance Responses and EEG

Dong-ni Pan<sup>1,2,3</sup> · Delhii Hoid<sup>2,4</sup> · Oliver T. Wolf<sup>3</sup> · Christian J. Merz<sup>3</sup> · Xuebing Li<sup>2</sup>

Received: 1 December 2023 / Accepted: 4 April 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

#### Abstract

The post-retrieval extinction paradigm, rooted in reconsolidation theory, holds promise for enhancing extinction learning and addressing anxiety and trauma-related disorders. This study investigates the impact of two reminder types, mild US-reminder (US-R) and CS-reminder (CS-R), along with a no-reminder extinction, on fear recovery prevention in a categorical fear conditioning paradigm. Scalp EEG recordings during reminder and extinction processes were conducted in a three-day design. Results show that the US-R group exhibits a distinctive extinction learning pattern, characterized by a slowed-down yet successful process and pronounced theta-alpha desynchronization (source-located in the prefrontal cortex) during CS processing, followed by enhanced synchronization (source-located in the anterior cingulate) after shock cancellation in extinction trials. These neural dynamics correlate with the subtle advantage of US-R in the Day 3 recovery test, presenting faster spontaneous recovery fading and generally lower fear reinstatement responses. Conversely, the CS reminder elicits CS-specific effects in later episodic tests. The unique neural features of the US-R group suggest a larger prediction error and subsequent effortful conflict learning processes, warranting further exploration.

Keywords Post-Retrieval Extinction · Fear Conditioning · Reconsolidation · EEG Oscillation · Prediction Error

# Introduction

Reactivating the original memory is believed to open the door to memory destabilization and subsequent "reconsolidation," wherein various manipulations can update the original memory trace, leading to diverse memory consequences (Lee et al. 2017). In the realm of fear conditioning, exposure

Handling Editor: Juha Silvanto

⊠ Xuebing Li lixb@psych.ac.cn

- <sup>1</sup> School of Psychology, Beijing Language and Culture University, Beijing 100083, China
- <sup>2</sup> Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, No 16 Lincui Rd Chaoyang District, Beijing 100101, China
- <sup>3</sup> Department of Cognitive Psychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, 44801 Bochum, Germany
- <sup>4</sup> Department of Psychology, School of Social Sciences, Tsinghua University, Beijing 100083, China

to a reminder of the conditioning experience (CS +) purportedly destabilizes a fear memory, making it susceptible to disruption through pharmacological treatments that induce fear "erasure" in rodents (Nader et al. 2000). This concept was later adapted to drug-free procedures (Monfils et al. 2009), where presenting the retrieval cue (CS +) before classical extinction training was found to enhance extinction learning and prevent fear relapse. This approach quickly extended to human participants (Schiller et al. 2010).

However, despite early replications (Bjorkstrand et al. 2016; Agren et al. 2012; Schiller et al. 2013; Johnson and Casey 2015), the robustness and generalizability of the superior effect of post-retrieval extinction training in humans are facing challenges. Evidence suggests that a simple CS + reminder before extinction training may not effectively prevent fear relapse for fear-relevant stimuli (Fricchione et al. 2016; Golkar et al. 2017, 2012) or in more generalized situations (Kroes et al. 2017; Drexler et al. 2014; Soeter and Kindt 2011). Moreover, preregistered studies failed to fully replicate the CS + reminder effect in an independent sample (Chalkia et al. 2010; Chalkia et al. 2020), underscoring the

need for a cautious approach toward the CS + reminder's advantage in preventing fear relapse.

Simultaneously, an alternative paradigm has been proposed, using unconditioned stimuli (US) with lower intensity as the reminder cue. This approach, known as the US-R paradigm, posits that subsequent extinction training can affect the fear response of all conditioned stimuli (CSs) associated with the US, potentially enhancing the efficiency of post-retrieval effects on the original fear memory (Liu et al. 2014). While the US-R paradigm shows promise in reducing cue specificity of memory reactivation (Dunbar and Taylor 2017a, 2017b) and has demonstrated the elimination of the return of extinguished fear responses to fear-relevant stimuli (Thompson and Lipp 2017), empirical evidence for this paradigm remains relatively limited.

Despite the paradigm of post-retrieval extinction being derived from the concept of reconsolidation, distinctions have been noted in the conceptual hierarchy between the reminder-dependent effect and reconsolidation processing (Schroyens et al. 2021; Hardwicke et al. 2016). Researchers face challenges in confirming whether the post-reminder effect, observed solely through behavioural observation, is attributable to reconsolidation. This dilemma has led to a call for evidence from neural/brain aspects to enrich the understanding of the real contribution of the post-reminder advantage (Haubrich et al. 2020).

Electroencephalography (EEG) technology presents an opportunity to explore these exciting inquiries due to its non-invasive nature and high temporal resolution, allowing for a close examination of specific cognitive processes. While EEG has not been previously applied to post-reminder extinction protocols, incorporating neural oscillation features into this research field could shed light on the underlying neural conditions of assumed fear memory destabilization and updating.

Specifically, theta band phase synchronization is recognized as a crucial mechanism facilitating synaptic plasticity and interregional communication within the brain (Fell and Axmacher 2011), demonstrating an increase in theta-phase synchronization during memory encoding and retrieval in rodents (Benchenane et al. 2010; Place et al. 2016) and humans (Summerfield and Mangels 2005; Weiss and Rappelsberger 2000). In the context of fear memory processing, numerous studies involving animals as well as intracranial electroencephalography recordings indicate the importance of synchronization in the theta band during different stages of fear acquisition and extinction, within and between key structures of the fear circuitry, such as the amygdala, hippocampus, and regions of the medial prefrontal cortex (mPFC) (Taub et al. 2018; Chen 2021; Lesting et al. 2011). In human studies, utilizing electroencephalogram and source estimation methods, fear expression was found specifically associated with theta activity in anterior midcingulate cortex (Mueller et al. 2014), thereby bridges findings from prior rodent research and human neuroimaging studies. Combined with fMRI techniques, changes in theta power were also observed to correspond with alterations in amygdala activation. Theta activity associated with fear and extinction recall accounted for 60% of the variance observed in the right amygdala (Sperl et al. 2019). These oscillation and synchronized activities facilitate precisely timed neural processes essential for synaptic plasticity (Bocchio et al. 2017; Seidenbecher et al. 2003), not only play a role of incidental signature of memory (Hanslmayr et al. 2019), but a causal glue that binds associative memory formation (Clouter et al. 2017).

In addition, alpha oscillations are also closely related to human emotional learning and unbinding. For instance, research has demonstrated that even successful pattern separation of emotional stimuli is associated with theta band. However, the discrimination failure was linked to alpha band (7–13 Hz)-coordinated unidirectional influence from the amygdala to the hippocampus, indicating that alpha band synchrony may impair emotional discrimination (Zheng, et al. 2019). Furthermore, changes in alpha power have long been associated with internally focused cognitive processes such as imagery, wakefulness, and anticipation (Bonnefond and Jensen 2012; Basar and Guntekin 2012). These classic observations led to some new insights indicating that dynamic changes in alpha-band neural activity are involved in updating threatening representations (Riels et al. 2022).

Notably, these medium-low frequency (i.e., alpha and theta) oscillations are typically associated with error monitoring and conflict resolution processing, particularly when involving activations from the prefrontal cortex and anterior cingulate (Cavanagh and Frank 2014; Carp and Compton 2009; Driel et al. 2012). These oscillations might offer a potential neural index for the prediction error of post-retrieval processing. Indeed, a considerable number of researchers have proposed the critical role of prediction error or expectation violation in post-retrieval updating (Sinclair and Barense 2019; Exton-McGuinness et al. 2015; Sevenster et al. 2014). The difficulty however lies in directly manipulating or measuring the amplitude of the underlying prediction error or violation degree at the behavioural level, creating a challenge in understanding the relationship between memory updating and prediction error (Schroyens et al. 2021). The amplitude of neural oscillations during task offers a potential indicator. While previous EEG studies have provided indicators of successful extinction, such as P300 or LPP amplitude reduction (Kummar et al. 2019; Danon-Kraun et al. 2021), they have not addressed the prediction error or expectation violation aspects during extinction. Considering a wealth of evidence suggests the multiplexing of theta and alpha rhythms and their unique association with fear learning, a complete time-frequency analysis of the entire extinction trial, focusing on key timing points, could be informative.

The current study aims to explore the effect of two postreminder extinction trainings (CS-R & US-R extinction) on preventing fear recovery compared to typical extinction training without a reminder. Using a three-day categorical fear conditioning design, we aim to create a more general scenario, allowing for measurement of changes in episodic memory related to conditioned fear (Kroes et al. 2017; Dunsmoor et al. 2015). Skin conductance response (SCR) will be used as an indicator of the fear response, and EEG will be recorded during the reminder day to examine the neural processing of post-reminder extinction procedures.

Given the repeatable crisis faced by the CS reminder paradigm, especially in preventing the return of categorical fear, and the apparent generalization advantage of the US reminder paradigm, our general hypothesis is that USreminder extinction might yield better behavioral outcomes in preventing the relapse of categorical fear. In terms of neural aspects, unique oscillation patterns during post-reminder extinction should be observed in the US reminder extinction condition, specifically associated with potential superior effects in preventing fear recovery.

# Methods

## **Participants**

A total of 127 Asian participants (70 females; mean age: 21.71, SD = 2.81) were enrolled in the study. All participants affirmed the absence of mental illness or chronic diseases. Six individuals withdrew prior to the acquisition phase due to an inability to tolerate the lowest electrical stimulation level. An additional 33 participants, exhibiting no observable changes in skin conductance responses (SCRs) and reporting no apparent fear of conditioned stimuli on the first day, were excluded. This led to a final sample size of 88 participants. Six participants voluntarily withdrew from the experiment (did not participate in extinction or testing days), and technical issues affected four participants during data collection. The analysis encompassed 78 participants, distributed among the No reminder-extinction group (No-R, n = 29), CS reminder-extinction group (CS-R, n = 23), and US reminder-extinction group (US-R, n = 26). Given the primary focus on changes in extinction learning with different reminders, no additional exclusion criteria were applied for extinction processing.

To determine the required sample size for the potential post-retrieval extinction effect, a power analysis using G\*Power 3.1 (Faul et al. 2007) was conducted. Based on a previous meta-analysis suggesting a moderate effect size, g=0.40 (Kredlow et al. 2016), corresponding to  $f \sim 0.20$ , the mixed ANOVA design necessitated at least 22 participants per group for 80% power. This design included group (US-R, CS-R, No-R) and CS type (CS+, CS-) as between- and within-subjects factors.

No significant group differences were observed in gender proportion (No-R = 57.3%, CS-R = 50.0%, US-R = 42.5%,  $\chi^2 < 2.0$ , p = 0.288), age (No-R = 21.81 ± 2.97, CS-R = 22.21 ± 2.83, US-R = 22.20 ± 2.51, *F* < 1.0, *p* = 0.469), education level (No-R = 15.72 ± 2.15, CS-R = 15.80 ± 3.33, US-R = 16.26 ± 2.32, *F* < 1.0, *p* = 0.351), or trait anxiety level (No-R = 24.06 ± 14.39, CS-R = 21.07 ± 13.70, US-R = 25.13 ± 12.40, F < 1.0, *p* = 0.464), as assessed by the Chinese version of Anxiety Sensitivity index-3 (Cai et al. 2018).

The study protocol (IPCAS2020004) received ethical approval from the Institute of Psychology at the Chinese Academy of Sciences, adhering to approved guidelines. All participants provided written informed consent and received compensation of 200 RMB for completing the entire experiment or 30 RMB if the experiment terminated on day 1.

#### **Experimental Procedure**

The study spanned three consecutive days: Day 1 involved fear acquisition training, Day 2 featured different extinction procedures based on group assignment (CS reminder plus extinction training [CS-R], mild US reminder plus extinction training [US-R], classical extinction training with no prior reminder [No-R]), and Day 3 included spontaneous recovery, reinstatement, and episodic memory tests. During each session, participants were seated behind a table with a computer monitor positioned at a distance of 50 cm in a sound-attenuated room. Skin conductance response (SCR) electrodes (see description below) were attached for all fear conditioning phases, and electroencephalogram (EEG) signals were recorded only on Day 2. Figure 1 outlines the general experimental procedure.

#### **Day 1: Fear Acquisition Training**

On the first day, participants completed demographic information, informed consent, and the anxiety sensitivity index. The primary task involved a partial reinforcement category discrimination fear conditioning task (Kroes et al. 2017; Dunsmoor et al. 2015). The conditioned stimulus (CS) comprised 32 trial-unique images of fish (16) and birds (16), with one category designated as CS + (e.g., birds) and the other as CS - (e.g., fish). The assignment of birds or fish as CS + was counterbalanced across participants. Each CSimage was displayed for 5 s, followed by an inter-trial interval (ITI) of 13 s, 14 s, or 15 s during which a fixation cross was shown. Before fear acquisition training, participants were informed that one type of image (bird or fish) had a



**Fig. 1** Experimental Procedure Overview. The study spanned three consecutive days, with a  $24\pm 2$  h interval between each session. Day 1 encompassed fear acquisition training through a category discrimination fear conditioning task with partial reinforcement. On Day 2, participants underwent group-specific manipulations involving reminder cues: 1) a prototype CS+reminder followed by extinction training (CS-R), 2) a mild US reminder followed by extinction

training (US-R), and 3) classical extinction training with no preceding reminder (No-R). Notably, extinction training did not incorporate the unconditioned stimulus (US). Day 3 served as the testing day for spontaneous recovery, reinstatement, and episodic memory. Skin conductance response (SCR) and shock electrodes were applied throughout all fear conditioning phases, while EEG signals were exclusively recorded on Day 2

probability of being paired with a shock, emphasizing the need to pay careful attention to each image.

The unconditioned stimulus (US) was a 200 ms transcutaneous electrical stimulation (Beijing SANXIA Science and Technology Co. Ltd) delivered to the right inner wrist through two Ag/AgCl electrodes co-terminating with the CS + (initiated 4.8 s after CS + presentation, delayed conditioning). The US intensity was set to a level that was "maximally uncomfortable but not painful" using a gradually increasing procedure (average of 4 mA). The reinforcement ratio was 38%, allowing 6 CS + trials to be paired with the US. The order of CS types was pseudo-randomized, with no more than two trials of the same type occurring in a row.

#### **Day 2: Reminder and Extinction Training**

On Day 2, different reminder manipulations were implemented among the three groups, along with identical extinction learning. The CS-R group received a reminder of the CS + category through the single presentation of a novel item, the most prototypical exemplar of the CS + category, for 6 s. The specific image for the bird CS + exemplar was a normal Paridae, and for the fish CS +, it was a Grass carp. These are common birds and fish in China with no strong symbolic significance or emotional valence. The US-R group received a reminder using an isolated mild electric shock with an intensity set to half of the fear acquisition training (around 2 mA) and a duration of 500 ms. Both reminder groups then listened to a soothing Piano concerto (Für Elise, composed by Ludwig van Beethoven from Richard Clayderman - Digital Concerto, CD 1990). Participants were required to keep their eyes closed for 8 min before the start of extinction training. The choice of listening to soothing music instead of watching a documentary was based on the particularities of the current research and some pilot investigations. This decision aimed to: 1) reduce accidental memory reactivation caused by visual input from a documentary, as neutral natural landscapes may be reminiscent of previous experimental materials (i.e., birds and fish); 2) minimize head movements during this relatively long interval to obtain better EEG data. Individuals in the No-R group also relaxed and listened to the same music before extinction training. For both reactivation groups, electrical stimulation devices were attached during the reminder and extinction phase, while for the No-R group, electrical stimulation devices were attached during the extinction training phase.

Extinction training comprised 17 CS + trials and 17 CS – trials without any US reinforcement. These trials involved novel unique items of fish and bird, presented with the same timing and intervals as during fear acquisition training. Before extinction training commenced, participants were attached to the experimental apparatus and instructed to carefully observe each animal image. Since there is evidence to suggest that the pattern and speed of the extinction process are related to the internal mechanisms (e.g., learning new safety signals or eliminating old unsafe signals) (Gershman and Hartley 2015), we will divide the entire extinction process into three stages based on time: early, mid, and late phase, and implement this distinction in the statistics.

#### **Day 3: Fear Recovery and Episodic Memory Tests**

Day 3 involved tests for spontaneous recovery, reinstatement, and episodic memory. Initially, three novel CS + and three novel CS- images, with no accompanying US presentation, served as a spontaneous recovery test. Subsequently, without any interruption (but 15 s after the last spontaneous recovery CS presentation), two un-signaled USs with an inter-trial interval (ITI) of 14 s were administered to reinstate the fear response. Next (15 s after the last un-signaled US), participants underwent the reinstatement test on novel and unique items, including 8 CS + and 8 CS - trials. During both spontaneous recovery and reinstatement tests, participants were attached with SCR and shock electrodes. Following this, participants performed a surprise item recognition memory test. All unique items from fear acquisition (16 CS+, 16 CS-) and extinction training (17 CS+, 17 CS-) were presented intermixed with twenty novel items for each category. Participants were required to decide whether the stimulus had been presented on Day 1, Day 2, or if it was new.

#### SCR Acquisition and Analyses

SCR was assessed using two Ag/AgCl electrodes of size  $20 \text{ mm} \times 16 \text{ mm}$  attached to the distal phalanges of the index and middle finger of the left hand. Data were collected using the BIOPAC Smart Canter System (Goleta, CA) and continuously recorded at 2000 samples per second.

SCR data were analyzed using a customized program written in Matlab (Version R2018a, Math-Works) and functions from PsPM (Psychophysiological modeling, http:// pspm.sourceforge.net, Version 4.0.2) for data pre-processing and LedaLab 3.4 (Benedek and Kaernbach 2010) for EDA analysis. Data were down-sampled at 40 Hz and median filtered with the parameter set to 40 to remove impulse artifacts. The Continuous Decomposition Analysis (CDA) within LedaLab was used for further event-related SCR extractions. This method, based on Standard Deconvolution, is fast and robust. We extracted sudomotor nerve (SN) peaks within a response window of 1-4.5 s after stimulus onset (excluding the response to US), using an amplitude threshold of 0.01 µS for response scoring. The output index of "CDA.SCR" from LedaLab represented the average phasic driver within the response window, considered to accurately reflect phasic activity. The raw SCRs were log-transformed  $[\ln (\mu S + 1)]$  to achieve a normal distribution of the data.

#### **EEG Acquisition and Analyses**

EEG was recorded during the reminder extinction session on Day 2. A 64Ag/AgCl electrodes cap (NeuroScan Inc., Herndon, VA, USA) was placed on participants' scalps based on the extended International 10/20 system, maintaining electrode impedances below  $5K\Omega$ . EEG activity was amplified using a NeuroScan Synamp2 Amplifier with AC mode of 0.05–100 Hz band-pass filter and continuously sampled at a rate of 1000 Hz, with recording reference via the left mastoid. Offline EEG data were processed using EEGLAB (Delorme and Makeig 2004). The signals were re-referenced to the average reference. Trials contaminated by eye blinks and movements were corrected using an Independent Component Analysis (ICA) algorithm (Makeig et al. 1996). EEG signals during extinction training were segmented into 10-s epochs relative to CS onset (-1~9 s) to investigate different cognitive processes during each single extinction trial: a) the initial capture of attention at the beginning of CS onset, b) the shock anticipatory phase with CS image presentation, c) the US omission phase, and d) the post conflict processing.

The short Time Fourier transform (STFT) was performed for the segmented signals to capture time-frequency dynamics during each extinction trial using custom MATLAB scripts. A function sub\_stft, provided by Zhang et al. (2020), was used to estimate time-varying complex spectrum for EEG data. Specifically, a time-frequency decomposition (TFD) of the EEG signal was obtained using a windowed Fourier transform (WFT) with a fixed 200-ms Hanning window. The WFT yielded, for each EEG epoch, a complex time-frequency estimate F (t, f) at each point (t, f) of the time-frequency plane, extending from -1000 to 9,000 ms (in steps of 1 ms) in the time domain, and from 1 to 45 Hz (in steps of 1 Hz) in the frequency domain. The resulting spectrogram, squared magnitude (in fact PSD) without phase, represents the signal power as a joint function of time and frequency at each time-frequency point. The spectrograms were baseline-corrected (reference interval: -800 to -200 ms relative to stimulus onset) at each frequency f using the subtraction approach, which avoids the positive bias introduced by the percentage approach. This reference interval was chosen to reduce the adverse influence of spectral estimates biased by windowing post-stimulus activity and padding values (Hu et al. 2014).

To further utilize multi-electrodes of scalp EEG and provide a reference for intracranial sources of neural oscillations, standardized low-resolution brain electromagnetic tomography (sLORETA) was used for distributed EEG source location (inverse solution) of detected time–frequency components. Via LORETA software, EEG subepochs of defined time windows were analyzed using standard spectral analyses for a defined frequency band via Discrete Fourier Transform using a boxcar windowing. This yielded cross spectra with a resolution of 0.5 Hz. Intracerebral current density was computed using the LORETA algorithm across 6239 voxels constrained to cortical gray matter. This found a unique inverse solution for the cortical sources of the current pattern at the scalp level, assuming that neighboring neurons are activated synchronously and produce smooth spatial distributions of intracerebral electrical activity (Pascual-Marqui et al. 1999; Pascual-Marqui et al. 1994). Spatial traceability was obtained through the built-in statistical module of sLORETA with one-sample t-tests and 5000 permutation tests, revealing brain voxels reaching the significant threshold (p < 0.01). Intracerebral electrical activity values of defined regions of interest (ROIs) were extracted for further analyses.

## **Data Analysis**

Mixed ANOVAs with designed factors were conducted for all variables, and post-hoc ANOVA or t-tests were performed for further simple analysis. Greenhouse–Geisser corrections of degrees of freedom were applied when assumptions of sphericity were violated. Effect sizes were reported as partial eta-squared.

#### Results

#### **Skin Conductance Responses**

#### Day 1 fear Acquisition Training

For SCRs during fear acquisition, mixed three-way ANOVA with CS-type (CS +/ CS-) and 16 trials as within-subjects factors and group (CS-R / US-R / No-R) as a between-subjects factor were conducted. Fear acquisition training led to a CS +/CS- differentiation over time as indicated by significant main effects of CS-type ( $F_{(1,75)}$ =133.16, p <0.001,  $\eta_p$ =0.64), trial ( $F_{(9.8,735.0)}$ =9.39, p <0.001,  $\eta_p$ =0.11), and a significant CS-type × trial interaction ( $F_{(9.4,701.9)}$ =9.84, p <0.001,  $\eta_p$ =0.12). No group related effects were found, indicating all three groups exhibited successful and comparable fear learning (cf. Figure 2a).



**Fig. 2** Skin conductance responses (SCRs) for the three experimental days. **a**) Day 1 fear acquisition training (ACQ): the three groups showed equivalent and successful fear acquisition. **b**) Day 2 reminder manipulation and extinction training (EXT): All three groups showed successful extinction training (i.e., CS + > CS- at the beginning, and the CS + /CS- differences disappeared at the end of extinction training). The US-R group showed subtle unique extinction learning pattern, as reflected in significantly lower SCRs during the first extinction trial (in light purple background) and relatively slow-down extinction learning, indexed by higher CS + /CS- differentiation during middle extinction (trials 7–12, light yellow background), while the NO-R and CS-R groups exhibited relatively faster extinction processing. **c**) Day 3 fear recovery (REC) tests including spontane-

ous recovery (the first three trials) and reinstatement test (two unsignaled US applications yellow bolt, followed by 8 re-extinction trials): All three groups showed the significant spontaneous recovery of fear (i.e., CS + > CS - for the first and the second recovery trails); while the US-R group showed relatively highest fear fading rate indexed by a complete lack of CS + /CS- differentiation during the third spontaneous recovery trial (in light grey background) compared with other two groups. In addition, the fear reinstatement was observed in all three groups (i.e., CS + > CS - at the first three trials after unsignaled USs), and re-extinctions were successful (the CS + /CS- difference disappeared at the end). \*\*\*: p < .001, \*\*: p < .01, \*: p < .05, ns: non-significant

#### **Day 2 Extinction Training**

For SCRs during fear extinction training, mixed threeway ANOVA with CS-type (CS+/CS-) and 17 trials as within-subjects factors and group (CS-R / US-R / No-R) as a between-subjects factor were conducted. Extinction training was successful in reducing the CS + /CS- differentiation (cf. Figure 2b) as indicated by a significant CS-type × trial interaction ( $F_{(8,2,617,8)} = 6.01$ , p < 0.001,  $\eta_p = 0.074$ ). Subtle group differences could be found when the extinction process was divided into early (1-6 trials), middle (7-12 trials)and late extinction (13-17 trials). All three groups showed significantly higher SCRs towards the CS+compared to CSduring early extinction, i.e., the significant main effect of CS type in the early extinction ( $F_{(1.75)} = 35.94$ , p < 0.001,  $\eta_p = 0.32$ , CS + /CS- differences for three group, No-R, p < 0.001; CS-R, p = 0.001, US-R, p = 0.010) and these differences vanished during late extinction, i.e., non-significant main effect of CS type in the late extinction ( $F_{(175)} = 1.49$ , p = 0.226,  $\eta_p = 0.019$ , CS +/CS- differences for three group, No-R, p = 0.236; CS-R, p = 0.806, US-R, p = 0.356), indicating overall successful extinction learning as suggested by above mentioned analysis. However, during middle extinction phase, a marginal significant CS type x group interaction effect ( $F_{(1.75)} = 2.63, p = 0.079, \eta_p = 0.066$ ) were shown. Further post hoc t-test revealed that SCRs towards CS + were still slightly higher compared to the CS- in the US-R group  $(t_{(25)} = 2.10, p = 0.046)$ , but not in the No-R  $(t_{(28)} = 1.18, p = 0.046)$ p = 0.247) or CS-R group ( $t_{(22)} = -0.50$ , p = 0.620), indicating relatively persisting arousal and thus slow extinction learning for the US-R group specifically (cf. Figure 2b, the 7-12 trials).

In addition, when focusing the first extinction trial as index of initial fear retrieval (in contrast to extinction learning emerging with subsequent trials), CS type (CS +/ CS-)×group ANOVA revealed a main effect of group *F* (2,75) = 10.24, p < 0.001,  $\eta_p = 0.21$  with lower SCR in the US-R group compared to the No-R group (p < 0.001) and CS-R group (p = 0.001). No significant CS type×group interaction ( $F_{(2,75)} = 1.61$ , p = 0.207,  $\eta_p = 0.041$ ) was revealed indicating a general reduced initial fear specifically for the US-R group on day 2.

#### Day 3 Spontaneous Recovery and Reinstatement test

For SCRs regarding spontaneous recovery mixed three-way ANOVA (2 CS-type×3 trials×3 group) were performed. During spontaneous recovery, CS + /CS- differentiation reoccurred and decreased over time (cf. Figure 2c) as indicated by a significant CS×trial interaction occurred (F (1.8,135.8) = 15.25, p < 0.001,  $\eta_p = 0.17$ ). The decomposed ANOVA of CS type×group for each trial revealed that SCRs towards the CS + were higher compared to the CS- in

all three groups at the first recovery trial, i.e., significant main effect of CS type ( $F_{(1,75)} = 50.59$ , p < 0.001,  $\eta_p = 0.40$ , CS +/CS- differentiation for three groups, No-R, p = 0.001; CS-R, p = 0.001; US-R, p < 0.001) and with the same pattern at the second recovery trial (i.e., main effect of CS type,  $F_{(1,75)} = 17.89$ , p < 0.001,  $\eta_p = 0.19$ ). These results indicated that fear recovery occurred in all groups. At the third recovery trial however, the CS +/CS- differentiation disappeared only in the US-R group ( $t_{(25)} = 0.04$ , p = 0.965), and remained intact in the CS-R ( $t_{(22)} = 2.28$ , p = 0.032) and No-R groups ( $t_{(28)} = 2.45$ , p = 0.021). This result indicated a relatively faster fear fading rate during spontaneous recovery in the US-R group (cf. Figure 2c, the last spontaneous recovery trial).

For the reinstatement test, CS + /CS- differentiation reappear after the US signaling and decreased over time (cf. Figure 2c) as indicated by a significant CS×trial interaction ( $F_{(5,2,298,2)} = 8.37$ , p < 0.001,  $\eta_p = 0.10$ ). Interestingly, a significant interaction of trial  $\times$  group was revealed (F  $(11.0.411.7) = 2.30, p = 0.010, \eta_p = 0.058)$ , indicating different temporal patterns of reinstatement responses in three groups. Given the response of the first couple trials after reinstated, US signaling should be a better indicator of fear retrieval, the first three trials of reinstatement phase were included to conduct the 2 CS type  $\times$  3 trial  $\times$  3 group ANOVA. Significant main effects of CS-type ( $F_{(1.75)} = 48.95, p < 0.001$ ,  $\eta_p = 0.40, \text{CS} + > \text{CS}$ -), trial ( $F_{(1.7,131,1)} = 20.41, p < 0.001$ ,  $\eta_p = 0.21$ , first > second > third trial) and group emerged (F  $(2.75) = 3.32, p = 0.042, \eta_p = 0.08)$ . Post hoc analyses revealed that, SCRs in the US-R group were significantly lower than in the No-R group (p = 0.048), regardless of trial or CS-type (cf. Figure 2c, first three reinstatement trials). This group difference was not observed in the late (4~8 trials) reinstatement,  $(F_{(2.75)}=0.91, p=0.408, \eta_p=0.024)$ . Indeed, during the late reinstatement, the main effect of CS type was also disappeared ( $F_{(1.75)} = 1.57$ , p = 0.214,  $\eta_p = 0.02$ ), indicating a successfully re-extinction from reinstatement for all participants.

#### **Day 2 EEG Oscillation Dynamics During Extinction**

EEG signals were segmented into 10 s epochs relative to CS onset. The short Time Fourier transform with hamming window of 200 ms was performed for the segmented signals to get the time-frequency (1–45 Hz) dynamics during each extinction trial. Time-frequency components for statistics were determined by grand mean graphs with bootstrapping test. Specifically, at each time point (t) or time-frequency point, we extracted a collection of numerical samples from all subjects and compared them with a similar collection from the baseline interval. The null hypothesis posited no difference between the means of the two samples, i.e., no disparity in mean magnitude values between pre-stimulus and

post-stimulus intervals. We calculated the pseudo-t statistic of two populations and estimated its probability distribution through permutation testing (5,000 times). After obtaining the distribution of the pseudo-t statistics from the baseline, we calculated the bootstrap p-values for the null hypothesis. This procedure identified the time-frequency regions in which the brain responses were significantly different relative to the baseline interval (Durka et al. 2004). To account for multiple comparisons across time and frequency, the significance level (expressed as p-value) was corrected using an FDR procedure (Benjamini and Hochberg 2018). This led to the identification of several significant clusters (as depicted in the figure), categorized as early event-related synchronization (ERS) component with low to medium frequency bands during the initial attention capture phase (0-400 ms, 2-10 Hz), followed by an event-related desynchronization (ERD) of the shock anticipatory phase (700-5000 ms, 4–12 Hz), then a low frequency ERS when the CS picture disappeared (5500 ms-6000 ms, 1-8 Hz); as well as the late ERS dominated by alpha for subsequent processing of anticipatory violation (8-14 Hz, 6000-9000 ms). According to corresponding cognitive processes, we named these time-frequency components as early ERS, US awaiting ERD, immediate US omission ERS, and late ERS respectively. Three-way ANOVAs with 3 temporal blocks (early/middle/late extinction)  $\times 2$  CS type (CS+/ (No-R/CS-R/US-R) were conducted for each time-frequency component respectively.

For the early ERS (i.e., initial picture presentation phase), the only significant effect was a CS×group interaction (*F*  $_{(2, 65)}$ =3.22, *p*=0.046,  $\eta_p$ =0.090): the relatively higher response for CS + (compared to CS-) was pronounced in the CS-R group compared to the No-R group (*p*=0.017) and less so compared to the US-R group (*p*=0.056), indicating some early CS cue-specific effect in the CS-R group.

For the US awaiting ERD (i.e., picture conscious presentation phase), there was a main effect of CS ( $F_{(1,65)}=6.31$ , p=0.015,  $\eta_p=0.088$ ), the ERD power (negativity) of CS + was significantly larger compared to the CS-; the main effect of time ( $F_{(1.7, 134.6)}=3.35$ , p=0.038,  $\eta_p=0.049$ ), while, no interaction effect between CS and time was found (p=0.061). Notably, the main effect of group was marginally significant  $F_{(2,65)}=2.24$ , p=0.10,  $\eta_p=0.064$ : the US-R group showed significant larger US awaiting ERD relative to the No-R group (p=0.046), while no differences with regard to the CS-R group emerged.

For the immediate US omission ERS, no significant effects were revealed ( $p_s > 0.05$ ). While for the late ERS, we observed a main effect of group with relatively large effect sizes ( $F_{(2,65)}=3.72$ , p=0.030,  $\eta_p=0.103$ ). Post-hoc comparisons showed that the US-R group showed significant larger US omission ERS relative to the No-R group (p=0.011) and CS-R group (p=0.049). No CS-related main effects or interaction effects were found. The grand average plot of neural oscillations during extinction in the three groups can be seen in Figure 3a.

In order to further utilize multi-electrodes of the scalp EEG and to provide reference for intracranial sources of neural oscillations, standardized low-resolution brain electromagnetic tomography (sLORETA) was used for distributed EEG source location (inverse solution) of detected time-frequency ROI signals (cf. Figure 3b). For the early ERS, the maximum probability of source activation was located at Brodmann area 11(max MNI cords, x = 20, y = 65, z = -15, including clusters in the superior frontal and medial frontal gyrus). For shock anticipation (US awaiting) oscillation, the maximum probability of source activation was located at Brodmann area 9 (max MNI cords, x = -35, y = 15, z = -35, including dorsolateral and medial prefrontal cortex). For the shock relief oscillation (late ERS), the maximum probability of source activation is located at Brodmann area 24 (max MNI cords, x = -5, y = 0, z = 30, mainly including the cingulate gyrus and anterior cingulate).

#### **Regression Analysis Between Day 2 and Day 3**

We conducted regression analysis with fear recovery indicators as dependent variables. These recovery indicators were determined by actual recovery patterns and observed inter-group differences in the current study, including fear fading rate during spontaneous retrieval (indexed by LN ((CS+1-CS-1) / (CS+3-CS-3+e) of spontaneous recovery phase) and the fear responses during the reinstatement test (indexed by combined CS+/CS- responses of initial reinstatement trials). These two indicators are trended to be negatively correlated (r=0.28, p=0.064).

The independent predictors included the general power of each frequency band after the reminder, and the power of different time–frequency components of extinction training. Results showed that power of late expectancy violation ERS can positively predict the fear fading rate ( $\beta = 0.42$ ,  $F_{(1,65)} = 7.76$ , p = 0.020). Individuals with higher late ERS power exhibited a higher fear fading rate. In addition, the late ERS power negatively predicted the fear reinstatement index ( $\beta = -0.24$ ,  $F_{(1,65)} = 4.00$ , p = 0.050): individuals who exhibited higher late ERS after US omission showed lower fear reinstatement response on day 3 (cf. Figure 3c). At the source level, the oscillation in Brodmann area 9 (prefrontal cortex regions) during the shock anticipatory phase positively predicted the fear fading rate ( $\beta = 0.30$ ,  $F_{(1,42)} = 4.08$ , p = 0.05). No other significant effects were revealed.

#### **Day 3 Episodic Memory Tests**

Regarding the episodic memory test, mixed three-way ANOVA was performed with the factors CS-type, memory source (Day 1 pictures / Day 2 pictures / New pictures)  $\times$  3 group for memory accuracy. We observed a significant main effect of memory source ( $F_{(1.4,106.1)}$ =39.82,



**Fig. 3** Neural oscillation correlations of post-reminder extinction. **a**) Grand mean time–frequency graph for the three groups. During the extinction phase, there was an early ERS component, followed by an ERD during the CS image persistence and shock anticipation. When the picture or the expected shock threat was surprisingly cancelled, a ERS occurred. The CS-R group showed a higher early ERS in particular regarding the CS +. The US-R group trended to show a larger

p < 0.001,  $\eta_p = 0.35$ ). The correct judgment for new pictures was the highest ( $p_s < 0.001$ ), and the memory accuracy for Day 1 (fear acquisition training) was significantly higher compared to Day 2 (extinction training) (p = 0.002). Furthermore, a significant interaction of CS-type and memory source ( $F_{(1.9,144.3)} = 61.71$ , p < 0.001,  $\eta_p = 0.45$ ), and an interaction of CS-type, memory source and group emerged ( $F_{(3.8,144.2)} = 3.03$ , p = 0.020,  $\eta_p = 0.075$ ). Post hoc t-tests revealed an interesting pattern: For Day 1 pictures, memory accuracy for CS + was significantly higher compared to CS- (No-R,  $t_{(28)} = 4.30$ , p = 0.001; CS-R,  $t_{(22)} = 6.37$ , p < 0.001; US-R  $t_{(25)} = 4.94$ , p < 0.001; Moreover, this CS + enhancement (calculated by Day 1 CS + ACC

shock anticipation ERD and exhibited a significantly higher expectancy violation ERS in 4-12 Hz bands; **b**) Three time–frequency components were source-located in the medial frontal gyrus, dorsolateral and medial prefrontal cortex and anterior cingulate respectively; **c**) the expectancy violation ERS during extinction training positively predicted the fear fading rate during spontaneous recovery and tends to be negatively correlated with the fear reinstatement response

– Day 1 CS-<sub>ACC</sub>) was modulated by group, and especially significant in the CS-R group relative to the No-R group ( $t_{(50)} = 2.16$ , p = 0.035). For Day 2 pictures, no accuracy difference between CS + and CS- was observed. For new pictures, CS + pictures tend to be more erroneously reported as seen previously (NO-R,  $t_{(28)} = -2.77$ , p = 0.010; CS-R,  $t_{(22)} = 6.58$ , p < 0.001; US-R,  $t_{(25)} = -2.70$ , p = 0.012); Further, this the enhanced false rate for new CS + pictures (CS + distortion) calculated by New CS-<sub>ACC</sub> – New CS + <sub>ACC</sub> was modulated by group, CS-R group showed higher CS + distortion than US-R group ( $t_{(47)} = 2.39$ , p = 0.021) and No-R group ( $t_{(50)} = 3.22$ , p = 0.004); cf. Figure 4.



**Fig.4** Episodic memory performance of category fear conditioning for the three reminder groups. The accuracy of recognizing new items was the highest, and the memory accuracy for the Day 1 was significantly higher than for Day 2 ( $\mathbf{a}$ ,  $\mathbf{b}$ ,  $\mathbf{c}$ ); the memory accuracy regarding

CS+on Day 1 was significantly higher compared to the CS- among all groups (a); CS+pictures tend to be more erroneously reported as seen previously, while this effect was more pronounced in the CS-R group (c)

### Discussion

In the context of the memory reconsolidation framework, employing extinction training following a retrieval cue holds promise for mitigating the risk of fear relapses. Our investigation explored the impact of introducing a reminder (CS reminder or US reminder) before extinction training, revealing limited efficacy in preventing fear recovery, particularly within the categorical fear paradigm. However, the innovative US-reminder paradigm, involving a lower intensity US before extinction training, exhibited a distinctive pattern of extinction learning. This paradigm showcased a relatively gradual yet effective extinction process, marked by the persistence of CS + /CS- differentiation in skin conductance responses (SCRs) during the middle of the extinction phase which was eliminated at the end of the extinction.

At the neural level, the US-reminder group displayed heightened theta-alpha desynchronization during the electric shock anticipatory phase and significantly stronger expectation violation synchronization after US omission during extinction. These neural oscillations were source-located in the dorsolateral prefrontal cortex and anterior cingulate, respectively. The unique extinction learning process of the US-reminder group correlated with a relative advantage in the fear recovery test. Despite the presence of fear recovery effects across all groups, the USreminder group exhibited a higher fear extinction rate (indicated by the earlier disappearance of CS+/CS- differences during spontaneous recovery) and the lowest reinstatement-induced fear responses (evidenced by a diminished fear response during initial reinstatement). Regression analysis highlighted that heightened error/conflict processing during extinction training at the neural level was associated with a faster fading rate and less fear reinstatement, suggesting that the subtle advantage of US-reminder extinction might be linked to amplified conflict processing during extinction learning.

Conversely, the CS-reminder group did not demonstrate significant trends in preventing fear relapse. However, it did manifest CS cue-specific effects, such as higher early neurosynchronization for CS + during extinction processing and notably higher false alarms for new CS + category pictures in episodic tests. These findings suggest that different reminders before extinction training may exert complex effects on extinction processing and later fear recovery patterns, selectively influencing the fear response or episodic components of a memory trace.

The enhancement of extinction learning and the reduction of fear recurrence following extinction training are clinically significant goals (Lee et al. 2017). While post-retrieval extinction within the reconsolidation framework shows promise, our findings reveal that a subtle modification, such as adding a simple reminder before classical extinction, has limited effectiveness in preventing fear recovery. The occurrence of spontaneous

recovery and fear reinstatement across all groups underscores the challenges in mitigating fear memories in practical experimental procedures. This difficulty may be influenced by various non-specifically boundary conditions (Silva and Soares 2018; Zuccolo and Hunziker 2019), including the strength of the targeted fear memory (Robinson and Franklin 2010). The robustness of fear acquisition, influenced by factors such as high stimulus intensity and a low reinforcement rate, may contribute to heightened tension and uncertainty, potentially impacting fear extinction (Kitamura et al. 2020). Additionally, under the categorical fear paradigm, contusions and novelty presentations might intensify fear arousal during the learning period, making extinction and fear fading more challenging due to generalization scenarios (Vervliet and Geens 2014). Our results align with previous findings in the same paradigm, where CS-reminder extinction failed to prevent category fear recovery (Kroes et al. 2017). This suggests a broader challenge in achieving the effect of post-reminder extinction in generalized situations.

In contrast to the CS-reminder group, which demonstrated nearly equal fear recovery as the control, our data provide evidence supporting the use of a US as a novel reminder mode to enhance the rate of fear fading and reduce general fear reinstatement responses in category conditioning. While these effects are subtle, their instructive nature becomes apparent, particularly when fear recovery is presented in a general manner. Importantly, these indicators of superior effects from the US-reminder group were significantly correlated with specific processing features during extinction, reinforcing the legitimacy of the US-reminder-dependent effect and the potential benefits of the US-reminder protocol.

Several mechanistic pathways may account for the advantageous effect of the US-reminder paradigm. One possibility involves enhanced prediction errors, considered crucial for memory trace updating (Exton-McGuinness et al. 2015; Sevenster et al. 2014; Sinclair and Barense 2019). A US reminder before extinction training may lead to a broad prediction error without cue-selectivity, influencing the original fear memory in a general manner (Dunbar and Taylor 2017b; Liu et al. 2014). Neural indicators, such as the greater trend of shock anticipatory desynchronization and expectation violation synchronization after US omission in the US-reminder group, link these processes. These time-frequency components, involving theta/alpha oscillations, are associated with conflict monitoring and cognitive control (Cavanagh and Frank 2014; Driel et al. 2012). The source location of these oscillations in the dorsolateral and medial prefrontal cortex aligns with their critical role in threat prediction (Furlong et al. 2010; Kroes et al. 2019). The implication of the expectation violation phase corresponds well with the key function of the anterior cingulate cortex (ACC) in error detection and online error updating (Botvinick et al. 2004; Brown and Braver 2005).

It's worth noting that certain reinforcement learning models predict that larger prediction errors result in stronger updates in associative strength, potentially leading to accelerated learning processes such as extinction learning. However, our observed findings contradict this notion: we observed that the US-reminder group demonstrated successful but relatively slower extinction learning (Holland and Schiffino 2016). Despite efforts in computational neuroscience to unify the understanding of prediction error's role in learning processes, different learning modalities, like reward and fear learning, operate through distinct mechanisms (Iordanova et al. 2021). Establishing a universal single criterion linking prediction error and learning rate can be challenging. For instance, some integrative models suggest that prediction errors indicating outcomes better than expected may lead to slower learning rates compared to those indicating worse-than-expected outcomes, possibly reflecting a form of risk aversion or conservative bias towards negative outcomes (Gershman 2015). In studies where prediction error is considered solely as an indicator of surprise level for specific outcomes rather than a motivational component (i.e., positive or negative PE), a subtle nonlinear relationship between prediction error magnitude and reconsolidation efficiency emerges: a lack of anticipated violation fails to initiate the updating process, while excessively large anticipatory violations result in independent traces rather than facilitating memory updating and integration (Chen et al. 2020; Pan et al. 2021; Sevenster et al. 2014).

When delving into computational models intricately tied to fear learning and spontaneous recovery phenomena, the slower learning rate observed in the present study is in complete agreement with Gershman's 2015 forecasts derived from the latent cause model (Gershman and Hartley 2015). The research revealed two participant groups: one grouping all fear acquisition and extinction trials into a single state, and another separating acquisition and extinction trials into distinct states. Notably, the single-state group showed no spontaneous recovery, whereas the two-state group demonstrated significant spontaneous recovery. These groups exhibited notably distinct learning dynamics. Despite both groups showing no fear discrepancy during the final extinction block, participants in the singlestate group displayed slower learning rates. This sluggishness, according to the authors, likely stems from both reinforced CSa and unreinforced CSb being attributed to the same state, diminishing effective reinforcement within it. In contrast, the two-state group's separation of CSa and CSb into distinct states facilitated quicker acquisition of a conditioned response to CSa. Moreover, the two-state group extinguished its CSa response more rapidly by hypothesizing the activation of a new state during extinction, potentially enhancing pattern separationa hippocampal-dependent process that distinguishes patterns in CS-outcome observations (Marr 1971)—thereby facilitating the formation of separate fear and extinction memories. In contrast, the single-state group attributed both phases to a single state, it may modify the original fear memory within the amygdala, continually updating it with newly acquired safety information throughout the original trace (Kalisch et al. 2006). The operation of the US reminder may, to some extent, increase the likelihood of this state fusion learning, thereby exhibiting simultaneously slowed extinction rates and relatively less fear recovery. This essentially mirrors the partial reinforcement extinction effect (Gershman and Niv 2012).

Indeed, some theories suggest that the extinction process may also be conceptualized as a form of inhibition or cognitive control (Anderson and Floresco 2022; Dillon and Pizzagalli 2007; Lebois et al. 2019; Marek et al. 2019). The slower yet successful extinction in the US reminder group suggests a demand for more deliberate processing, involving ongoing management of conflict. This error/conflict processing is typically mediated by the medial frontal cortex and is characterized by oscillatory changes in alpha-theta frequencies in the medial prefrontal cortex and anterior cingulate cortex (Brown and Braver 2005; Carp and Compton 2009; Cavanagh and Frank 2014; Driel et al. 2012; Iordanova et al. 2021; Ridderinkhof et al. 2004; Taub et al. 2018), as observed in our results. Previous literature has emphasized the crucial role of medial prefrontal theta oscillations in the extinction process, highlighting their interaction with the amygdala and their association with the retrieval-stopping of fear memories, regulation of physiological responses, and adaptive control for unpleasant events (Mueller et al. 2014; Sperl et al. 2019). While some studies suggest that alpha and theta bands may play different roles in error processing, with theta band more related to cognitive control and alpha band more associated with sustained attention and error monitoring, they both contribute to error processing and serve as an integrated explanation for the phenomenon of post-error/conflict slowing (Driel et al. 2012). Individuals exhibit more careful and slowed response strategies after errors and conflict (Notebaert et al. 2009), which can also serve as an additional explanation for the slower learning observed in the US reminder group. Notably, our data further reveal that the neural oscillation energy during expectation violation/error processing directly correlates with the extent of the fear recovery index. This provides evidence that sufficient conflict processing and cognitive control can be critical to enhancing extinction learning and reducing fear relapses.

An alternative explanation for the advantage of the USreminder paradigm could be the attenuation of the electrical stimulation's intensity as a reminder. This weaker threat stimulus may enhance an individual's perception and belief that the impending threat is less painful and controllable (Dunbar and Taylor 2017b). The coping response to a future fear cue, based on the cognition of "Oh, no big deal," may be independent of subsequent extinction training. The significantly smaller CS response during the first extinction trial in the US-reminder group seems to reflect this conjecture. Fear expression in the USreminder group was already weaker than in the other two groups before complete extinction training. Indeed, this phenomenon is associated with "US devaluation," which involves reducing the aversiveness of the unconditioned stimulus (experiencing a decrease in US intensity), leading to a reduction in conditioned fear (Hosoba et al. 2001). Recent research has shown that employing various US devaluation techniques effectively diminishes fear responses to the CS or promotes safety behaviours (Woelk et al. 2021; Wong et al. 2024), and have shown promise in improving clinical anxiety (Kip et al. 2023). Notably, within the retrieval-extinction paradigm, behavioural evidence suggests that exposure to a reminder of the unconditioned stimulus, with or without subsequent intervention, subtly influences the reduction of fear recovery responses (Meir Drexler et al. 2019). This finding is consistent with real-life experiences, where individuals exposed to less intense reminders of traumatic events exhibit reduced recall of the original trauma (Weems et al. 2014).

Additionally, the effect of the US-reminder paradigm may involve pre-extinction emotional involvement. Studies have shown that pre-extinction stress exposure can enhance extinction learning (Drexler et al. 2017). The US reminder may induce a stress-like response in individuals, albeit to a lesser extent, contributing to devaluation processing toward threat signals and improving overall coping abilities to negative events (Meir Drexler et al. 2019).

Taken together, the advantages and underlying neural correlates of the US-reminder paradigm suggest the importance of considering prediction errors and error processing in updating original emotions and beliefs during extinction training. Introducing the arousal of real fear, at a controllable level, and integrating safety learning related to it might lead to sufficient cognitive or emotional expectation violation and subsequent error resolution processing, thereby enhancing extinction learning. However, it is crucial to note that the advantage of the US-reminder is minor, and any translation attempt needs to be approached cautiously.

The uniqueness of our paradigm, where all CS images are unique, allows us to investigate the effects of different reminders on episodic memory. Notably, the CS reminder tended to enhance individuals' original episodic memory for the CS+, consistent with prior reports (Kroes et al. 2017). However, the CS reminder also resulted in false alarms for new CS+images compared to the other groups. This cue specificity in the CSreminder group, where individuals saw a prototype of the CS+image before extinction training, increased the arousal or vigilance of the CS + category. This specificity was reflected at the neural level, with a unique increase of early expectation violation synchronization for CS + images during extinction. Interestingly, this specificity did not significantly contribute to preventing the return of fear. These findings underscore the potential divergent effects of reminders on different memory modalities, such as fear conditioning vs. episodic memory, emphasizing the need for further exploration.

Several limitations should be acknowledged in our research methods and result interpretation. Firstly, our procedural choices, such as a slightly shorter reminder-extinction interval and different buffer fillings, aimed to enhance EEG data capturing and reduce conceptual implicit memory reactivation. While these adjustments may introduce some confusions in result interpretations, they were made to strike a balance between standard procedures and experimental adjustments for specific situations. Secondly, the subtle relative advantage effect of the US-reminder is noteworthy. Rather than entailing the complete disappearance of spontaneous fear, it demonstrates a comparatively more successful recession during re-extinction compared to the other two groups. Moreover, the neural correlates associated with the US advantage effect are not straightforward; specifically, neural desynchronization and synchronization related to expectations and their violations are predominantly observed in the US reminder group, indicating a generally greater neural response in the US group on extinction day, irrespective of CS type. When combined with the slower yet effective behavioral extinction observed, this phenomenon may suggest a distinctive extinction learning process, potentially involving non-cuespecific, generalized intentional integrational learning or deliberate conflict resolution. Future studies could utilize multiple indicators, such as subjective ratings and physiological indexes, to authenticate the US-reminder-dependent extinction pattern and explore its potential advantages with more comprehensive approaches. Indeed, the retrieval dependence effect identified in our study is not synonymous with reconsolidation. Our study did not aim to prove the existence of the reconsolidation process or verify whether the memory genuinely loses stability in the storage aspect. Alternative explanations for the different extinction pattern and their effects outside the reconsolidation framework can exist. Finally, the selection of specific methods for time-frequency decomposition, such as windowed fast Fourier transform or wavelet analysis, and the choice of specific parameters can have a certain impact on the results (trade-off between time and frequency resolution), particularly when dealing with EEG epochs requiring fine differentiation of extinction processing. Moreover, the selection of frequency bands of interest based on grand average spectra may lack precision, especially in the context of complex intergroup factor designs. In the future, more sophisticated parameter comparison designs and model fitting methods can be employed based on fundamental conditioned fear paradigms to accurately identify neural components associated with fear learning and extinction processes. Nevertheless, our current dataset remains a valuable reference point notwithstanding these considerations.

In conclusion, our study demonstrates that the general effect of the reminder plus extinction procedure on preventing the recovery of categorical fear is limited. However, a mild US presentation before extinction training can subtly facilitate the rate of fading and reduce fear reinstatement, thereby alleviating fear recovery. These US reminder-dependent effects, associated with slowed extinction learning and enhanced error processing indexed by PFC/ACC-theta/alpha oscillations, offer constructive evidence for enhancing the effectiveness of extinction learning and reducing fear relapse. The specific effect and mechanism of post-reminder extinction warrant more detailed investigation and clarification. The US-reminder as a novel adaptive extinction paradigm requires further attention.

Acknowledgements We greatly appreciate the excellent work of the technical support staff at the Institutional Center for Shared Technologies and Facilities of Institute of Psychology, CAS. Thanks for the additional equipment support from Hu's lab IPCAS, data analysis guidance from Hui Zhang, Ruhr University Bochum, and technical support from Zhong Yang, Xiao-lei Xia, & Zhi-ling Qiao.

Author Contributions D.N.P was responsible for the design and execution of the experiment, data analysis and article draft writing; D.H assisted in data collection and assessment, O.T.W and C.J.M reviewed the draft article and gave suggestions for modification, X.B.L was responsible for experimental design, supervision, and revision of the article.

**Funding** This research was supported by National Nature Science Foundation of China [31530031, 31671136], Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences as well as by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) within the SFB 1280 Extinction Learning (grant no. 316803389—SFB1280; project A09 [OTW, CJM]). This research project is supported by Science Foundation of Beijing Language and Culture University (supported by "the Fundamental Research Funds for the Central Universities") (23YJ250002) (22YBB29) for Dong-ni Pan. The funding agencies had no role in study design, collection, analysis and interpretation of data, writing of the manuscript, or in the decision to submit the paper for publication.

**Data Availability** On reasonable request, the corresponding author will provide access to the data utilized in the current work.

## Declarations

**Ethics Approval** All procedures performed in the study involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All experimental protocols were approved by the ethics committee of the Institute of Psychology at the Chinese Academy of Sciences (IPCAS2020004). Informed consent was obtained from all individual participants included in the study. The research team placed a high emphasis on ensuring participant privacy and confidentiality, ensuring that all obtained data was exclusively utilized for research objectives.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.

# References

- Agren T et al (2012) Disruption of reconsolidation erases a fear memory trace in the human amygdala. Science 337:1550–1552. https:// doi.org/10.1126/science.1223006
- Anderson MC, Floresco SB (2022) Prefrontal-hippocampal interactions supporting the extinction of emotional memories: the retrieval stopping model. Neuropsychopharmacol 47:180–195. https://doi.org/10.1038/s41386-021-01131-1

- Basar E, Guntekin B (2012) A short review of alpha activity in cognitive processes and in cognitive impairment. Int J Psychophysiol 86:25–38. https://doi.org/10.1016/j.ijpsycho.2012.07.001
- Benchenane K et al (2010) Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning. Neuron 66:921–936. https://doi.org/10.1016/j.neuron. 2010.05.013
- Benedek M, Kaernbach C (2010) A continuous measure of phasic electrodermal activity. J Neurosci Methods 190:80–91. https://doi.org/ 10.1016/j.jneumeth.2010.04.028
- Benjamini Y, Hochberg Y (2018) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc: Ser B (methodol) 57:289–300. https://doi.org/10.1111/j. 2517-6161.1995.tb02031.x
- Bjorkstrand J et al (2016) Disrupting Reconsolidation Attenuates Long-Term Fear Memory in the Human Amygdala and Facilitates Approach Behavior. Curr Biol 26:2690–2695. https://doi.org/10. 1016/j.cub.2016.08.022
- Bocchio M, Nabavi S, Capogna M (2017) Synaptic Plasticity, Engrams, and Network Oscillations in Amygdala Circuits for Storage and Retrieval of Emotional Memories. Neuron 94:731–743. https:// doi.org/10.1016/j.neuron.2017.03.022
- Bonnefond M, Jensen O (2012) Alpha oscillations serve to protect working memory maintenance against anticipated distracters. Curr Biol 22:1969–1974. https://doi.org/10.1016/j.cub.2012.08.029
- Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn Sci 8:539–546. https://doi.org/10.1016/j.tics.2004.10.003
- Brown JW, Braver TS (2005) Learned predictions of error likelihood in the anterior cingulate cortex. Science 307:1118–1121. https:// doi.org/10.1126/science.1105783
- Cai W et al (2018) Reliability, validation and norms of the Chinese version of Anxiety Sensitivity Index 3 in a sample of military personnel. PLoS ONE 13:e0201778. https://doi.org/10.1371/journ al.pone.0201778
- Carp J, Compton RJ (2009) Alpha power is influenced by performance errors. Psychophysiology 46:336–343. https://doi.org/10.1111/j. 1469-8986.2008.00773.x
- Cavanagh JF, Frank MJ (2014) Frontal theta as a mechanism for cognitive control. Trends Cogn Sci 18:414–421. https://doi.org/10. 1016/j.tics.2014.04.012
- Chalkia A, Van Oudenhove L, Beckers T (2020) Preventing the return of fear in humans using reconsolidation update mechanisms: A verification report of Schiller Schiller et al. (2010). Cortex 129:510–525. https://doi.org/10.1016/j.cortex.2020.03.031
- Chalkia A et al (2020) No persistent attenuation of fear memories in humans: A registered replication of the reactivation-extinction effect. Cortex 129:496–509. https://doi.org/10.1016/j.cortex.2020. 04.017
- Chen W et al (2020) Destabilizing Different Strengths of Fear Memories Requires Different Degrees of Prediction Error During Retrieval. Front Behav Neurosci 14:598924. https://doi.org/10. 3389/fnbeh.2020.598924
- Chen S et al (2021) Theta oscillations synchronize human medial prefrontal cortex and amygdala during fear learning. Sci Adv 7. https://doi.org/10.1126/sciadv.abf4198
- Clouter A, Shapiro KL, Hanslmayr S (2017) Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. Curr Biol 27:3143-3148.e3146. https://doi.org/10.1016/j.cub.2017.09.001
- Danon-Kraun S et al (2021) Return of fear following extinction in youth: An event-related potential study. Dev Psychobiol 63:e22189. https://doi.org/10.1002/dev.22189
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21. https://doi.org/10. 1016/j.jneumeth.2003.10.009

- Dillon DG, Pizzagalli DA (2007) Inhibition of Action, Thought, and Emotion: A Selective Neurobiological Review. Appl Prev Psychol 12:99–114. https://doi.org/10.1016/j.appsy.2007.09.004
- Drexler SM et al (2014) Effects of Postretrieval-Extinction Learning on Return of Contextually Controlled Cued Fear. Behav Neurosci 128:474–481. https://doi.org/10.1037/a0036688
- Drexler SM, Hamacher-Dang TC, Wolf OT (2017) Stress before extinction learning enhances and generalizes extinction memory in a predictive learning task. Neurobiol Learn Mem 141:143– 149. https://doi.org/10.1016/j.nlm.2017.04.002
- Driel JV, Ridderinkhof KR, Cohen MX (2012) Not All Errors Are Alike: Theta and Alpha EEG Dynamics Relate to Differences in Error-Processing Dynamics. J Neurosci 32:16795–16806. https://doi.org/10.1523/jneurosci.0802-12.2012
- Dunbar AB, Taylor JR (2017a) Garcinol Blocks the Reconsolidation of Multiple Cocaine-Paired Cues after a Single Cocaine-Reactivation Session. Neuropsychopharmacol 42:1884–1892. https://doi.org/10.1038/npp.2017.27
- Dunbar AB, Taylor JR (2017b) Reconsolidation and psychopathology: Moving towards reconsolidation-based treatments. Neurobiol Learn Mem 142:162–171. https://doi.org/10.1016/j.nlm. 2016.11.005
- Dunsmoor JE, Murty VP, Davachi L, Phelps EA (2015) Emotional learning selectively and retroactively strengthens memories for related events. Nature 520:345. https://doi.org/10.1038/nature14106
- Durka PJ, Zygierewicz J, Klekowicz H, Ginter J, Blinowska KJ (2004) On the statistical significance of event-related EEG desynchronization and synchronization in the time-frequency plane. IEEE Trans Biomed Eng 51:1167–1175. https://doi.org/ 10.1109/TBME.2004.827341
- Exton-McGuinness MTJ, Lee JLC, Reichelt AC (2015) Updating memories-The role of prediction errors in memory reconsolidation. Behav Brain Res 278:375–384. https://doi.org/10.1016/j. bbr.2014.10.011
- Faul F, Erdfelder E, Lang AG, Buchner A (2007) G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 39:175–191. https://doi.org/10.3758/bf03193146
- Fell J, Axmacher N (2011) The role of phase synchronization in memory processes. Nat Rev Neurosci 12:105–118. https://doi. org/10.1038/nrn2979
- Fricchione J et al (2016) Delayed extinction fails to reduce skin conductance reactivity to fear-conditioned stimuli. Psychophysiology 53:1343–1351. https://doi.org/10.1111/psyp.12687
- Furlong TM, Cole S, Hamlin AS, McNally GP (2010) The role of prefrontal cortex in predictive fear learning. Behav Neurosci 124:574–586. https://doi.org/10.1037/a0020739
- Gershman SJ (2015) Do learning rates adapt to the distribution of rewards? Psychon Bull Rev 22:1320–1327. https://doi.org/10. 3758/s13423-014-0790-3
- Gershman SJ, Hartley CA (2015) Individual differences in learning predict the return of fear. Learn Behav 43:243–250. https://doi. org/10.3758/s13420-015-0176-z
- Gershman SJ, Niv Y (2012) Exploring a latent cause theory of classical conditioning. Learn Behav 40:255–268. https://doi.org/10. 3758/s13420-012-0080-8
- Golkar A, Bellander M, Olsson A, Ohman A (2012) Are fear memories erasable?-reconsolidation of learned fear with fear-relevant and fear-irrelevant stimuli. Front Behav Neurosci 6:80. https:// doi.org/10.3389/Fnbeh.2012.00080
- Golkar A, Tjaden C, Kindt M (2017) Vicarious extinction learning during reconsolidation neutralizes fear memory. Behav Res Ther 92:87–93. https://doi.org/10.1016/j.brat.2017.02.004
- Hanslmayr S, Axmacher N, Inman CS (2019) Modulating Human Memory via Entrainment of Brain Oscillations. Trends Neurosci 42:485–499. https://doi.org/10.1016/j.tins.2019.04.004

- Hardwicke TE, Taqi M, Shanks DR (2016) Postretrieval new learning does not reliably induce human memory updating via reconsolidation. Proc Natl Acad Sci 113:5206–5211. https://doi.org/ 10.1073/pnas.1601440113
- Haubrich J, Bernabo M, Baker AG, Nader K (2020) Impairments to Consolidation, Reconsolidation, and Long-Term Memory Maintenance Lead to Memory Erasure. Annu Rev Neurosci 43:297– 314. https://doi.org/10.1146/annurev-neuro-091319-024636
- Holland PC, Schiffino FL (2016) Mini-review: Prediction errors, attention and associative learning. Neurobiol Learn Mem 131:207–215. https://doi.org/10.1016/j.nlm.2016.02.014
- Hosoba T, Iwanaga M, Seiwa H (2001) The effect of UCS inflation and deflation procedures on "fear" conditioning. Behav Res Ther 39:465–475. https://doi.org/10.1016/s0005-7967(00)00025-5
- Hu L, Xiao P, Zhang ZG, Mouraux A, Iannetti GD (2014) Single-trial time-frequency analysis of electrocortical signals: baseline correction and beyond. Neuroimage 84:876–887. https://doi.org/10. 1016/j.neuroimage.2013.09.055
- Iordanova MD, Yau JO, McDannald MA, Corbit LH (2021) Neural substrates of appetitive and aversive prediction error. Neurosci Biobehav Rev 123:337–351. https://doi.org/10.1016/j.neubiorev. 2020.10.029
- Johnson DC, Casey BJ (2015) Extinction during memory reconsolidation blocks recovery of fear in adolescents. Sci Rep-Uk 5(1):8863. https://doi.org/10.1038/srep08863
- Kalisch R et al (2006) Context-Dependent Human Extinction Memory Is Mediated by a Ventromedial Prefrontal and Hippocampal Network. J Neurosci 26:9503–9511
- Kip A, Schoppe L, Arntz A, Morina N (2023) Efficacy of imagery rescripting in treating mental disorders associated with aversive memories - An updated meta-analysis. J Anxiety Disord 99:102772. https://doi.org/10.1016/j.janxdis.2023.102772
- Kitamura H, Johnston P, Johnson L, Strodl E (2020) Boundary conditions of post-retrieval extinction: A direct comparison of low and high partial reinforcement. Neurobiol Learn Mem 174:107285. https://doi.org/10.1016/j.nlm.2020.107285
- Kredlow MA, Unger LD, Otto MW (2016) Harnessing Reconsolidation to Weaken Fear and Appetitive Memories: A Meta-Analysis of Post-Retrieval Extinction Effects. Psychol Bull 142:314–336. https://doi.org/10.1037/bul0000034
- Kroes MCW, Dunsmoor JE, Lin Q, Evans M, Phelps EA (2017) A reminder before extinction strengthens episodic memory via reconsolidation but fails to disrupt generalized threat responses. Sci Rep-Uk 7:10858. https://doi.org/10.1038/s41598-017-10682-7
- Kroes MCW et al (2019) Patients with dorsolateral prefrontal cortex lesions are capable of discriminatory threat learning but appear impaired in cognitive regulation of subjective fear. Soc Cogn Affect Neurosci 14:601–612. https://doi.org/10.1093/scan/nsz039
- Kummar AS, Correia H, Fujiyama H (2019) A Brief Review of the EEG Literature on Mindfulness and Fear Extinction and its Potential Implications for Posttraumatic Stress Symptoms (PTSS). Brain Sci 9. https://doi.org/10.3390/brainsci9100258
- Lebois LAM, Seligowski AV, Wolff JD, Hill SB, Ressler KJ (2019) Augmentation of Extinction and Inhibitory Learning in Anxiety and Trauma-Related Disorders. Annu Rev Clin Psychol 15:257– 284. https://doi.org/10.1146/annurev-clinpsy-050718-095634
- Lee JLC, Nader K, Schiller D (2017) An Update on Memory Reconsolidation Updating. Trends Cogn Sci 21:531–545. https://doi.org/ 10.1016/j.tics.2017.04.006
- Lesting J et al (2011) Patterns of coupled theta activity in amygdala-hippocampal-prefrontal cortical circuits during fear extinction. PLoS ONE 6:e21714. https://doi.org/10.1371/journal.pone.0021714
- Liu JF et al (2014) An Unconditioned Stimulus Retrieval Extinction Procedure to Prevent the Return of Fear Memory. Biol Psychiat 76:895–901. https://doi.org/10.1016/j.biopsych.2014.03.027

- Makeig S, Bell AJ, Jung T.-P, Sejnowski TJ (1996) Independent Component Analysis of Electroencephalographic Data 8
- Marek R, Sun Y, Sah P (2019) Neural circuits for a top-down control of fear and extinction. Psychopharmacology 236:313–320. https:// doi.org/10.1007/s00213-018-5033-2
- Marr D (1971) Simple memory: a theory for archicortex. Philos Trans R Soc Lond B Biol Sci 262:23–81. https://doi.org/10.1098/rstb. 1971.0078
- Meir Drexler S, Merz CJ, Lissek S, Tegenthoff M, WolfOT (2019) Reactivation of the Unconditioned Stimulus Inhibits the Return of Fear Independent of Cortisol. Front Behav Neurosci 13. https:// doi.org/10.3389/fnbeh.2019.00254
- Monfils MH, Cowansage KK, Klann E, LeDoux JE (2009) Extinction-Reconsolidation Boundaries: Key to Persistent Attenuation of Fear Memories. Science 324:951–955. https://doi.org/10.1126/ science.1167975
- Mueller EM, Panitz C, Hermann C, Pizzagalli DA (2014) Prefrontal oscillations during recall of conditioned and extinguished fear in humans. J Neurosci 34:7059–7066. https://doi.org/10.1523/ JNEUROSCI.3427-13.2014
- Nader K, Schafe GE, Le Doux JE (2000) Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature 406:722–726. https://doi.org/10.1038/35021052
- Notebaert W et al (2009) Post-error slowing: an orienting account. Cognition 111:275–279. https://doi.org/10.1016/j.cognition.2009. 02.002
- Pan DN, Hoid D, Wolf OT, Li X (2021) Brain activities of reconsolidation: Nuances in post-retrieval interference led to optimal alterations of episodic memories. Neurobiol Learn Mem 185:107531. https://doi.org/10.1016/j.nlm.2021.107531
- Pascual-Marqui RD, Michel CM, Lehmann D (1994) Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int J Psychophysiol 18:49–65. https://doi. org/10.1016/0167-8760(84)90014-x
- Pascual-Marqui RD et al (1999) Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. Psychiatry Res 90:169–179. https://doi.org/10.1016/s0925-4927(99)00013-x
- Place R, Farovik A, Brockmann M, Eichenbaum H (2016) Bidirectional prefrontal-hippocampal interactions support context-guided memory. Nat Neurosci 19:992–994. https://doi.org/10.1038/nn.4327
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. Science 306:443–447. https://doi.org/10.1126/science.1100301
- Riels K, Ramos Campagnoli R, Thigpen N, Keil A (2022) Oscillatory brain activity links experience to expectancy during associative learning. Psychophysiology 59:e13946. https://doi.org/10.1111/psyp.13946
- Robinson MJF, Franklin KBJ (2010) Reconsolidation of a morphine place preference: Impact of the strength and age of memory on disruption by propranolol and midazolam. Behav Brain Res 213:201–207. https://doi.org/10.1016/j.bbr.2010.04.056
- Schiller D et al (2010) Preventing the return of fear in humans using reconsolidation update mechanisms. Nature 463:49–53. https:// doi.org/10.1038/nature08637
- Schiller D, Kanen JW, LeDoux JE, Monfils MH, Phelps EA (2013) Extinction during reconsolidation of threat memory diminishes prefrontal cortex involvement. P Natl Acad Sci USA 110:20040– 20045. https://doi.org/10.1073/pnas.1320322110
- Schroyens N, Sigwald EL, Van Den Noortgate W, Beckers T, Luyten L (2021) Reactivation-Dependent Amnesia for Contextual Fear Memories: Evidence for Publication Bias. eNeuro 8. https://doi. org/10.1523/ENEURO.0108-20.2020
- Seidenbecher T, Laxmi TR, Stork O, Pape HC (2003) Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. Science 301:846–850. https://doi.org/10. 1126/science.1085818

- Sevenster D, Beckers T, Kindt M (2014) Prediction error demarcates the transition from retrieval, to reconsolidation, to new learning. Learn Mem 21:580–584. https://doi.org/10.1101/lm.035493.114
- Silva MB, Soares AB (2018) Reconsolidation of human motor memory: From boundary conditions to behavioral interventions-How far are we from clinical applications? Behav Brain Res 353:83–90. https://doi.org/10.1016/j.bbr.2018.07.003
- Sinclair AH, Barense MD (2019) Prediction Error and Memory Reactivation: How Incomplete Reminders Drive Reconsolidation. Trends Neurosci 42:727–739. https://doi.org/10.1016/j.tins.2019.08.007
- Soeter M, Kindt M (2011) Disrupting reconsolidation: Pharmacological and behavioral manipulations. Learn Memory 18:357–366. https://doi.org/10.1101/lm.2148511
- Sperl MFJ et al (2019) Fear Extinction Recall Modulates Human Frontomedial Theta and Amygdala Activity. Cereb Cortex 29:701– 715. https://doi.org/10.1093/cercor/bhx353
- Summerfield C, Mangels JA (2005) Coherent theta-band EEG activity predicts item-context binding during encoding. Neuroimage 24:692–703. https://doi.org/10.1016/j.neuroimage.2004.09.012
- Taub AH, Perets R, Kahana E, Paz R (2018) Oscillations Synchronize Amygdala-to-Prefrontal Primate Circuits during Aversive Learning. Neuron 97:291–298293. https://doi.org/10.1016/j.neuron.2017.11.042
- Thompson A, Lipp OV (2017) Extinction during reconsolidation eliminates recovery of fear conditioned to fear-irrelevant and fear-relevant stimuli. Behav Res Ther 92:1–10. https://doi.org/10.1016/j. brat.2017.01.017
- Vervliet B, Geens M (2014) Fear generalization in humans: impact of feature learning on conditioning and extinction. Neurobiol Learn Mem 113:143–148. https://doi.org/10.1016/j.nlm.2013.10.002
- Weems CF et al (2014) Memories of Traumatic Events in Childhood Fade After Experiencing Similar Less Stressful Events: Results From Two Natural Experiments. J Exp Psychol Gen 143:2046– 2055. https://doi.org/10.1037/xge0000016
- Weiss S, Rappelsberger P (2000) Long-range EEG synchronization during word encoding correlates with successful memory performance. Brain Res Cogn Brain Res 9:299–312. https://doi.org/10. 1016/s0926-6410(00)00011-2
- Woelk M, Krans J, Raes F, Vervliet B, Hagenaars MA (2021) Imagery Rescripting Versus Extinction: Distinct and Combined Effects on Expectancy and Revaluation Learning. Clinical Psychological Science 10:622–639
- Wong AHK, Franzen M, Wieser MJ (2024) Unconditioned stimulus devaluation decreases the generalization of costly safety behaviors. J Anxiety Dis 102847:102847. https://doi.org/10.1016/j. janxdis.2024.102847
- Zhang L, Li Z, Zhang F, Gu R, Peng W, Hu L (2020) Demystifying signal processing techniques to extract task-related EEG responses for psychologists. Brain Sci Adv 6(3):171–188. https://doi.org/10. 26599/BSA.2020.9050018
- Zheng J et al (2019) Multiplexing of Theta and Alpha Rhythms in the Amygdala-Hippocampal Circuit Supports Pattern Separation of Emotional Information. Neuron 102:887-898.e885. https://doi. org/10.1016/j.neuron.2019.03.025
- Zuccolo PF, Hunziker MHL (2019) A review of boundary conditions and variables involved in the prevention of return of fear after post-retrieval extinction. Behav Process 162:39–54. https://doi. org/10.1016/j.beproc.2019.01.011

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.