Title:

Dissociable memory modulation mechanisms facilitate fear amnesia at different timescales

Ye Wang[1*], Yinmei Ni[1*], Zijian Zhu[2], Jingchu Hu[3], Daniela Schiller[4], Jian Li[1, 5]

[1]School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

[2]School of Psychology, Shaanxi Normal University, Xi'an, China

[3]Department of Anxiety Disorders, Shenzhen Kangning Hospital, Shenzhen Mental Health Institute, Shenzhen, China.

[4]Departments of Psychiatry and Neuroscience, and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[5]IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

* Equal contribution

Correspondance: leekin@gmail.com

Keywords: fear memory, memory retrieval, reconsolidation, amnesia, cTBS, dlPFC

Abstract

Memory reactivation renders consolidated memory fragile and preludes memory reconsolidation. However, whether memory retrieval facilitates update mechanisms other than memory reconsolidation remains unclear. We tested this hypothesis in three experiments with healthy human participants. First, we demonstrate that memory reactivation prevents the return of fear shortly after extinction training in contrast to the memory reconsolidation effect which takes several hours to emerge and such a short-term amnesia effect is cue independent (Study 1, N = 57 adults). Furthermore, memory reactivation also triggers fear memory reconsolidation and produces cue-specific amnesia at a longer and separable timescale (Study 2, N = 79 adults). Finally, using continuous theta-burst stimulation (Study 3, N = 75 adults), we directly manipulated brain activities in the dorsolateral prefrontal cortex and found that both memory retrieval and intact prefrontal cortex functions were necessary for the short-term fear amnesia. The temporal scale and cue-specificity results of the short-term fear amnesia are clearly dissociable from the amnesia related to memory reconsolidation, and suggest that memory retrieval and extinction training trigger distinct underlying memory update mechanisms. These findings raise the possibility of concerted memory modulation processes related to memory retrieval and shed light to clinical treatment of persistent maladaptive memory.

Introduction

Memory retrieval provides a unique window through which consolidated memories can be modified or even neutralized. Studies of memory reconsolidation suggest that memories are rendered fragile each time they are retrieved [1-3], possibly due to the unbinding of the old memory trace. The reconsolidation of the new memory trace also requires *de novo* protein synthesis, which usually takes hours to complete. Pharmacologically blocking protein synthesis and behavioral interventions can both eliminate the original fear memory expression in the long-term (24 hours later) memory test retrieval[4-7], resulting in a cue-specific fear memory deficit [2, 4, 8]. For example, during the reconsolidation window, old fear memory can be updated via extinction training following fear memory retrieval (the retrieval-extinction paradigm) [5-7].

However, the exact functional role played by the memory retrieval remains unclear. For example, other than the long-term reconsolidation effects induced by pharmacological manipulations in the fear memory tasks [4, 9, 10], the short-term amnesia effect of behavioral manipulation such as retrieval-extinction is often overlooked in the fear memory literature [5-7]. This is in stark contrast with evidence from semantic and episodic memory studies where the behavioral manipulations such as the retrieval-induced forgetting (RIF) protocol or direct suppression would cause short-term memory deficit within 30 minutes [11, 12]. More recently, behavioral strategies originally developed in the declarative memory studies such as direct suppression had also been successfully applied to the fear memory paradigms in humans and prevented the return of fear shortly afterwards (30mins), suggesting that the fear memory might be amenable to a more immediate effect, in addition to what the memory reconsolidation theory prescribes [13].

Parallel to the behavioral manifestation of long- and short-term memory deficits, concurrent neural evidence supporting memory reconsolidation theory emphasizes the long-term effect of memory retrieval by hypothesizing that synapse degradation and *de novo* protein synthesis are required for reconsolidation. These processes typically take several hours, if not longer, and are often aided by overnight sleep [9]. On the contrary, previous behavioral manipulations engendering the short-term declarative memory effect hinges on the instact activities in brain areas such as dorsolateral prefrontal cortex (cognitive control) and the its functional coupling with specific brain regions such as hippocampus (memory retrieval). The declarative amnesia effect emerges much earlier due to the online functional activity modulation, instead of the much slower protein synthesis in these brain areas [13]. However, it remains unclear whether memory retrieval might also precipitate a short-term amnesia effect for the fear memory, in addition to the long-term prevention orchestrated by memory reconsolidation.

To fully comprehend the temporal dynamics of the memory retrieval effect, we first carried out a simple fear memory study by pairing a neutral cue (conditioned stimulus, CS) with the electric shock (unconditioned stimulus, US) to directly test whether the fear retrieval-extinction paradigm would yield a short-term effect on fear expression. We then performed another double cue CS-US fear memory experiment to study the temporal characteristics of fear memory malleability to behavioral extinction at the short- (30mins), medium- (6 hours) and long-term (24 hours) time scales.

We hypothesize that the labile state triggered by the memory retrieval may facilitate different

memory update mechanisms following extinction training, and these mechanisms can be further disentangled through the lens of temporal dynamics and cue-specificities. Specifically, memory reconsolidation effect will only be evident in the long-term (24h) memory test due to its requirement of new protein synthesis and is cue-dependent [14]. In contrast, if more immediate memory update mechanisms are involved, then the fear memory deficit should be observed relatively early (30mins) after extinction training [15, 16]. There may also be a "limbo" state: when the immediate memory update effects dissipate but reconsolidation updating is not yet fully in effect (6h), the conditioned stimuli should still elicit a fear response, regardless of the reminder (cue). Previous research in the short-term declarative and fear memory deficits highlighted the role of cognitive control. Subjects with better control abilities to the intrusive thoughts or higher functional connectivities between brain regions such as the dorsolaterlal prefrontal cortex (dlPFC) and hippocampus yielded more severe amnesia in the memory suppression tasks[13, 17, 35]. Correspondingly, more immediate mechanisms might be associated with individual difference in control abilities over intrusive thoughts [17], which can be measured via the Thought-Control Ability Questionnaire (TCAQ) scale and tested against the individual associative fear memory deficits.

Indeed, through a series of experiments, we identified a short-term fear amnesia effect following memory retrieval, in addition to the fear reconsolidation effect that appeared much later. Finally, by applying the continuous theta-burst stimulation (cTBS) protocol over the dlPFC, the brain region critical for cognitive control and goal maintainance, we showed that both memory retrieval and intact dlPFC function were necessary for the more immediate fear amnesia effect.

Open Practices Statement

Neither of the studies reported in this article was preregistered. The data for both studies are publicly accessible at https://osf.io/9agvk/..

Methods

Participants

All participants were students recruited from Peking University. They were right-handed with normal vision and had not participated in electric shock-related experiments before. None of the particpants had current or past history of psychiatric illness, nor any contraindication on TMS[18, 19]. All participants provided informed consent and were remunerated for their participation. This study was approved by the institutional review board of school of psychological and cognitive sciences at Peking University.

We conducted a power analysis (G*Power) to determine the number of participants sufficient to detect a reliable effect [20]. Based on the effect size of reinstatement effect between treatment group and control group on fear memories reported in the previous literatures ($\eta^2 = 0.19, 0.216, 0.24$; N = 40, 55, 30 total sample size, respectively [9, 13, 21], 52 participants for both groups (study 1: reminder group and no-reminder group) were enough to detect a significant interaction effect ($\alpha = 0.05$, $\beta = 0.9$, 2 (groups) × 2 (phases) two-way ANOVA interaction effect). Similarly, a total of 72 participants for three groups (study 2: 30-min group, 6-h group and 24-h group) were required to detect a reliable effect ($\alpha = 0.05$, $\beta = 0.9$, 3 (groups) × 2 (phases) × 2 (CS+) three-way ANOVA interaction effect) [22]. For the rTMS study, based on the effect size of reinstatement effect between treatment group and

control group on fear memories reported in the previous literatures (η^2 = 0.144, 0.19, 0.216, 0.24; N =79, 40, 55, 30 total sample size, respectively [9, 13, 21]), a total of 72 participants (study 3) were needed to detect a significant interaction effect (α = 0.05, β = 0.8, 4 (groups) × 2 (phases) × 2 (CS+) three-way ANOVA interaction effect).

In study 1, we first recruited a total of 75 human subjects (41 females; mean age= 21.3, SD= 2.21). Eight participants discontinued after Day 1 due to their non-measurable spontaneous conditioned stimulus (CS) signal (the skin conductance response, SCR) towards different CS (non-responders, SCR response to any CS was less than 0.02 μS, n = 0 & 8; reminder group and no-reminder group, respectively). Ten participants (7 females; mean age = 20.6, SD = 1.79) who were "non-learners" during fear acquisition (n = 3 & 1 in both groups; respectively) or extinction (n = 3 & 3, respectively) phase were also excluded from further analysis. The criteria for "non-learners" in the fear acquisition include smaller mean SCR of CS+ (CS that was partially paired with the electric shock) than that of the CS- (CS that was never paired with shock) in the latter half trials of acquisition, and a smaller mean SCR difference between CS+ and CS- in the latter than the first half trials of fear acquisition on day 1. Similarly, the criteria for "non-learners" in the fear extinction are larger mean SCRs of CS+ than the CS- responses in both the last trial and the latter half trials of extinction and a larger mean SCR difference between CS+ and CS- in the latter than the first half trials during fear extinction on day 2 (see the exclusion criteria Table S1 in the Supplemental Material). Therefore, we had a total of 57 subjects for data analysis in study 1 (30 for the reminder group, 17 females, mean age = 21.4, SD = 2.37, and 27 for the no-reminder group, 15 females, mean age = 21.62, SD = 2.13). The group × gender interaction effect on participants' age was not statistically significant ($F_{1,49}$ = 0.137, P = 0.713, η^2 = 0.003, two-way ANOVA), and gender ($\chi_1^2 = 0.007$, P = 0.933) and age (t_{55} = -0.386, P = 0.708) were not significantly different between two groups.

In study 2, we first recruited 119 participants (56 females; mean age= 20.85, SD= 2.59). Thirty-seven participants discontinued after Day 1 for they were "non-responders" (n = 19, 5 and 13; 30-min group, 6-h group and 24-h group, respectively). Additionally, two participants (n = 1&1; 30-min group and 24-h group; respectively) failed to show the evidence of fear acquisition and one participant (30-min group) failed to show the evidence of fear extinction on day 2 and were also excluded from further analysis. Finally, we had a total of 79 subjects for data analysis in study 2 (27 for the 30-min group, 15 females, mean age = 21.4, SD = 2.46; 26 for the 6-h group, 12 females, mean age = 20.54, SD = 1.74; and 26 for the 24-h group, 12 females, mean age = 20.96, SD = 2.990). The interaction effect of gender × group on participants' age was not statistically significant ($F_{2,73}$ = 0.213, P = 0.809, η^2 = 0.006, two-way ANOVA), and gender (χ_2^2 = 0.021, P = 0.990) or age ($F_{2,76}$ = 0.661, P = 0.519, η^2 = 0.017) was not significantly different across three groups.

In study 3, we recruited a total of 90 participants (47 females; mean age = 21.01, SD = 2.44). Fifteen participants (8 females; mean age = 21.2, SD = 1.22) were excluded from further analysis since they failed to show SCR response to the any stimuli (CS1+, CS2+ or CS-) (non-responders, n = 3,4,4,4; R rdlPFC group, R vertex group, NR rdlPFC group and NR vertex group, respectively, see Supplemental Table S1)[6, 7, 23-28]. Our final sample included 75 participants for data analysis (19 for the Reminder right dlPFC group, 10 females, mean age = 22.05, SD = 2.44; 18 for the Reminder vertex group, 9 females, mean age = 20.67, SD = 2.68; 18 for the No-reminder right dlPFC group,

10 females, mean age = 20.28, SD = 2.85; and 20 for the No-reminder vertex group, 10 females, mean age = 20.85, SD = 2.43). A two-way ANOVA showed no significant effect of age of TMS ($F_{1,71}$ = 0.459, P = 0.500, η^2 = 0.006), reminder ($F_{1,71}$ = 1.756, P = 0.189, η^2 = 0.024), nor their interaction($F_{1,71}$ = 2.659, P = 0.107, η^2 = 0.036).

Conditioned stimuli and psychophysiological stimulation

For study 1, we employed two squares with different colors (yellow and blue) that served as CS+ and CS-. During fear acquisition, the CS+ was paired with a mild electrical shock (US) on a 37.5% partial reinforcement scheme (6 CS+ paired with electric shock and 10 CS+ with no shock), and the CS- was never paired with a shock (10 CS-). A pseudo-random CS delivery order was generated for the fear acquisition, extinction and test phase for two groups with the rule that no same trial-type (CS+ and CS-) repeated more than twice. Assignment of square colors to the CS (CS+ and CS-) was counterbalanced across participants.

For study 2 & 3, three squares with different colors (yellow, red and blue) were used as CS1+, CS2+ and CS-, respectively. During the fear acquisition phase, both CS1+ and CS2+ were paired with electric shocks (US) on a 37.5% partial reinforcement scheme, and the CS- was not paired with shocks (10 non-reinforced presentations of CS1+, CS2+ and CS- each, intermixed with an extra of 6 CS1+ and 6 CS2+ that co-terminated with the shock). Again, a pseudo-random stimulus order was generated for fear acquisition and extinction phases of three groups with the rule that no same trial-type (CS1+, CS2+ and CS-) repeated more than twice. In the test phase, to exclude the possibility that the difference between CS1+ and CS2+ was simply caused by the presentation sequence of CS1+ and CS2+, half of the participants completed the test phase using a pseudo-random stimuli sequence and the identities of CS1+ and CS2+ reversed in the other half of the participants. For all three groups, the CS colors were counterbalanced across participants.

The US was a mild electrical shock delivered to the right wrist of participants via a DS-5 Isolated Bipolar Constant Current Stimulator (Digitimer, Welwyn Garden City, UK). The US levels were set by the participants themselves, starting from a low level of shock (5v) and gradually increased and settled to a level that they described as 'uncomfortable, but not painful' (with the maximum level of 10v). The duration of all electric shocks was 200ms with 50 pulses per second.

SCR measurements were collected using two Ag-AgCI electrodes attached to the tips of the index and middle fingers of each subject' left hand. All skin conductance data were recorded via the Biopac® MP160 BioNomadix System, and analyzed using the Acknowledgement 5.0 software. The SCR level for each trial was calculated as the amplitude difference (in microsiemens) from peak to trough during the 0.5 to 4.5s window after the CS (CS SCR) or US (US SCR) onset and responses below 0.02μS were encoded as zero. The raw SCR scores were divided by each subject's mean US responses and then square-root-transformed to normalize SCR across individuals and groups (Wang et al., 2021).

Continuous theta burst stimulation (cTBS) was applied with a Magstim Rapid 2 system (Magstim, Whitland, Wales, UK). Before the experiments, we determined the stimulation intensity of the cTBS

protocol that was at 80% of the individual active motor threshold (AMT). Each transcranial magnetic stimulation (TBS) burst consisted of three stimuli pulses at 50HZ, with each train being repeated every 200ms (5Hz). In this cTBS protocol, a 40s train of uninterrupted TBS is given (600 pluses)[24-25]. During the stimulation, the cTBS was applied either over the right dlPFC or the vertex, which was determined by standard F4 and Cz locations based on international electroencephalogram 10-20 system measurements[29].

The control ability over intrusive thoughts was measured by the 25-item Thought-Control Ability Questionnaire (TCAQ) scale [30]. Participants were asked to rate on a five-point Likert-type scale the extent to which they agreed with the statement from 1 (*completely disagree*) to 5 (*completely agree*). At the end of the experiments, all participants completed the TCAQ scale to assess their perceived control abilities over intrusive thoughts in daily life [17]. The internal consistency (Cronbach's alpha) of TCAQ scale was 0.92, and the reliability coefficient was satisfactory (r = 0.88).

Experimental procedure

Study 1, 2 and 3 were carried out for two or three consecutive days with three phases. Before the experiments, all participants gave informed consent. During the experiments, subjects were required to stay relaxed and still, focus on the computer screen and pay attention to the relationship between color stimuli (CS) and the electric shock (US).

Day 1: Acquisition. For study 1, two groups of participants underwent a fear-conditioning paradigm (acquisition) where the CS+ was partially (37.5%) paired with the US and the CS- was never paired with the shock. For study 2 & 3, participants underwent fear conditioning using three color squares as the CS (CS1+, CS2+ and CS-). The CS were presented for 4s each with a 10-12s variable ITI in all the studies (1, 2 and 3).

Day 2: Retrieval-extinction phase. For study 1, the fear memory was reactivated before extinction in the reminder group. During reactivation, the CS+ was presented once without the delivery of the US. Followed by a 10-min break, the reminder group underwent extinction training with 10 CS+ and 11 CS- presentation non-reinforced. In the no-reminder group, the fear memory was not reactivated (no CS+ reminder). To match with the timeline of the reminder group, the no-reminder group subjects still needed to wait for 10-min after the beginning of the experiment on day 2 and then underwent extinction training with 11 CS+ and 11 CS+ non-reinforced presentation. During the 10-min break, all participants were asked to rest and stay still.

For study 2, only one of the CS+ (CS1+) was presented once without the US during reactivation in all three groups. After a 10-min break, all three groups underwent extinction training with 10 CS1+, 11CS2+ and 11CS- non-reinforced presentation.

For study 3, for the reminder groups (reminder-dlPFC and reminder-vertex groups), only one of the CS+ (CS1+) was presented once (without the US delivery) as the retrieval trial. No CS+ reminder was delivered for the no-reminder groups (no-reminder dlPFC and no-reminder vertex groups). Additionally, during the 10-min break (8 mins after the supposed CS+ reminder delivery), cTBS

was applied either over the right dlPFC (dlPFC groups: PFC) or the vertex (vertex groups: VER), resulting in a 2 (reminder vs. no-reminder) x 2 (dlPFC vs. vertex) of 4 different groups (reminder-dlPFC: R-PFC, reminder-vertex: R-VER, no-reminder dlPFC: NR-PFC, no-reminder vertex: NR-VER).

Day 2 or Day 3: Reinstatement and test phases. For study 1, 30 minutes after the extinction training, both groups received four un-signaled electric shocks with 10s to 12s ITI (reinstatement). Eighteen seconds after the reinstatement phase, all subjects were presented with CS+ and CS- 10 times each without electric shocks and their SCRs were recorded (test phase).

For study 2, participants were assigned into three groups: the 30min group, the 6h group and the 24h group. All participants received four un-signaled electric shocks with 10s to 12s ITI (reinstatement) after their corresponding time delays (30min, 6h or 24h) between extinction and reinstatement. Eighteen seconds after the reinstatement phase, all subjects were tested via the presentation of CS1+, CS2+ and CS- 10 times each without electric shocks associated.

For study 3, the reinstatement and testing procedures were the same as in study 2 except that they were conducted one hour after the extinction training.

Results

The short-term effect of the fear memory retrieval-extinction procedure.

To test the memory-retrieval-induced fear memory deficits at different time scales in humans, we designed two experiments to examine whether the fear memory retrieval-extinction procedure would block the return of fear memory.

In the first study, we aimed to test whether there is a short-term amnesia effect of fear memory retrieval following the fear retrieval-extinction paradigm. We measured subjects' SCRs for the fear acquisition, extinction and test phases to assess the associative fear memory and the recovery of fear memory. At each phase, the differential fear response was calculated as the difference between SCRs to the CS+ and CS-.

To assess fear acquisition across groups (Fig. 1b&c), we conducted a mixed two-way ANOVA (reminder vs. no-reminder) x time (early vs. late part of the acquisition) on the differential fear SCR. Our results showed a significant main effect of time (early vs. late; $F_{1,55} = 6.545$, $P = 0.013$, $\eta^2 = 0.106$), suggesting successful fear acquisition in both groups. There was no main effect of group (reminder vs. no-reminder) or the group x time interaction (group: $F_{1,55} = 0.057$, $P = 0.813$, $\eta^2 = 0.001$; interaction: $F_{1,55} = 0.066$, $P = 0.798$, $\eta^2 = 0.001$), indicating similar levels of fear acquisition between two groups. Post-hoc *t*-tests confirmed that the fear responses to the CS+ were significantly higher than that of CS- during the late part of acquisition phase in both groups (reminder group: $t_{29} = 6.642$, $P < 0.001$; no-reminder group: $t_{26} = 8.522$, $P < 0.001$; Fig. 1c). Importantly, the levels of acquisition were equivalent in both groups (early acquisition: $t_{55} = -0.063$, $P = 0.950$; late acquisition: $t_{55} = -0.318$, $P = 0.751$; Fig. 1c).

To ensure equivalent and successful extinction across groups (Fig. 1b & d), we performed similar mixed two-way ANOVA (group x trial) on the differential SCR in the extinction phase. Our results showed significant main effect of trial ($F_{10,550} = 2.825$, $P = 0.002$, $\eta^2 = 0.049$), but no effect of group ($F_{1,55} = 0.063$, $P = 0.802$, $\eta^2 = 0.001$) or their interaction ($F_{10,550} = 1.77$, $P = 0.063$, $\eta^2 = 0.031$). Follow-up *t*-tests further confirmed that in the last trial of extinction there was no difference between fear responses to CS+ and CS- within group (reminder group: $t_{29} = -1.147$, $P = 0.261$; no-reminder group: $t_{26} = 0.261$, $P = 0.796$; Fig. 1d). Importantly, the levels of extinction did not show significant difference beteween two groups ($t_{55} = -1.074$, $P = 0.288$; Fig. 1d).

Fear recovery was assessed using a two-way ANOVA with main effects of group (reminder vs. no-reminder) and time (last trial of extinction vs. first trial of test) and it showed a significant time × group interaction ($F_{1,55} = 4.087$, $P = 0.048$, $\eta^2 = 0.069$). We defined the fear recovery index as the SCR difference between the first test trial and the last extinction trial for a specific CS. Post hoc *t*-tests showed that fear memories were resilient after regular extinction training, as demonstrated by the significant difference between fear recovery indexes of the CS+ and CS- for the no-reminder group ($t_{26} = 7.441$, $P < 0.001$; Fig. 1e), while subjects in the reminder group showed no difference of fear recovery between CS+ and CS- ($t_{29} = 0.797$, $P = 0.432$, Fig. 1e). Finally, there is no statistical difference between the differential fear recovery indexes between CS+ in the reminder and no-reminder groups ($t_{55} = -2.022$, $P = 0.048$; Fig. 1c, also see Supplemental Material for direct test for the test phase). Also, our results remain robust even with the "non-learners" included in the analysis (Fig. S1 in the Supplemental Material).

Together, these results indicate that the retrieval before extinction procedure prevents the return of fear expression in the short-term test, distinct from the well-established fear reconsolidation effect at a much later time [7]. In the current study, fear recovery was tested 30 minutes after extinction training, whereas the effect of memory reconsolidation was generally evident only several hours later and possibly with the help of sleep [9, 10], leaving open the possibility of a different cognitive mechanism for the short-term fear dementia related to the retrieval-extinction procedure. Importantly, such a short-term effect is also retrieval dependent, suggesting the labile state of memory is necessary for the short-term memory update to take effect (Fig. 1e).

Temporal and generalizing characteristics of the retrieval-extinction effect

Given the findings of short-term fear amnesia following the retrieval-extinction paradigm, we set out to map the temporal dynamics of fear amnesia as well as its cue-specificities - whether the amnesia triggered by a specific CS+ reminder (for example CS1+) generalizes to the fear memory associated with the other CS+ (CS2+). We adopted a double-cue fear learning paradigm and examined fear recovery in the short- (30 mins), medium- (6 hours) and long-term (24 hours) fear memory test in three separate groups of participants after the extinction training (Fig. 2a).

Fear acquisition was assessed using a mixed three-way ANOVA with group (30min, 6h and 24h), CS+ (CS1+ vs. CS2+, defined as the mean SCR differences between each CS+ and CS-) and trial numbers (10 trials) as main factors. This analysis showed a significant main effect of trial ($F_{9,684} = 12.782$, $P < 0.001$, $\eta^2 = 0.144$), but no effect of group ($F_{2,76} = 2.121$, $P = 0.127$, $\eta^2 = 0.053$), CS+

($F_{1,76}$ = 0.575, P = 0.45, η^2 = 0.008) or their interactions (all Ps > 0.05; Fig. 2b). Furthermore, to ensure equivalent fear acquisition across three groups, a mixed two-way ANOVA showed no effect of group (30min, 6h and 24h; $F_{2,76}$ = 1.79, P = 0.174, η^2 = 0.045), CS+ (CS1+ vs. CS2+; $F_{1,76}$ = 1.337, P = 0.251, η^2 = 0.017) or their interaction ($F_{2,76}$ = 0.959, P = 0.388, η^2 = 0.025) on the late phase of acquisition (last 5 trials). Post-hoc t-tests confirmed that the SCRs elicited by CS1+ and CS2+ were significantly higher than that of CS- among three groups on the late phase of acquisition (CS1+: t_{26} = 6.806, P < 0.001; CS2+: t_{26} = 6.853, P < 0.001 for the 30min group; CS1+: t_{25} = 9.822, P < 0.001; CS2+: t_{25} = 9.841, P < 0.001 for the 6h group and CS1+: t_{25} = 6.909, P < 0.001; CS2+: t_{25} = 6.956, P < 0.001 for the 24h group), indicating successful and similar fear acquisition across all groups (Fig. 2c).

We then set out to examine whether participants underwent successful extinction training on day 2. The similar mixed three-way ANOVA showed a significant effect of trial ($F_{10,760}$ = 8.219, P < 0.001, η^2 = 0.098), but no effect of group ($F_{2,76}$ = 1.822, P = 0.169, η^2 = 0.046), CS+ ($F_{1,76}$ = 0.91, P = 0.343, η^2 = 0.012) or their interactions (all Ps > 0.1; Fig. 2b). To test whether participants across all groups achieved similar levels of extinction, a mixed two-way ANOVA showed no effect of group (30min, 6h and 24h; $F_{2,76}$ = 0.059, P = 0.943, η^2 = 0.002), CS+ (CS1+ vs. CS2+; $F_{1,76}$ = 0.258, P = 0.613, η^2 = 0.003) or their interaction ($F_{2,76}$ = 0.504, P = 0.606, η^2 = 0.013) on the last trial of extinction training. Post-hoc t-tests showed that there was no difference between CS+ and CS- responses in all groups on the last trial of extinction (CS1+: t_{26} = - 0.397, P = 0.695; CS2+: t_{26} = -0.505, P = 0.618 for the 30min group; CS1+: t_{25} = -0.054, P = 0.957; CS2+: t_{25} = 0.140, P = 0.890 for the 6h group and CS1+: t_{25} = 0.435, P = 0.667; CS2+: t_{25} = -0.317, P = 0.754 for the 24h group; Fig. 2d).

To assess the effects of the retrieval-extinction procedure as a function of the time delay between fear extinction and test phases, we conducted a 3-way ANOVA with the within-subject factor CS+ (CS1+ vs. CS2+, defined by the mean SCR differences between each CS+ and CS-), time (last trial of extinction vs. first trial of test) and between-subject factor group (30min, 6h and 24h) and found a significant three-way interaction ($F_{2,76}$ = 6.376, P = 0.003, η^2 = 0.144). To disentangle this interaction, we further examined the effects in each group.

In the 30min group, there was no significant effect of time ($F_{1,26}$ = 0.208, P = 0.652, η^2 = 0.008), CS+ ($F_{1,26}$ = 0.888, P = 0.355, η^2 = 0.033) or their interaction ($F_{1,26}$ = 1.039, P = 0.317, η^2 = 0.038). Post-hoc t-tests showed that the retrieval-extinction training diminished fear responses to both the reminded CS+ (differential fear recovery index between CS1+ and CS-, t_{26} = -0.787, P = 0.438; Fig. 2c) and the non-reminded CS+ (differential fear recovery index between CS2+ and CS-, t_{26} = 0.088, P = 0.93; Fig. 2e), suggesting a cue-independent short-term amnesia effect. This result indicates that the short-term amnesia effect observed in Study 2 is not reminder-cue specific and can generalize to the non-reminded cues. In addition, there was no significant difference between the fear recovery index associated with CS1+ and CS2+ (t_{26} = -1.019, P = 0.317; Fig. 2e, also see the supplemental material for direct test of the test phase SCRs).

In contrast to the short-term effect, there was a significant effect of time ($F_{1,25}$ = 9.654, P = 0.005, η^2 = 0.279), but no effect of CS+ ($F_{1,25}$ = 0.036, P = 0.85, η^2 = 0.001) or their interaction ($F_{1,25}$ = 0.028, P = 0.869, η^2 = 0.001) in the 6h (medium-term) group. Post-hoc t-tests showed that fear

memory recovered for the reminded CS+ (differential fear recovery index between CS1+ and CS-, $t_{25} = 2.382$, $P = 0.05$), as well as the non-reminded CS+ (differential fear recovery index between CS2+ and CS-, $t_{25} = 3.525$, $P = 0.004$) in the medium-term test of the retrieval-extinction procedure effect, suggesting the failure to suppress the return of extinguished fear memory at such a time scale (Fig. 2e). There was also no significant difference measured by the fear recovery index between CS1+ and CS2+ ($t_{25} = 0.166$, $P = 0.869$; Fig. 2e).

Finally, in line with previous research on fear memory reconsolidation, significant SCR difference of the fear reinstatement effect between the reminded CS+ and non-reminded CS+ (time × CS+ interaction: $F_{1,25} = 16.924$, $P < 0.001$, $\eta^2 = 0.404$) was detected in the 24h group via two-way ANOVA. More specifically, the retrieval-extinction procedure only diminished fear responses to the reminded CS+ (differential fear recovery index between CS1+ and CS-, $t_{25} = 0.25$, $P = 0.805$), whereas the fear response to the non-reminded CS+ remained significant (differential fear recovery index between CS2+ and CS-, $t_{25} = 3.269$, $P = 0.009$; Fig. 2e). Post-hoc t-test showed that the fear recovery index between CS1+ and CS2+ was significantly different ($t_{25} = -4.114$, $P = 0.001$; Fig. 2c), aligning well with previous literature that suggested the cue-specificity of the fear reconsolidation effect [5, 14]. Similarly, these results remain robust even with the "non-learners" included in the analysis (Fig. S2 in the Supplemental Material).

Altogether, in line with our hypothesis, the return of fear memory measured by the reinstatement effect showed distinct patterns at different time scales. The retrieval-extinction procedure diminished fear SCRs for both the reminded and non-reminded CS+, engendering cue-independent amnesia in the short-term test (30min). On the other hand, at the long-term (24h) fear memory test, fear responses were only eradicated for the reminded CS+ (but not the non-reminded CS+) with a cue-specific effect, consistent with previous literatures on memory reconsolidation. Interestingly, SCR measures of the fear memory re-appeared in the medium-term (6h) test and there was no difference between fear responses elicited by the reminded and non-reminded CS+, suggesting that the 6-hour time-scale may be a "limbo" time point where both the immediate memory update mechanism and memory reconsolidation are not effective.

Fear amnesia effect as a function of the thought-control ability.

Finally, to assess whether the fear reinstatement effects at different time scale is related to the thought-control ability, we first ran a one-way ANOVA to confirm that subjects among different groups did not differ in their thought-control abilities according to the TCAQ scores ($F_{2,76} = 0.154$, $P = 0.857$, $\eta^2 = 0.004$). We then ran separate linear regressions between the differential fear recovery index (differential fear reocery index between CS+ and CS-) and thought-control abiliites (TCAQ scores) in the three gourps. This analysis yielded significant correlation between TCAQ score and fear recovery index for both CS+ in the 30min group (Fig. 3a, main effect of thought-control ability, $t_{51} = -3.446$, $P = 0.003$, Bonferroni correction), indicating that subjects endowed with higher thought-conrol abilities were less likely to experience the return of fear response. However, no significant correlation was observed in the 6h or the 24h group (Fig. 3b & c; Ps > 0.4). Post-hoc analysis of the 30min group showed significant correlation between the fear recovery index of both CS1+ and CS2+ with subjects' thought-control abilities (Fig. 3a, CS1+: r = -0.462, $P = 0.015$; CS2+: r = -0.413, $P = 0.032$). No significant TCAQ and CS+ interaction effect in the three groups (Ps >

0.4).

It is worth noting that the correlations between TCAQ scores and fear responses in all three groups were not significant in the late phase of acquisition or the last trial of extinction (all *Ps* > 0.10), indicating that the effect of thought-control ability was specifically reflected in the fear recovery index.

In conclusion, thought-control ability only affected fear recovery in the short-term (30min) test for both the reminded and non-reminded CS+; however, the association between thought-control ability and fear recovery index was not identified at longer time scales (6h and 24h). These results are consistent with previous research which suggested that people with better capability to resist intrusive thoughts also performed better in motivated dementia in both declarative and associative memories [13, 17].

dlPFC dependent short-term fear memory amnesia

In study 3 (Fig4. a), we tested that in addition to the memory retrieval cue (reminder), whether the intact right dlPFC activity was necessary for the short-term fear amnesia. To test this hypothesis, we set up a reminder CS+ (CS1+ – CS- vs. CS2+ – CS-) x reminder (reminder vs. no-reminder) x cTBS (rdlPFC vs. vertex) x trial number four-way ANOVA against the SCRs for the acquisition and extinction phases. In the acquisition phase, as expected, there were no significant main effects of reminder ($F_{1,71}$ = 1.291, P = 0.260, η^2 = 0.018), cTBS ($F_{1,71}$< 0.001, P = 0.990, η^2 < 0.001), CS+ ($F_{1,71}$ = 1.927, P = 0.169, η^2 = 0.026) or their interactions (all *Ps* > 0.1; Fig. 4b&c) but a significant main effect of trial number ($F_{9,639}$ = 8.054, P < 0.001, η^2 = 0.102), indicating successful learning of CS+ stimuli. Indeed, post-hoc tests confirmed that the SCRs of CS+ was significantly higher than that of the CS- on the late phase (last 5 trials) of acquisition in all four groups (CS1+: t_{18} = 7.735, P < 0.001; CS2+: t_{18} = 8.546, P < 0.001 for the R-PFC group; CS1+: t_{17} = 8.568, P < 0.001; CS2+: t_{17} = 8.478, P < 0.001 for the R-VER group; CS1+: t_{17} = 6.032, P < 0.001; CS2+: t_{17} = 7.722, P < 0.001 for the NR-PFC group and CS1+: t_{19} = 9.110, P < 0.001; CS2+: t_{19} = 6.538, P < 0.001 for the NR-VER group, Fig. 4c) .

The effects of cTBS over the right dlPFC after the memory reactivation were assessed using the similar mixed-effect four-way ANOVA, which showed significant effects of trial ($F_{10,710}$ = 8.566, P < 0.001, η^2 = 0.108), cTBS ($F_{1,71}$ = 6.269, P = 0.015, η^2 = 0.081) and cTBS × trial interaction ($F_{10,710}$ = 2.292, P = 0.012, η^2 = 0.031), but no effect of reminder ($F_{1,71}$ = 1.894, P = 0.173, η^2 = 0.026), CS+ ($F_{1,71}$ = 0.398, P = 0.530, η^2 = 0.006) or other interactions (all *Ps* > 0.1; Fig. 4b). To test whether participants across all groups achieved similar levels of extinction, a mixed-effect three-way ANOVA was performed on the last trial of extinction training and there was no significant effect of reminder (reminder and no-reminder; $F_{1,71}$ = 3.112, P = 0.082, η^2 = 0.043), cTBS ($F_{1,71}$ = 0.929, P = 0.338, η^2 = 0.013), CS+ (CS1+ vs. CS2+; $F_{1,71}$ = 1.623, P = 0.207, η^2 = 0.022) or their interactions (all *Ps* > 0.1; Fig. 4d). Furthermore, post-hoc *t*-tests showed that there was no difference between CS+ and CS- responses in all groups on the last trial of extinction (CS1+: t_{18} = 1.310, P = 0.207; CS2+: t_{18} = 1.297, P = 0.211 for the R-PFC group; CS1+: t_{17} = 0.839, P = 0.413; CS2+: t_{17} = 1.090, P = 0.291 for the R-VER group; CS1+: t_{17} = 0.380, P = 0.709; CS2+: t_{17} = 0.624, P = 0.541 for the NR-PFC group and CS1+: t_{19} = -1.434, P = 0.168; CS2+: t_{19} = -0.535, P = 0.599 for the NR-

VER group; Fig. 4d).

To assess the effects of the right dlPFC on the memory retrieval related short-term amnesia, we conducted a mixed-effect 3-way ANOVA with the within-subject factors CS+ (CS1+ vs. the CS2+), between-subject factors reminder (reminder vs. no-reminder) and cTBS (dlPFC vs. vertex) on fear recovery index (the difference between the first test trial and the last extinction trial for CS1+ and CS2+), which revealed a significant effect of reminder ($F_{1,71} = 8.920$, $P = 0.004$, $\eta^2 = 0.112$), cTBS ($F_{1,71} = 7.180$, $P = 0.009$, $\eta^2 = 0.092$) and reminder × cTBS interaction ($F_{1,71} = 10.613$, $P = 0.002$, $\eta^2 = 0.130$), but no effect of CS+ ($F_{1,71} = 0.518$, $P = 0.474$, $\eta^2 = 0.007$) or other interactions (all $Ps > 0.1$; Fig. 4c). Post-hoc *t*-tests showed that there were significant fear reinstatement responses in the R-PFC group (CS1+: $t_{18} = 5.182$, $P < 0.001$; CS2+: $t_{18} = 5.258$, $P < 0.001$), the NR-PFC group (CS1+: $t_{17} = 3.496$, $P = 0.024$; CS2+: $t_{17} = 4.628$, $P < 0.001$) and the NR-VER group (CS1+: $t_{19} = 4.885$, $P < 0.001$; CS2+: $t_{19} = 3.262$, $P = 0.016$), but not in the R-VER group (CS1+: $t_{17} = -1.640$, $P = 0.119$; CS2+: $t_{17} = -0.744$, $P = 0.467$) (Fig. 4e), suggesting that both memory retrieval and normal dlPFC function are necessary for the short-term fear amnesia.

Finally, we went on to examine whether the fear reinstatement responses was related to the thought-control ability. A two-way ANOVA of reminder (reminder vs. no-reminder) x cTBS (rdlPFC vs. vertex) confirmed that there were no significant effects of reminder ($F_{1,71} = 0.365$, $P = 0.547$, $\eta^2 = 0.005$), cTBS ($F_{1,71} = 0.606$, $P = 0.439$, $\eta^2 = 0.008$) or their interaction ($F_{1,71} = 1.216$, $P = 0.274$, $\eta^2 = 0.017$) on the TCAQ scores. We then conducted simple linear regressions on different groups (as in study 2) and found a significant correlation between the thought-control ability and fear recovery index in the R-VER group (Fig. 5b, $t_{33} = -3.283$, $P = 0.008$, Bonferroni correction across 4 groups), consistent with the results of the short term (30min) amnesia effect in study 2 (Fig. 3a). Further post-hoc analysis of the individual CS+ showed that the correlations between thought-control ability and fear recovery index were significant for both CS+ (CS1+: $r = -0.557$, $P = 0.016$; CS2+: $r = -0.483$, $P = 0.042$). However, there was no such correlation in the R-PFC group (Fig. 5a, $t_{35} = 0.719$, $P = 0.477$), the NR-PFC group (Fig. 5c, $t_{33} = -1.334$, $P = 0.192$) or the NR-VER group (Fig. 5d, $t_{37} = -0.025$, $P = 0.980$) and the interaction of thought-control ability and CS+ effect was not significant in all 4 groups (Ps > 0.3). Importantly, there was no significant correlation between thought-control ability and SCR in the late phase of acquisition or the last trial of extinction in four groups (all $Ps > 0.10$), suggesting a specific relationship between thought-control ability and short-term fear recovery.

Discussion

Our results provide direct evidence to support the hypothesis that the memory reactivation might engage distinct cognitive mechanisms for fear memory modulation, which can be separated by their temporal dynamics, cue specificity and the dependence of thought-control ability. In line with previous research on the reconsolidation of fear memory [31], our research replicated the classic results that behavioral extinction training during reconsolidation window specifically blocked the fear responses to the reminded CS+ while leaving non-reminded CS+ response intact (cue-specificity) in the long-term test (Fig. 6a&b). Moreover, thought-control ability was not related to the fear memory reconsolidation effect (Fig. 6c). However, findings of the short-term fear amnesia effect suggest that the reconsolidation framework may fall short to accommodate this more immediate

effect (Fig. 6a&b).

Research in both declarative and associative emotional memories suggest that memory retrieval is not a simple act of obtaining a copy of stored information, but provides a critical window where the old memory is eligible to be integrated with new information [14]. Research in the field of declarative memory has shown that the motivated forgetting effect can appear early (10-30 minutes) after the memory suppression practices such as the the influential think/no-think (TNT) paradigm [15], suggesting that memory retrieval may indeed also facilitate short-term memory update [32, 33]. Unlike the slow protein synthesis process involved in memory reconsolidation, the motivated forgetting effect was accompanied by heightened dorsolateral prefrontal cortex (dlPFC) activity, diminished hippocampal activation and altered functional connectivity between these brain regions and hence more temporally adaptive [34, 35].

Previous studies indicate that a suppression mechanism can be characterized by three distinct features: first, the memory suppression effect tends to emerge early, usually 10-30 mins after memory suppression practice and can be transient [36, 37]; second, the memory suppression practice seems to directly act upon the unwanted memory itself [38], rendering it cue-independent such that other cues originally associated with the unwanted memory also fail in memory recall; third, the magnitude of memory suppression effects is believed to be associated with individual difference in control abilities over intrusive thoughts [17]. The short-term fear dementia effects we identified are fundamentally different from the memory reconsolidation effects and fit closely with the characteristics of the memory suppression effect. Inspired by the similarities between our results and suppression-induced declarative memory amnesia [39], we hypothesize that the retrieval-extinction procedure might facilitate a spontaneous memory suppression process and thus yield a short-term amnesia effect. Accordingly, the activated fear memory induced by the retrieval cue would be subjected to an automatic fear memory suppression through the extinction training [16].

In our experiments, subjects were not explicitly instructed to suppress their fear expression, yet the retrieval-extinction training significantly decreased short-term fear expression. These results are consisitent with the short-term dementia induced with the more explicit suppression intervention[11-13, 15, 35]. It is worth noting that the standard memory suppression paradigm involve consciously repelling unwanted memory, which might not be necessary to engage the suppression mechanism. For example, in the well-known retrieval-induced forgetting (RIF) phenomenon, the recall of a stored memory can impair the retention of related long-term memory and this forgetting effect emerges as early as 20 minutes after the retrieval procedure, suggesting memory suppression or inhibition can occur in a more spontaneous and automatic manner [40]. Moreover, subjects with greater history of trauma exhibited more suppression-induced forgetting of both negative and neutral memories than those with little or no trauma [41]. Similarly, people with higher self-reported thought-control capabilities showed more severe cue-independent recall deficit, suggesting that suppression mechanism is associated with individual differences in spontaneous control abilities over intrusive thoughts [17]. It has been suggested that similar automatic mechanisms might be involved in organic retrograde amnesia of the traumatic childhood memories [42, 43].

Taken together, we propose that similar brain mechanisms, such as the dynamic brain activity and

connectivity modulation, involved in memory suppression may be deployed spontaneously and act as a general cognitive mechanism to cope with intrusive memories in daily life. Indeed, recently researchers have proposed that the standard retrieval-extinction paradigm to study fear memory might also engage automatic memory suppression [16]. Consistent with this hypothesis, we showed that the retrieval cue was necessary to produce the short-term amnesia effect (Figs. 1&2) and critically, our results fitted well with the three key characteristics of the active suppression mechanism: temporal dynamics, cue-specificity and the dependence of thought-control ability (Figs. 2&3).

To calibrate the temporal boundaries of the short and long-term effects, we also included the medium-term group where fear expression was tested 6 hours after the retrieval-extinction training. Surprisingly, our results showed that fear memory returned at this time point (Fig. 6a&b), consistent with previous animal and human studies that suggest the 6-hour lower time boundary for memory reconsolidation [4, 8-10, 44]. However, the return of fear expression in the medium-term group also indicates that the short-term fear amnesia is transient, laying out the upper time boundary for the hypothesized fear memory suppression effect. Although mixed results have been reported regarding the durability of suppression effects in the declarative memory studies [45, 46], future research will be needed to investigate whether the short-term effect we observed is specifically related to associative memory or the spontaneous nature of suppression. Nevertheless, our results provide clear evidence that the short-term amnesia effect can be temporally separated from the reconsolidation effect (Fig. 6c). Previous research also showed that separate mnemonic processes can be involved after emotional memory retrieval. For example, by varying the duration length and repetition number of memory cue re-exposure, different levels of memory persistence (reconsolidation, extinction or a "limbo" state in between) can be induced, suggesting a neural titration mechanism for memory de-stabilization and re-stabilization [47, 48]. These works identified important "boundary conditions" of memory retrieval in affecting the retention of the maladaptive emotional memories. In our study, however, we showed that even within a boundary condition previously thought to elicit memory reconsolidation, mnemonic processes other than reconsolidation could also be at work, and these processes jointly shape the persistence of fear memory.

Interestingly, our results also corroborate with a recent study showing that blocking fear reconsolidation using post-retrieval repetitive transcranial magnetic stimulation (rTMS) did not influence the short-term fear memory expression [49]. However, as we demonstrated in studies 2 & 3, memory retrieval and intact rdlPFC function were both necessary for the short-term reinstatement fear amnesia after the extinction training. Therefore, instead of "intact short-term fear expression by the rTMS over the dlPFC", their results could also be interpreted such that the rTMS on the dlPFC demolished the otherwise short-term fear amnesia induced by the memory retrieval cue. Indeed, a series of neuroimaging studies have revealed that the prefrontal neural network including anterior cingulate cortex, anterior ventrolateral prefrontal cortex and the dlPFC is directly engaged in purging the unwanted declarative memory from consciousness[33, 34, 38]. Moreover, previous research has established a causal relationship between dlPFC activity and the memory suppression mechanism underlying retrieval-induced forgetting using transcranial direct current stimulation (tDCS)[50-52]. In our experiment (studies 2&3), when the dlPFC function was intact, subjects' thought-control abilities were positively correlated with fear amnesia in the test phase. Such correlations

disappeared in the no-reminder groups (NR-VER and NR-PFC) and the R-PFC group (Figs. 3&5), establishing a causal link between the dlPFC activity and short-term fear amnesia.

These results help shed light on the dynamics of memory modulation after memory activation and designing novel treatment of psychiatric disorders caused by excessive fear or anxiety [53]. Memory reconsolidation and suppression both have been studied thoroughly in separate fields. Specifically, aversive memories are usually associated with different retrieval cues, and interfering memory reconsolidation only blocks learned fear memory associated with the reactivated memory probe but not the others [4, 8-10, 44]. On the other hand, spontaneous memory suppression is cue-independent and directly suppresses the fear memory trace yet only effective within limited durability [36]. Although future research is clearly needed to understand the brain mechanisms of the short-term effect and its connection to the memory suppression effect, especially in relevance to clinical populations such as posttraumatic stress disorder (PTSD) and phobia patients [54, 55], our results provide a general framework to highlight the potential of memory retrieval to triggering different memory updating mechanisms with unique temporal dynamics.

Transparency: This work was supported by the National Science and Technology Innovation 2030 Major Program (2021ZD0203702), National Natural Science Foundation of China Grants (32071090, 31871140) to J.L.

The authors have declared no conflict of interests.

Data availability

All the data of this study are available for download at https://osf.io/9agvk/.

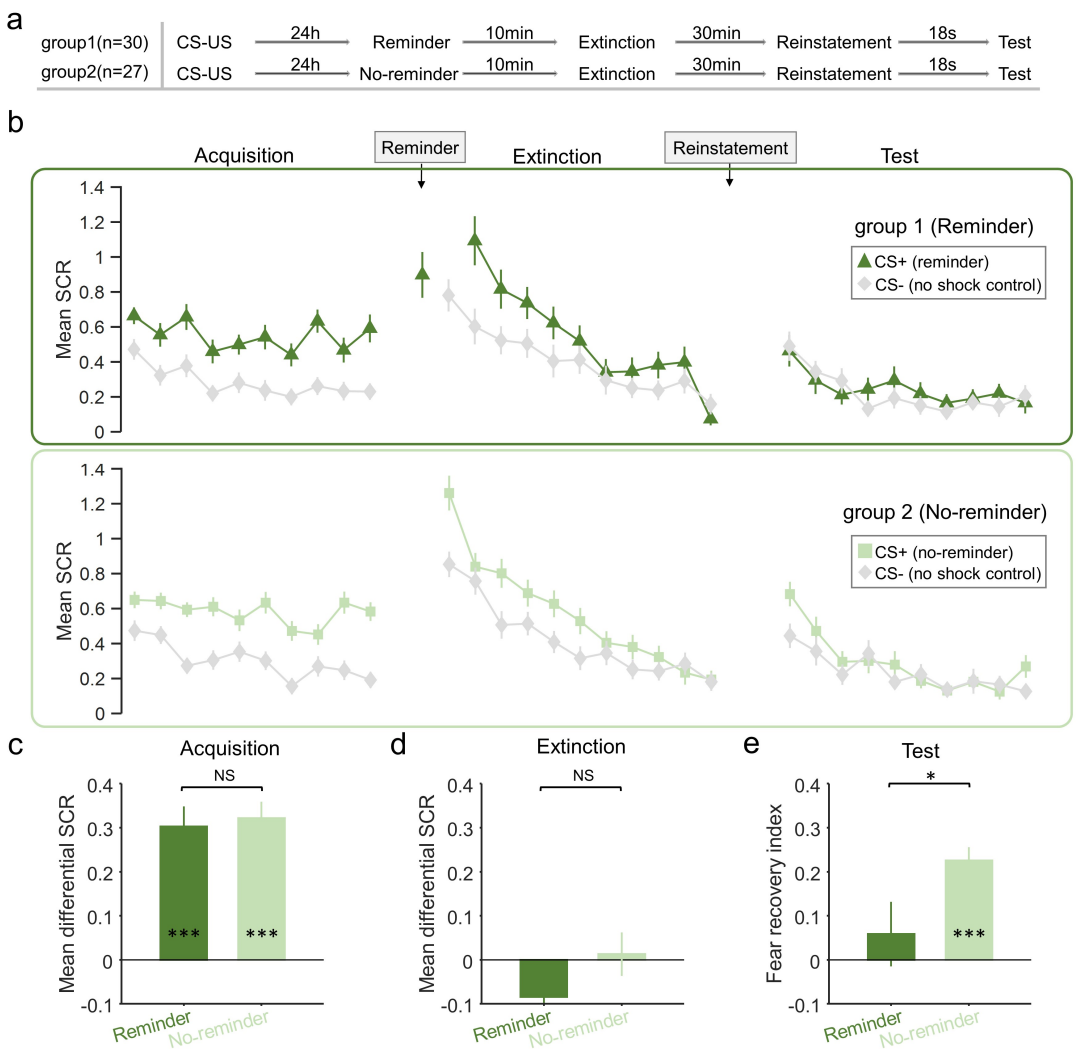

Fig.1 | SCR responses to conditioned stimuli in two groups of study 1. a, Experimental design and timeline. **b**, Mean SCRs of fear conditioned stimuli (CS+) and the control stimulus (CS-) across fear acquisition, extinction and test phases for two groups (reminder and no-reminder). **c**, Mean differential SCRs (CS+ minus CS-) in the acquisition phase (latter half trials). **d**, Mean differential SCRs (CS+ minus CS-) in the extinction phase (last trial). **e**, Differntial fear recovery index between CS+ and CS- in the test phase. ***$P < 0.001$, *$P < 0.05$. NS: Non-significant. Error bars represent standard errors.

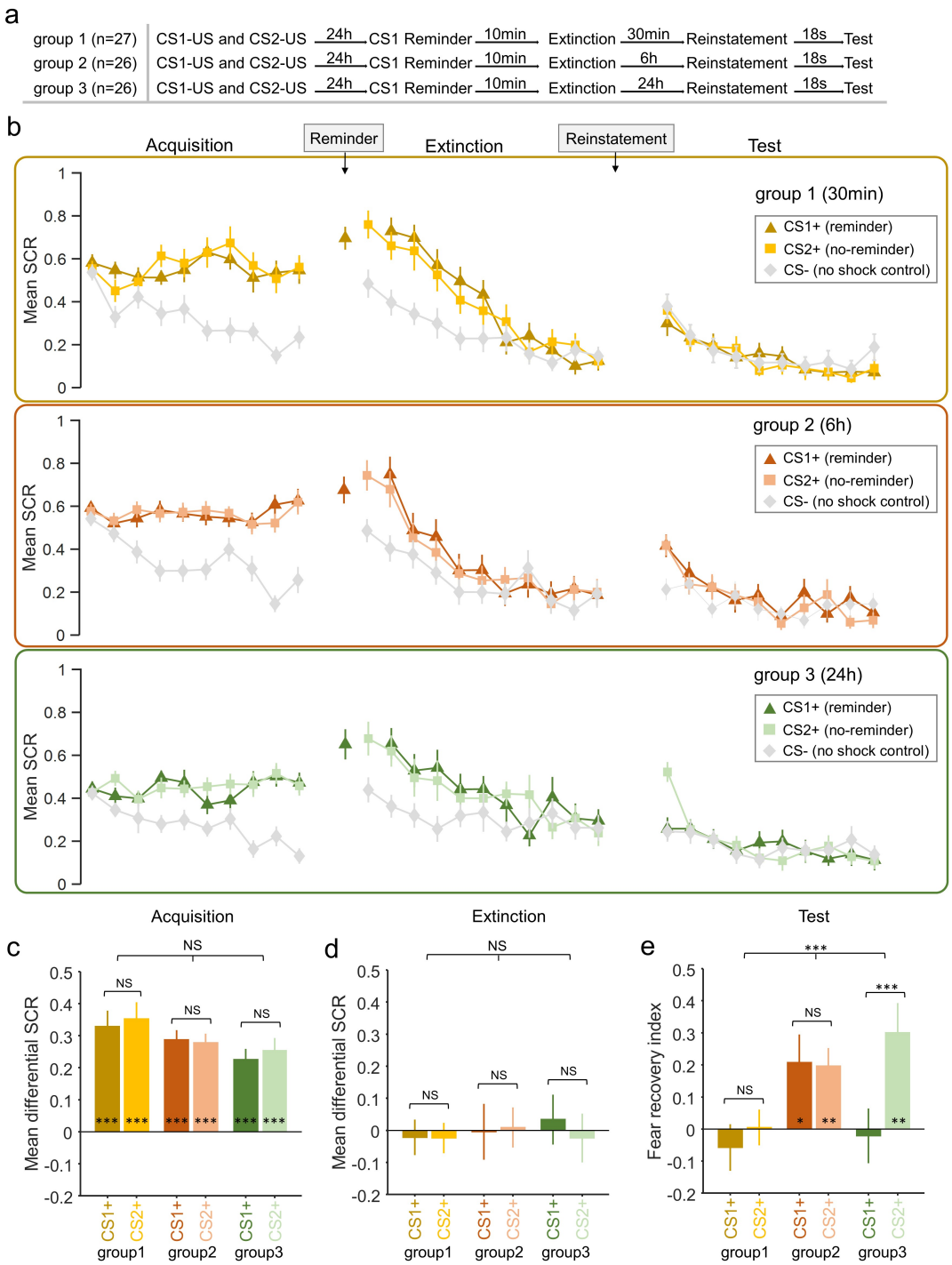

Fig.2 | SCR responses to conditioned stimuli in three groups of study 2. a, Experimental design and timeline. **b**, Mean SCRs of fear conditioned stimuli CS1+ (reminder) and CS2+ (No-reminder), and the control stimulus (CS-) across the fear acquisition, extinction and test phases for each group (30min, 6h and 24h tests). **c,** Mean differential SCRs (CS1+ minus CS- and CS2+ minus CS-) in the acquisition phase (latter half trials). **c**, Mean differential SCRs in the acquisition phase (latter half trials). **d**, Mean differential SCRs in the extinction phase (last trial). **e**, Differntial fear recovery index between CS+ and CS- in the test phase.***$P < 0.001$. *$P < 0.05$. NS: Non-significant. Error bars represent standard errors.

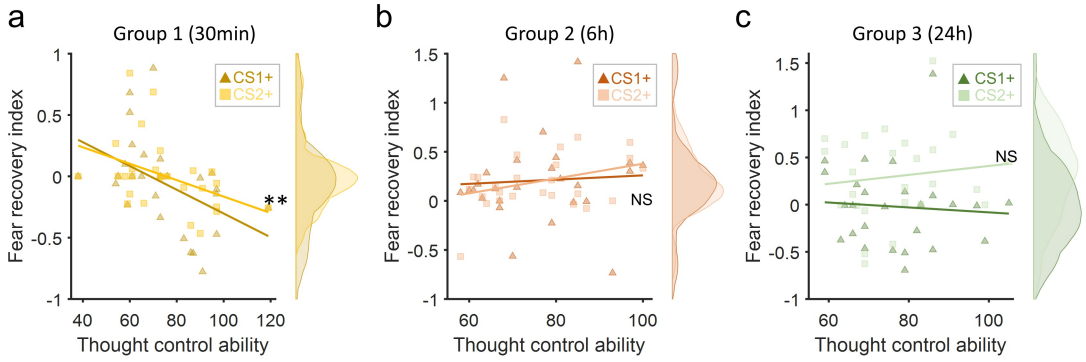

Fig.3 | Fear recovery as a function of thought-control abilities. (**a**)Thought control ability was significantly correlated with fear recovery index in the 30min group (P = 0.003, Bonferroni correction), but not in the 6 hour (**b**) or 24 hour group (**c**, Ps > 0.7). The violin graphs indicate the distribution of fear recovery index across subjects (Fig. 2e). **P < 0.01. NS: Non-significant.

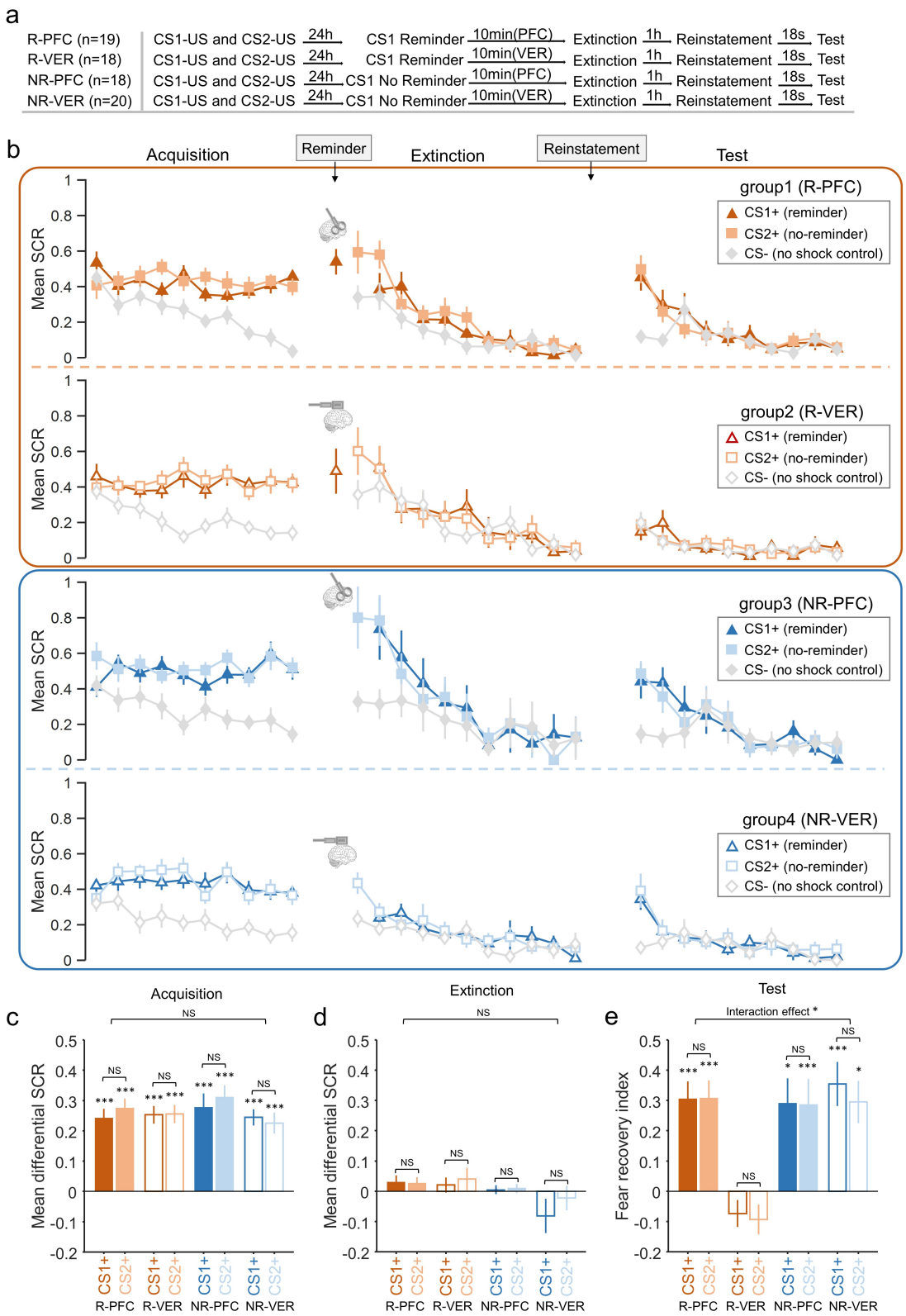

Fig.4 | SCR responses to conditioned stimuli in four cTBS groups. a, Experimental design and timeline. **b**, Mean SCRs of fear conditioned stimuli CS1+ (reminder) and CS2+ (No-reminder), and the control stimulus (CS-) across the fear acquisition, extinction and test phases for each group (R - PFC, R-VER, NR-PFC and NR-VER groups). **c,** Mean differential SCRs (CS1+ minus CS- and CS2+ minus CS-) in the acquisition phase (latter half trials). **d**, Mean differential SCRs in the

extinction phase (last trial). **e**, Differntial fear recovery index between CS+ and CS- in the test phase. ***$P < 0.001$. *$P < 0.05$. NS: Non-significant. Error bars represent standard errors.

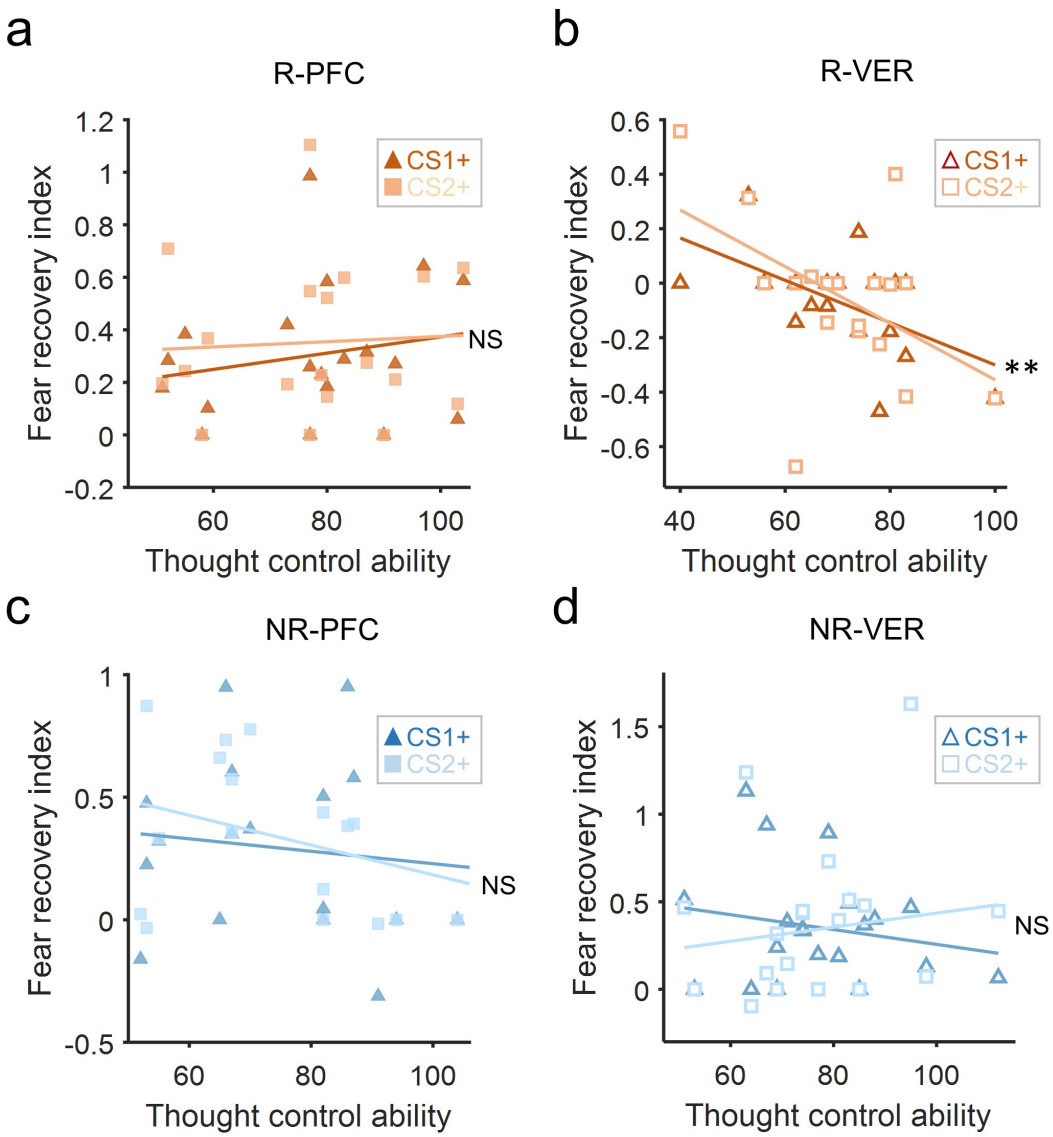

Fig.5 | The correlation between fear recovery index and thought-control abilities in four cTBS groups. In the R-VER group,the correlation between thought control ability scores and fear recovery index was significant ($P = 0.008$, Bonferroni correction), with high thought-control ability participants showing less fear recovery for both CS+; whereas there was no significant correlation between fear recovery index and thought-control abilities in other three groups ($Ps > 0.4$, R-PFC, NR-PFC and NR-VER groups). ** $P < 0.01$, NS: Non-significant.

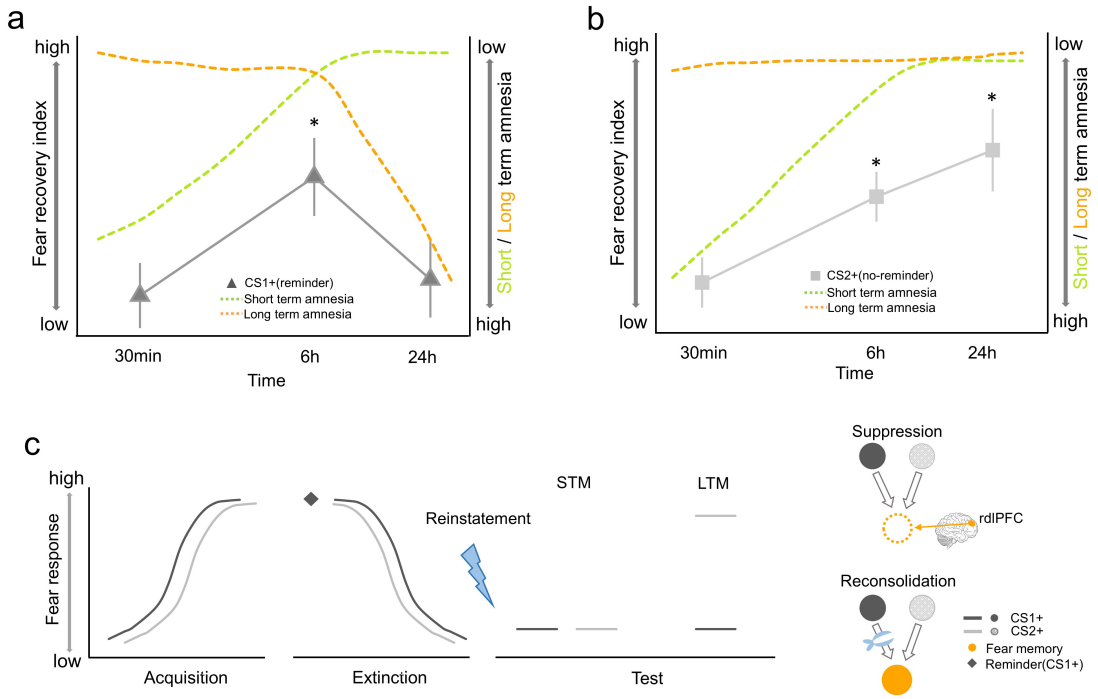

Fig.6 | Time courses of fear memory suppression and reconsolidation. a, At the short interval (30min to 1h), fear recovery of the reminded CS (CS1+) is inhibited (green). As time progresses (from 6h to 24h), amnesia effect is mainly due to the later emerged fear memory reconsolidation effect (orange). The actual SCR data in black solid line. **b**, Fear memory suppression of the non-reminded CS (CS2+) is only evident at the 30-min interval and such effect starts to decay as test interval increases (green). However, fear reconsolidation does not affect CS2+ due to the cue-specificity of the reconsolidation effect (orange). The observed SCR data (black solid line) paralleled the prediction from both the short-term and the reconsolidation effects as the interval length increased (from 30min to 24h). **c,** Schematics demonstrating the effect of cue-reminder on fear memory retention. After both CS1+ (black) and CS2+ (grey) successfully elicit fear responses in the acquisition phase, CS1+ is reminded (black) before both CS+ go through the extinction training. The lack of both CS1+ and CS2+ fear responses in the short-term memory (STM) test (30 min) can be explained by the dlPFC dependent direct suppression effect (dotted circle of US representation). However, the cue-specific fear amnesia effect in the long-term memory (LTM) test (24h) of CS1+ but not CS2+ is attributed to the reconsolidation effect specific to CS1+. * $P < 0.05$, Error bars represent standard errors.

Reference

1. Alberini CM. Mechanisms of memory stabilization: are consolidation and reconsolidation similar or distinct processes? *Trends Neurosci* **28**, 51-56 (2005).

2. Nader K, Schafe GE, LeDoux JE. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* **406**, 722-726 (2000).

3. Dudai Y. Reconsolidation: the advantage of being refocused. *Curr Opin Neurobiol* **16**, 174-178 (2006).

4. Lee JL. Memory reconsolidation mediates the strengthening of memories by additional learning. *Nat Neurosci* **11**, 1264-1266 (2008).

5. Lee JLC, Nader K, Schiller D. An Update on Memory Reconsolidation Updating. *Trends Cogn Sci* **21**, 531-545 (2017).

6. Schiller D, Monfils MH, Raio CM, Johnson DC, Ledoux JE, Phelps EA. Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* **463**, 49-53 (2010).

7. Schiller D, Kanen JW, LeDoux JE, Monfils MH, Phelps EA. Extinction during reconsolidation of threat memory diminishes prefrontal cortex involvement. *Proc Natl Acad Sci U S A* **110**, 20040-20045 (2013).

8. Debiec J, LeDoux JE, Nader K. Cellular and systems reconsolidation in the hippocampus. *Neuron* **36**, 527-538 (2002).

9. Kindt M, Soeter M. Pharmacologically induced amnesia for learned fear is time and sleep dependent. *Nat Commun* **9**, 1316 (2018).

10. Speer ME, Ibrahim S, Schiller D, Delgado MR. Finding positive meaning in memories of negative events adaptively updates memory. *Nat Commun* **12**, 6601 (2021).

11. Anderson MC, Bjork RA, Bjork EL. Remembering can cause forgetting: retrieval dynamics in long-term memory. *J Exp Psychol Learn Mem Cogn* **20**, 1063-1087 (1994).

12. Anderson MC, Bjork EL, Bjork RA. Retrieval-induced forgetting: evidence for a recall-specific mechanism. *Psychon Bull Rev* **7**, 522-530 (2000).

13. Wang Y, Zhu Z, Hu J, Schiller D, Li J. Active suppression prevents the return of threat memory in humans. *Commun Biol* **4**, 609 (2021).

14. Monfils MH, Cowansage KK, Klann E, LeDoux JE. Extinction-reconsolidation boundaries:

key to persistent attenuation of fear memories. *Science* **324**, 951-955 (2009).

15. Anderson MC, Green C. Suppressing unwanted memories by executive control. *Nature* **410**, 366-369 (2001).

16. Anderson MC, Floresco SB. Prefrontal-hippocampal interactions supporting the extinction of emotional memories: the retrieval stopping model. *Neuropsychopharmacology* **47**, 180-195 (2022).

17. Küpper CS, Benoit RG, Dalgleish T, Anderson MC. Direct suppression as a mechanism for controlling unpleasant memories in daily life. *J Exp Psychol Gen* **143**, 1443-1449 (2014).

18. Rossi S, Hallett M, Rossini PM, Pascual-Leone A. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol* **120**, 2008-2039 (2009).

19. Rossini PM*, et al.* Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. *Clin Neurophysiol* **126**, 1071-1107 (2015).

20. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* **41**, 1149-1160 (2009).

21. Kindt M, Soeter M, Vervliet B. Beyond extinction: erasing human fear responses and preventing the return of fear. *Nat Neurosci* **12**, 256-258 (2009).

22. Heo M, Leon AC. Sample sizes required to detect two-way and three-way interactions involving slope differences in mixed-effects linear models. *J Biopharm Stat* **20**, 787-802 (2010).

23. Schiller D, Raio CM, Phelps EA. Extinction training during the reconsolidation window prevents recovery of fear. *J Vis Exp*, e3893 (2012).

24. Dunsmoor JE, Campese VD, Ceceli AO, LeDoux JE, Phelps EA. Novelty-facilitated extinction: providing a novel outcome in place of an expected threat diminishes recovery of defensive responses. *Biol Psychiatry* **78**, 203-209 (2015).

25. Dunsmoor JE, Otto AR, Phelps EA. Stress promotes generalization of older but not recent threat memories. *Proc Natl Acad Sci U S A* **114**, 9218-9223 (2017).

26. Raio CM, Hartley CA, Orederu TA, Li J, Phelps EA. Stress attenuates the flexible updating of aversive value. *Proc Natl Acad Sci U S A* **114**, 11241-11246 (2017).

27. Hu J, Wang W, Homan P, Wang P, Zheng X, Schiller D. Reminder duration determines threat memory modification in humans. *Sci Rep* **8**, 8848 (2018).

28. de Voogd LD, Phelps EA. A cognitively demanding working-memory intervention enhances extinction Research article. *Sci Rep* **10**, 7020 (2020).

29. Censor N, Dimyan MA, Cohen LG. Modification of existing human motor memories is enabled by primary cortical processing during memory reactivation. *Curr Biol* **20**, 1545-1549 (2010).

30. Wells A, Davies MI. The Thought Control Questionnaire: a measure of individual differences in the control of unwanted thoughts. *Behav Res Ther* **32**, 871-878 (1994).

31. Liu J*, et al.* An unconditioned stimulus retrieval extinction procedure to prevent the return of fear memory. *Biol Psychiatry* **76**, 895-901 (2014).

32. Engen HG, Anderson MC. Memory Control: A Fundamental Mechanism of Emotion Regulation. *Trends Cogn Sci* **22**, 982-995 (2018).

33. Benoit RG, Anderson MC. Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron* **76**, 450-460 (2012).

34. Anderson MC*, et al.* Neural systems underlying the suppression of unwanted memories. *Science* **303**, 232-235 (2004).

35. Wimber M, Alink A, Charest I, Kriegeskorte N, Anderson MC. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nat Neurosci* **18**, 582-589 (2015).

36. MacLeod MD, Macrae CN. Gone but not forgotten: the transient nature of retrieval-induced forgetting. *Psychol Sci* **12**, 148-152 (2001).

37. Saunders J, MacLeod MD. New evidence on the suggestibility of memory: the role of retrieval-induced forgetting in misinformation effects. *J Exp Psychol Appl* **8**, 127-142 (2002).

38. Levy BJ, Anderson MC. Inhibitory processes and the control of memory retrieval. *Trends Cogn Sci* **6**, 299-305 (2002).

39. Gagnepain P, Hulbert J, Anderson MC. Parallel Regulation of Memory and Emotion Supports the Suppression of Intrusive Memories. *J Neurosci* **37**, 6423-6441 (2017).

40. Imai H, Kim D, Sasaki Y, Watanabe T. Reward eliminates retrieval-induced forgetting. *Proc*

Natl Acad Sci U S A **111**, 17326-17329 (2014).

41. Hulbert JC, Anderson MC. What doesn't kill you makes you stronger: Psychological trauma and its relationship to enhanced memory control. *J Exp Psychol Gen* **147**, 1931-1949 (2018).

42. Schacter DL, Koutstaal W, Norman KA. Can cognitive neuroscience illuminate the nature of traumatic childhood memories? *Curr Opin Neurobiol* **6**, 207-214 (1996).

43. Schacter DL, Addis DR, Hassabis D, Martin VC, Spreng RN, Szpunar KK. The future of memory: remembering, imagining, and the brain. *Neuron* **76**, 677-694 (2012).

44. Nader K, Schafe GE, Le Doux JE. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* **406**, 722-726 (2000).

45. Storm B, Bjork E, Bjork R. On the durability of retrieval-induced forgetting. *Journal of Cognitive Psychology* **24**, 617-629 (2012).

46. Meier B, König A, Parak S, Henke K. Suppressed, but not forgotten. *Swiss Journal of Psychology / Schweizerische Zeitschrift für Psychologie / Revue Suisse de Psychologie* **70**, 5-11 (2011).

47. Merlo E, Milton AL, Goozée ZY, Theobald DE, Everitt BJ. Reconsolidation and extinction are dissociable and mutually exclusive processes: behavioral and molecular evidence. *J Neurosci* **34**, 2422-2431 (2014).

48. Faliagkas L, Rao-Ruiz P, Kindt M. Emotional memory expression is misleading: delineating transitions between memory processes. *Current Opinion in Behavioral Sciences* **19**, 116-122 (2018).

49. Borgomaneri S, Battaglia S, Garofalo S, Tortora F, Avenanti A, di Pellegrino G. State-Dependent TMS over Prefrontal Cortex Disrupts Fear-Memory Reconsolidation and Prevents the Return of Fear. *Curr Biol* **30**, 3672-3679 e3674 (2020).

50. Penolazzi B, Stramaccia DF, Braga M, Mondini S, Galfano G. Human memory retrieval and inhibitory control in the brain: beyond correlational evidence. *J Neurosci* **34**, 6606-6610 (2014).

51. Stramaccia DF, Penolazzi B, Altoè G, Galfano G. TDCS over the right inferior frontal gyrus disrupts control of interference in memory: A retrieval-induced forgetting study. *Neurobiol Learn Mem* **144**, 114-130 (2017).

52. Valle TM, Bajo MT, Gómez-Ariza CJ. Cathodal transcranial direct current stimulation over the right dorsolateral prefrontal cortex cancels out the cost of selective retrieval on subsequent analogical reasoning. *Neuropsychologia* **141**, 107431 (2020).

53. Chen W*, et al.* Retrieval-extinction as a reconsolidation-based treatment for emotional disorders:Evidence from an extinction retention test shortly after intervention. *Behav Res Ther* **139**, 103831 (2021).

54. Homan P*, et al.* Neural computations of threat in the aftermath of combat trauma. *Nat Neurosci* **22**, 470-476 (2019).

55. Stramaccia DF, Meyer AK, Rischer KM, Fawcett JM, Benoit RG. Memory suppression and its deficiency in psychological disorders: A focused meta-analysis. *J Exp Psychol Gen* **150**, 828-850 (2021).