nature

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41586-018-0405-7

In the format provided by the authors and unedited.

Addendum: Preventing the return of fear in humans using reconsolidation update mechanisms

Daniela Schiller, Marie-H. Monfils, Candace M. Raio, David C. Johnson, Joseph E. LeDoux & Elizabeth A. Phelps

Revised Supporting Online Material*

*revised text is underlined

Materials and Methods

Experiment 1: Long-term blockade of the recovery of fear memories

Participants

One hundred and twenty six subjects were recruited by poster or email advertisements. Sixty-one participants were excluded at different stages of data collection because they did not acquire fear conditioning or they did not extinguish fear responses, as assessed by SCR. We used these exclusion criteria because we could not assess fear recovery when fear responses were not successfully acquired and/or extinguished. Forty participants were excluded after Day 1 of testing for being a non-responder (i.e., failing to show evidence of an SCR response on over 75% of the trials) or failing to show evidence of acquisition. An additional 2 participants were excluded after Day 2 for failing to show evidence of extinction. Thirteen participants who completed all study procedures, were excluded prior to the investigators completing data collection for the study for failing to acquire (n=5) or extinguish (n=8). Six subjects were eliminated from statistical analysis after data collection was complete (2 from the 10 min group, 1 from the 6 hr group, and 3 from the no-reminder group) for failing to acquire or extinguish . Finally, there were 5 participants who were excluded because their data collection was incomplete due to technical problems. Thus, the final sample included 65 subjects (41 females, age range = 18-48). All subjects gave informed consent and were paid for their participation. Data for all participants for whom we have complete data sets can be found on Open Science Framework: https://osf.io/jhu5c/.

The exclusion or inclusion of participants during the study was based on a qualitative assessment of trial-by-trial SCR as reflecting a pattern consistent with successful acquisition and extinction. Below, we provide a more exact characterization of the final

data set, corresponding to the qualitative evaluations made in the process of data collection and analysis.

The characterization of the data that emerged from the qualitative judgments made during the study is described below. Specifically, "successful acquisition and extinction" for the participants in Study 1 can be translated to a chain of logical "IF" statements:

- If during acquisition the differential CS response (CS+ minus CS-) was below an individually standardized cut-off (value of 0.1 divided by the mean US response) on: (1) the first half of acquisition, (2) the second half acquisition, (3) last trial of acquisition, and (4) the increase from the first to last trial of acquisition. When *all* of these criteria were met then acquisition was deemed as failed, otherwise the participant was included.
- If during extinction the differential CS response (CS+ minus CS-) was above an individually standardized cut-off on: 1) the second half of extinction, 2) the last trial of extinction, and 3) below cutoff in the decrease from the first to the last trials or halves of extinction. When *all* of these criteria were met then extinction was deemed as failed, otherwise the participant was included.
- In a limited number of cases, a decision was made to (a) include 2 participants that did not meet criteria for acquisition because the numerical distance from the cutoff was negligible (<0.005), or all measurable differential responses were positive or equal to zero, (b) to include 1 participant that did not meet criteria for extinction because the numerical distance from the cutoff for a single criterion was negligible (.01), (c) to exclude 3 participants despite not meeting all acquisition exclusion criteria for idiosyncratic responses not observed in other participants in the sample (e.g., CR on last acquisition trial was negative or almost equal to zero) and (d) to exclude 6 participants despite not meeting all extinction exclusion criteria, for idiosyncratic responses not observed in other

participants in the sample (e.g., full reversal of the differential response, which might inflate the recovery index).

Behavioral Paradigm

The paradigm consisted of three consecutive stages conducted 24 hr apart: Day 1 - Acquisition, Day 2 - Reactivation and Extinction, and Day 3 - Re-extinction (Fig. 1a). During acquisition, three randomly assigned groups of subjects underwent a Pavlovian discrimination fear conditioning paradigm with partial reinforcement. The CSs were yellow and blue squares and the US was a mild shock to the wrist. The CS+ was paired with the shock on a 38% partial reinforcement schedule and the CS- was never paired with shock. Subjects were instructed to pay attention to the computer screen and try to figure out the relationship between the stimuli appearing on the screen and the shocks. A day later, all three groups underwent extinction training where the CS+ and CS- were repeatedly presented without the US. In two groups, the fear memory was reactivated prior to extinction. During reactivation, the CS+ was presented once (unreinforced), followed by a 10 min break. One group (n = 20) underwent extinction after the 10 min break (within the reconsolidation window). The second group (n = 23) underwent extinction 6 hr after the reactivation (outside of the reconsolidation window). The third group (n = 22) was not reactivated but directly proceeded to the 10 min break. Following the break, extinction immediately followed for half the subjects in this group, or was conducted 6 hr later for the other half. During the break, all participants watched a pre-selected TV show episode. Day 3 consisted of re-extinction where participants were presented with nonreinforced presentations of the stimuli. During all sessions (acquisition, reminder, extinction and re-extinction), with the exception of the breaks, the participants were attached to the SCR and shock electrodes, and the shock stimulator was set to the ‘On’ position.

The stimuli were presented for 4 sec each with a 10-12s variable inter-trial-interval (ITI). Acquisition consisted of 10 nonreinforced presentations of each of the CSs, intermixed with an additional 6 CS+ presentations that co-terminated with the shock (200 ms, 50 pulses/s). Extinction for the two reactivation groups consisted of 10 nonreinforced CS+ and 11 CS- presentations. Since the CS- was not presented during the reactivation session, an extra presentation was added to the extinction session so all CSs were presented an equal number of times during Day 2. Extinction for the no-reminder group consisted of 11 nonreinforced presentation of CS+ and CS-. Re-extinction was similar for all groups and consisted of nonreinforced presentation of the stimuli (10 CS+, 11 CS-). The first CS- trial in each phase was disregarded due to the orienting response at the beginning of the session. Two orders were used to counterbalance for trial order and designations of colored squares (blue or yellow) as CS+ or CS-. During the breaks the shock stimulator was set to the 'Off' position.

In an attempt to examine how long the blockade of memory persists, we invited the participants of the experiment to come back to the lab after about a year (10-14 months). Twenty three participants were located (10 min group, n = 10; 6 hr group, n = 5; no-reminder group, n = 8). As mentioned above, following the spontaneous recovery test, subjects were re-extinguished using 10 non-reinforced presentations of the stimuli, which allowed us to test again their recovery of fear. We used a more potent recovery essay, namely, reinstatement, in which subjects were exposed to 4 unsignaled shocks, followed by non-reinforced presentations of the same CSs that were used in the spontaneous recovery experiment (10 presentations each, using two randomized orders counterbalanced across subjects). The index of fear recovery was the difference in the conditioned fear response at the end of re-extinction following the initial spontaneous recovery test and the conditioned fear response immediately after reinstatement a year later. Specifically, a differential SCR score (CS+ minus CS-) was calculated for the end

of re-extinction (mean of last 2 trials) and post-reinstatement (mean of first 4 trials). The first CS+ and CS- trial were disregarded due to the orienting response at the beginning of the session. We collapsed subjects from the two groups previously showing spontaneous recovery (i.e., 6 hr and no-reminder) into one group. Subjects that failed to re-extinguish following the spontaneous recovery test (differential SCR score > 0.2) or showed no measurable responses to the shocks during reinstatement were not included in the analysis (4 subjects). The final analysis included 19 subjects (10 min group, n = 8; 6 hr / no-reminder group, n = 11). Throughout the session, the participants were attached to the SCR and shock electrodes, and the shock stimulator was set to the 'On' position. They were instructed to sit still and pay attention to the screen. The stimuli were presented for 4 sec each with a 10-12s variable ITI.

Psychophysiological stimulation and assessment

Mild shocks were delivered through a stimulating bar electrode attached with a Velcro strap to the right inner wrist. A Grass Medical Instruments stimulator (West Warwick, Rhode Island) charged by a stabilized current was used. Subjects determined the level of the shock themselves, beginning at a very mild level of shock (10 V) and gradually increasing the level until the shock reached the maximum level that they determined was uncomfortable, but not painful (the maximum level was 60 V). All shocks were given for 200ms, with a current of 50 pulses per second.

Skin conductance response (SCR) was assessed using two Ag-AgCl electrodes, which were connected to a BioPac Systems (Santa Barbara, California) skin conductance module. The electrodes were attached to the first and second fingers of the left hand, between the first and second phalanges. SCR waveforms were analyzed offline, using AcqKnowledge 3.9 software (BIOPAC Systems Inc., Goleta, California). SCR amplitudes to the CS and US were the dependent measures of conditioned and

unconditioned responses, respectively. The level of SCR response was determined by taking the base to peak difference for the first waveform (in micro siemens, μs) in the 0.5 to 4.5 s window following stimulus onset. The minimal response criterion was 0.02 μs. The raw SCR scores were square root transformed to normalize distributions. These normalized scores were scaled according to each subjects unconditioned response by dividing each response by the mean square-root-transformed US response.

Statistical Analysis

To assess expectations for the aversive outcome separate from unconditioned responses to the shocks themselves, we included only non-reinforced trials of CS+ in the analysis. The differential fear response was assessed by subtracting responses to the CS- from responses to the CS+ in corresponding trials. The differential scores were averaged across subjects. For each phase, we conducted two-way ANOVAs with main factors of group (10 min, 6hr, and no-reminder) and time (early and late phase). Follow up *t*-tests were used to assess acquisition of fear by comparing the mean differential response during the second half of the acquisition session. Extinction was assessed by comparing the mean differential response during the last trial of extinction. The change in fear response from acquisition to extinction was assessed by comparing the differential response in the second half of acquisition to the last trial of extinction. To test for spontaneous recovery of fear, the differential response in the first re-extinction trial was compared with the last extinction trial. Two-tailed *t*-tests and an alpha level of 0.05 were used for all statistical comparisons unless mentioned otherwise.

Experiment 2: Preventing the reinstatement of fear memories

Participants

Seventy subjects were recruited by poster or email advertisements. Fifty-two subjects were excluded at different stages of data collection for either being a non-responder or failing to acquire or extinguish conditioned responses to both CS+ stimuli. We used these criteria because we could not assess fear recovery when fear responses were not successfully and equally acquired to the two conditioned stimuli. Like in Experiment 1, the exclusion or inclusion of participants was based on a qualitative assessment of trial-by-trial SCR as reflecting a pattern consistent with successful acquisition and extinction. Thus, the final sample included 18 subjects (10 females, age range = 18-34). All subjects gave informed consent and were paid for their participation. Data for all the subjects included in the final sample can be found on Open Science Framework; https://osf.io/jhu5c/.

Behavioral Paradigm

The paradigm consisted of three consecutive stages conducted 24 hr apart: Day 1 - Acquisition, Day 2 - Reactivation and Extinction, and Day 3 - Reinstatement and Re-extinction, in a within-subject design (Fig. 2a). During Acquisition, subjects underwent fear conditioning using three colored squares. Two squares (CS*a*+ and CS*b*+) were paired with the shock on a 38% partial reinforcement schedule. The third square (CS-) was never paired with the shock. Subjects were instructed to pay attention to the computer screen and try to figure out the relationship between the stimuli appearing on the screen and the shocks. Day 2 consisted of reactivation and extinction. During reactivation, the CS*a*+ and the CS- were each presented once (unreinforced), in a counterbalanced fashion. Participants were then given a 10 min break where they watched a pre-selected TV show episode. Extinction immediately followed and consisted

of nonreinforced presentations of the three stimuli (CS*a*+, CS*b*+, CS-). Day 3 consisted of reinstatement and re-extinction. During reinstatement, subjects were administered with 4 unsignalled shocks. After a 10 min break, a re-extinction session commenced where participants were presented with nonreinforced presentations of the three stimuli (CS*a*+, CS*b*+, CS-). During all sessions, (acquisition, reminder, extinction, reinstatement and re-extinction), with the exception of the breaks, the participants were attached to the SCR and shock electrodes, and the shock stimulator was set to the 'On' position. During the breaks the shock stimulator was set to the 'Off' position.

The stimuli were presented for 4 sec each with a 10-12s variable inter-trial-interval (ITI). Acquisition consisted of 8 nonreinforced presentations of CS*a*+, CS*b*+ and CS- each, intermixed with an additional 5 CS*a*+ and 5 CS*b*+ presentations that co-terminated with the shock (200 ms, 50 pulses/s). Extinction consisted of nonreinforced presentations of 10 CS*a*+, 11 CS*b*+, and 11 CS-. Since CS*b*+ was not presented during the reactivation session, an extra presentation was added into the extinction session, so the CS*a*+ and CS*b*+ were presented an equal number of times during Day 2. Re-extinction consisted of nonreinforced presentation of the stimuli (10 CS*a*+, 10 CS*b*+, and 11 CS-). The first CS- trial in each phase was disregarded due to the orienting response at the beginning of the session. Two orders were used to counterbalance for trial order and designations of colored squares into CS*a*+, CS*b*+, and CS- (Order A: yellow = CS*a*+, red = CS*b*+, blue = CS- ; Order B: Orange = CS*a*+, Green = CS*b*+, Pink = CS-).

Psychophysiological stimulation and assessment

Same as in Experiment 1 (see above).

Statistical Analysis

To assess expectations for the aversive outcome separate from unconditioned responses to the shocks themselves, we included only non-reinforced trials of CS*a*+ or CS*b*+ in this analysis. For each phase, we conducted two-way ANOVAs with main factors of stimulus (CS*a*+, CS*b*+, CS-) and time (early and late responses within each phase). Follow up *t*-tests were conducted to assess fear acquisition by comparing each of the conditioned stimuli (CS*a*+ or CS*b*+) to the CS- during the second half of the acquisition session (last 4 trials). Extinction was assessed by comparing each of the conditioned stimuli (CS*a*+ or CS*b*+) to the CS- during the last trials of extinction. We also compared CS*a*+ to CS*b*+ to verify that the two stimuli did not differ in acquisition or extinction of fear. The change in fear response from acquisition to extinction was assessed by comparing the mean second half of acquisition with the last trial of extinction for each stimulus (CS*a*+, CS*b*+, CS-). To test for fear reinstatement, the response to each stimulus in the first re-extinction trial was compared with that of the last extinction trial. Two-way *t*-tests and an alpha level of 0.05 were used for all statistical comparisons.

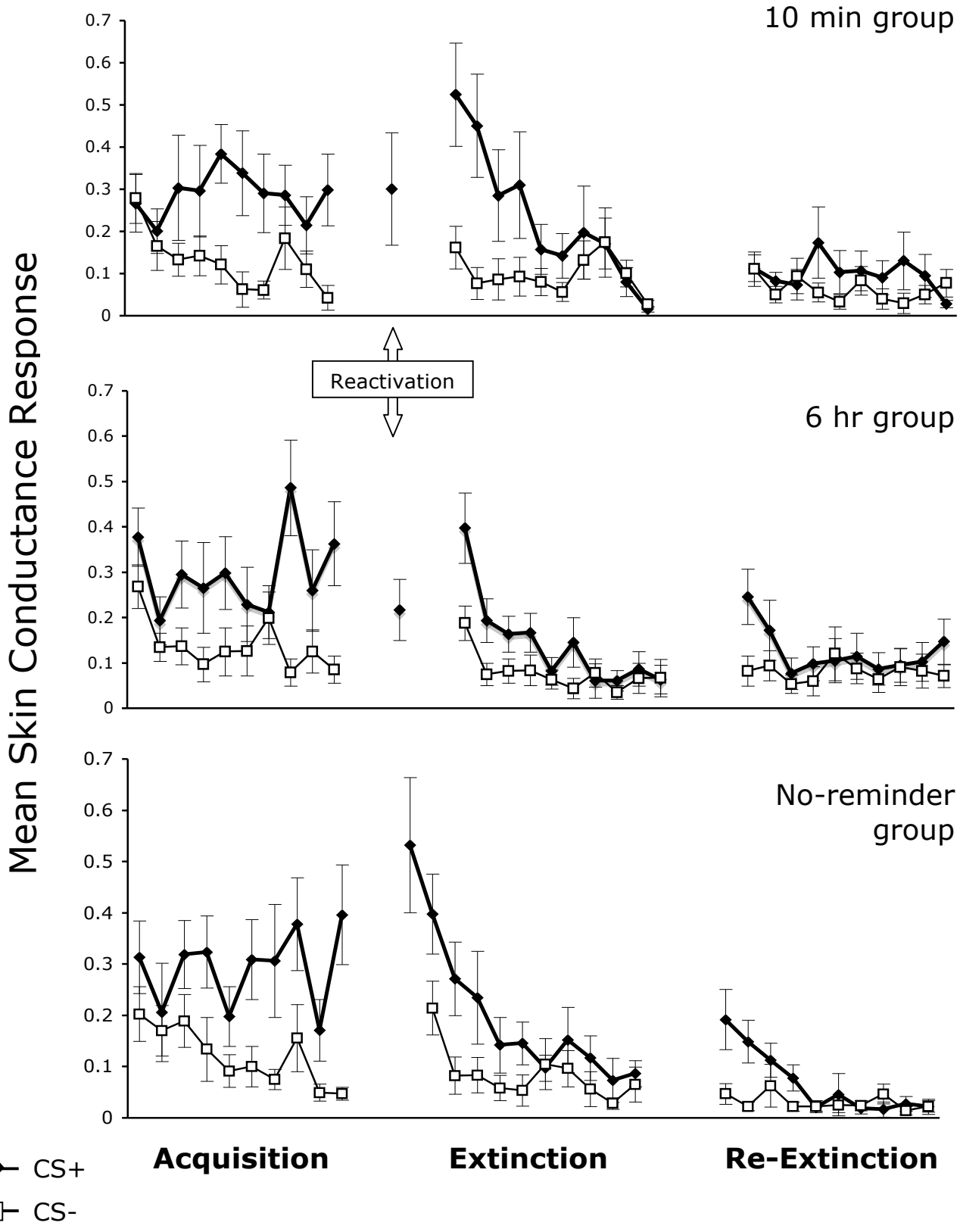
Supplementary Figure legends

Figure S1. Extinction during reconsolidation prevents spontaneous recovery of extinguished fear. Mean SCRs for CS+ and CS- presentations throughout acquisition (excluding the CS+ trials paired with the shock), reactivation, extinction and re-extinction are presented for each experimental group (10 min reminder, 6 hr reminder, and no-reminder). Error bars represent standard errors.

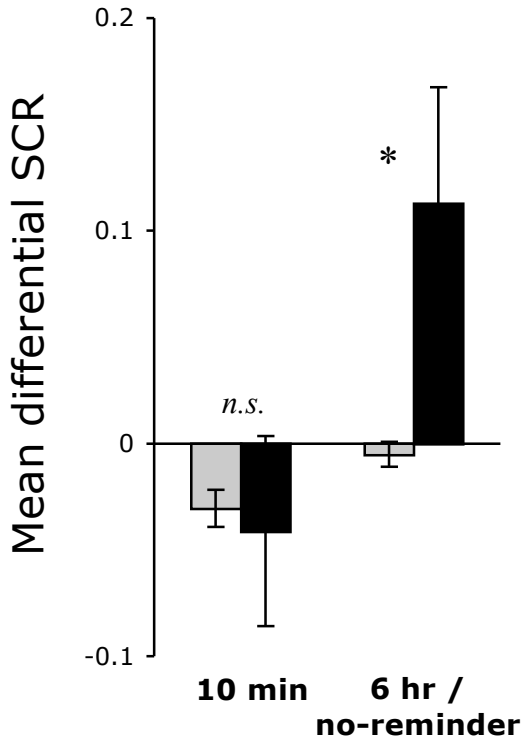
Figure S2. Blockade of the return of fear persists one year later. Mean conditioned SCR response (CS+ minus CS-) of re-extinction at the end of the spontaneous recovery experiment (gray bars) and re-extinction a year later following reinstatement (black bars) for each experimental group (10 min group, 6 hr / no-reminder group). Both groups ended the previous experiment showing no conditioned fear. A year later, subjects in the 6 hr / no-reminder group, but not the 10 min group, showed significant recovery of fear.

Figure S3. Extinction during reconsolidation prevents reinstatement of extinguished fear. Mean SCRs throughout acquisition (excluding the CS+ trials paired with the shock), reactivation, extinction and re-extinction are presented for each stimulus (CS*a*+, CS*b*+, CS-). Responses to the CS*a*+ and CS- reminders are collapsed due to proximity of presentation. Error bars represent standard errors.

Supplementary Figure 1



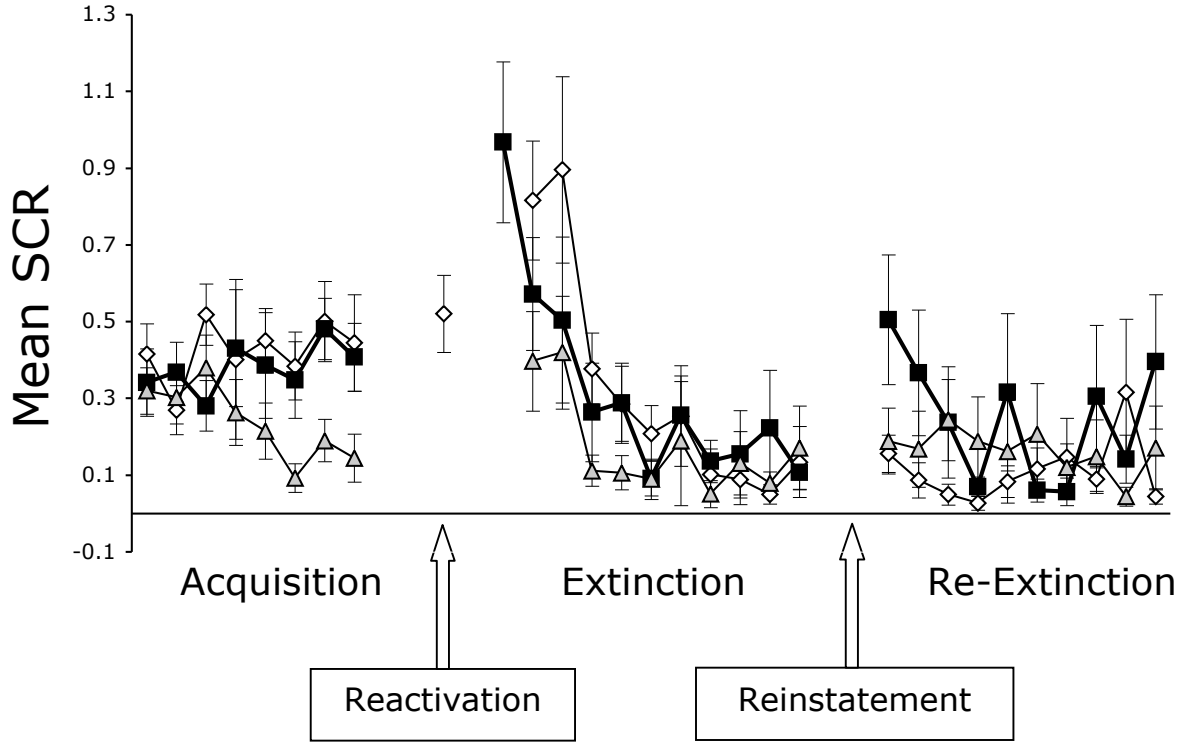
Supplementary Figure 2



Re-extinction

Post-Reinstatement

Supplementary Figure 3



CS*a*+ (reminded)

CS*b*+ (not reminded)

CS-