



# Plasticity of Fear and Safety Neurons of the Amygdala in Response to Fear Extinction

Susan Sangha<sup>1,2\*</sup>

<sup>1</sup> Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA, <sup>2</sup> Ernest Gallo Clinic and Research Center, University of California, San Francisco, Emeryville, CA, USA

Fear inhibition learning induces plasticity and remodeling of circuits within the amygdala. Most studies examine these changes in nondiscriminative fear conditioning paradigms. Using a discriminative fear, safety, and reward conditioning task, Sangha et al. (2013) have previously reported several neural microcircuits within the basal amygdala (BA) which discriminate among these cues, including a subpopulation of neurons responding selectively to a safety cue and not a fear cue. Here, the hypothesis that these “safety” neurons isolated during discriminative conditioning are biased to become fear cue responsive as a result of extinction, when fear behavior diminishes, was tested. Although 41% of “safety” neurons became fear cue responsive as a result of extinction, the data revealed that there was no bias for these neurons to become preferentially responsive during fear extinction compared to the other identified subgroups. In addition to the plasticity seen in the “safety” neurons, 44% of neurons unresponsive to either the fear cue or safety cue during discriminative conditioning became fear cue responsive during extinction. Together these emergent responses to the fear cue as a result of extinction support the hypothesis that new learning underlies extinction. In contrast, 47% of neurons responsive to the fear cue during discriminative conditioning became unresponsive to the fear cue during extinction. These findings are consistent with a suppression of neural responding mediated by inhibitory learning, or, potentially, by direct unlearning. Together, the data support extinction as an active process involving both gains and losses of responses to the fear cue and suggests the final output of the integrated BA circuit in influencing fear behavior is a balance of excitation and inhibition, and perhaps reversal of learning-induced changes.

## OPEN ACCESS

### Edited by:

Denise Manahan-Vaughan,  
Ruhr University Bochum, Germany

### Reviewed by:

Marie H. Monfils,  
University of Texas at Austin, USA  
Stephen Maren,  
Texas A&M University, USA  
Stefan Herlitze,  
Ruhr-University Bochum, Germany

### \*Correspondence:

Susan Sangha  
sangha@purdue.edu

**Keywords:** amygdala, fear, safety, extinction

## INTRODUCTION

Environmental cues signifying danger, safety, or reward availability can have a potent effect in emotion regulation. Accurately discriminating among these cues is important in initiating the proper emotional response in order to guide behavior. Maladaptive emotion regulation can lead to a wide-range of clinical problems, such as anxiety disorders and addiction. Since potentially rewarding and dangerous stimuli often occur simultaneously leading to opposing behaviors of approach or avoidance, respectively, reward- and fear-related circuits must interact in order to mediate these antagonistic behaviors. Approach and avoidance behaviors can also be modulated

**Received:** 27 August 2015

**Accepted:** 06 December 2015

**Published:** 24 December 2015

### Citation:

Sangha S (2015) Plasticity of Fear and Safety Neurons of the Amygdala in Response to Fear Extinction. *Front. Behav. Neurosci.* 9:354. doi: 10.3389/fnbeh.2015.00354

by signals that inform the organism if the environment is safe or not. The inability to discriminate among danger, safety, and reward cues can lead to generalized fear responses that are enhanced in Post-traumatic Stress Disorder (PTSD) patients (Jovanovic et al., 2012).

Behavioral therapy for maladaptive fear often involves repeated exposures to the danger cue in the absence of an aversive outcome, a procedure known as extinction. Through repeated exposures, the subject feels an increasing sense of control over the situation and fear diminishes. Safety conditioning is another method of reducing fear. During safety conditioning, a safety cue in conjunction with a danger cue signifies no aversive outcome whereas the danger cue on its own does result in an aversive outcome. Thus, extinction and safety conditioning are related but distinct phenomena. Safety cues can even act as positive reinforcers, suggesting the mechanisms of safety learning may overlap with reward learning (Christianson et al., 2012; Sangha et al., 2013).

The amygdala has been consistently implicated in processing and regulating a myriad of emotional responses (for review see Janak and Tye, 2015). The basal amygdala (BA) in particular is important for discriminating among sensory stimuli that signal multiple outcomes of a similar valence (Málková et al., 1997; Corbit and Balleine, 2005; Balleine and Killcross, 2006), and it possesses neuronal populations selective for valence (Schoenbaum et al., 1999; Paton et al., 2006; Belova et al., 2007; Shabel and Janak, 2009; Sangha et al., 2013).

Evidence suggests that fear extinction learning induces plasticity and remodeling of inhibitory circuits and synapses within the amygdala (Heldt and Ressler, 2007; Lin et al., 2009; Sangha et al., 2012), as well as decreased synaptic efficacy in the medial prefrontal cortex-BA pathway (Cho et al., 2013). Within the BA, “extinction” neurons have been reported (Herry et al., 2008). These are neurons that are unresponsive to a fear cue before extinction but become responsive to the fear cue after extinction, when fear behavior is diminished. Diminished fear behavior is also seen during safety conditioning in response to a safety cue. Using a discriminative conditioning task that allows assessment of fear, safety and reward cue learning together, Sangha et al. (2013) demonstrated significant suppression of freezing behavior in response to a compound fear+safety cue compared to the high freezing seen in response to a fear cue. In addition, this study also reported several neural microcircuits within the BA that showed a discriminative response to these cues. In particular, 24% of recorded neurons were responsive to the compound fear+safety cue but unresponsive to the fear cue when presented alone suggesting these neurons are encoding safety. Similar to these “safety” neurons, the “extinction” neurons reported by Herry et al. (2008) were also unresponsive to the fear cue before extinction training. Since safety conditioning and extinction are related phenomena, neurons classified as “safety” neurons in Sangha et al. (2013) were here examined through extinction to see if they became “extinction” neurons, similar to the neurons reported by Herry et al. (2008).

To do this, firing rates of neurons classified as discriminative, nondiscriminative or unresponsive during discriminative conditioning (DC), based on their responses to the fear cue

alone and the compound fear+safety cue, were examined in response to the fear cue during extinction training and recall as fear behavior decreased. The hypothesis tested is that there is a bias for the neurons that are safety cue responsive during DC to become responsive to the fear cue as fear extinction progresses.

## MATERIALS AND METHODS

### Subjects

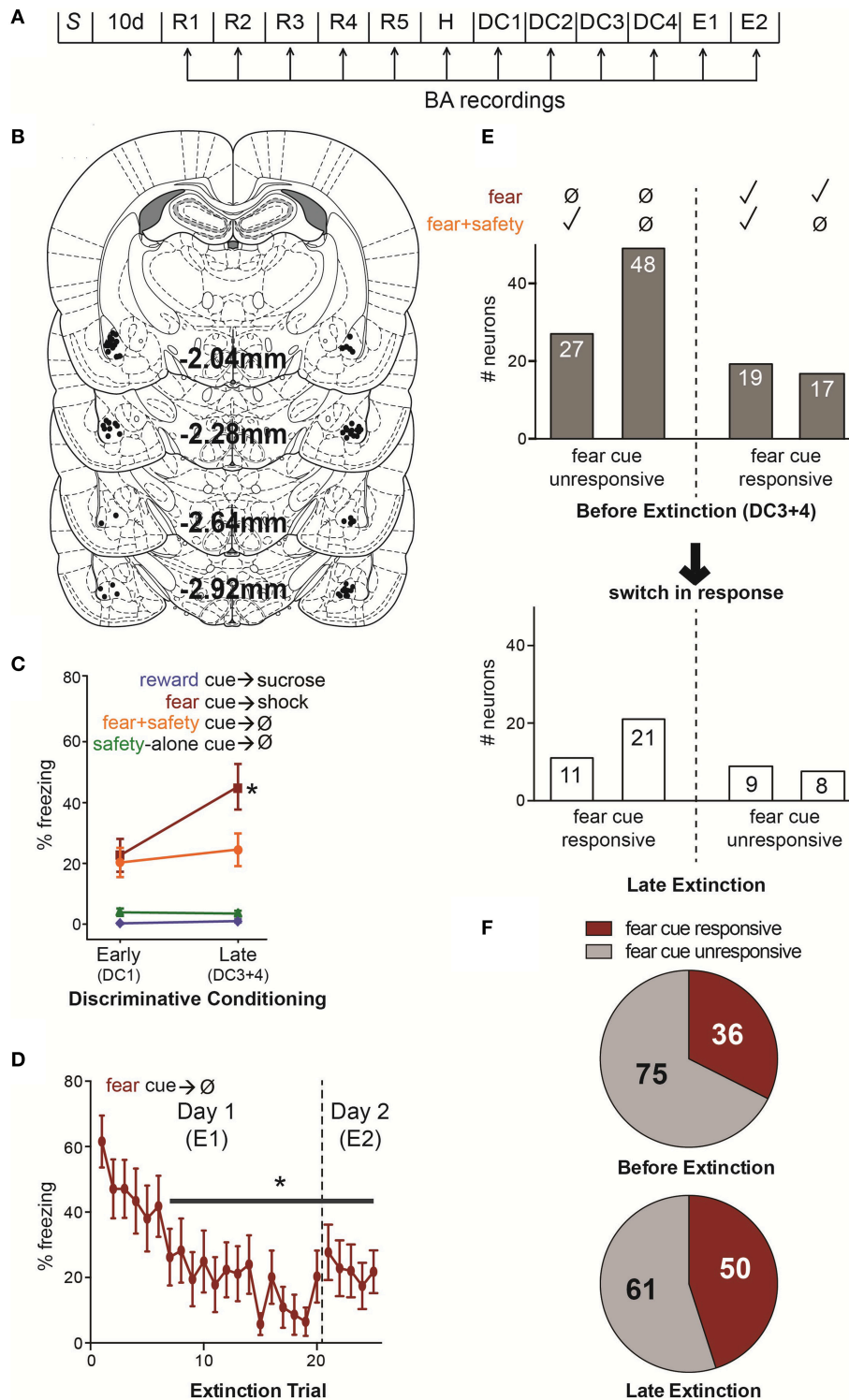
Fourteen Long Evans male rats (Harlan) weighing 350–400 g at the beginning of experiments were single housed under a 12 h light/dark cycle (lights on 07:00) and handled for 1 week before commencing experiments. All procedures were performed during the light cycle and approved by the Gallo Center Institutional Animal Care and Use Committee in accordance with the National Institute of Health guidelines. Rats had *ad libitum* access to food and water up until the third reward learning session, at which point they were restricted to 22 g of food per day for the remainder of the experiment.

### Behavioral Apparatus

The experimental chambers, used in all experiments and obtained from MedAssociates, were Plexiglas boxes (32 cm length × 31 cm width × 35 cm height) encased in sound-attenuating shells. A recessed port 3 cm above the floor and located in the center of one wall was used to deliver sucrose. Two lights (28 V, 100 mA) located 12 cm from the floor on the wall opposite the port provided constant illumination. A light (28 V, 100 mA) located 33 cm above the floor on the wall opposite the port served as the 20 s continuous light cue. A high-frequency “tweeter” speaker (ENV-224BM) located 25 cm from the floor on the wall opposite the port was used to deliver the auditory cues. Footshock was delivered through a grid floor via a constant current aversive stimulator (ENV-414S). A video camera located at the top of the sound-attenuating shell recorded the rat’s behavior for offline video analysis.

### Discriminative Conditioning

The three cues signifying reward, fear or safety were a 20 s continuous 3 kHz tone (70 dB), a 20 s pulsing 11 kHz tone (200 ms on, 200 ms off; 70 dB) or a 20 s continuous light (28 V, 100 mA), counterbalanced across subjects, with the caveat that the light cue was reserved for the safety cue in most subjects, 12 out of 14 rats. Training first consisted of five reward sessions (Figure 1A; R1–5), in which a 20 s reward cue was paired with 3 s delivery of a 10% sucrose solution (100  $\mu$ L) into a port accessible to the rat (3 s sucrose delivery commenced pseudorandomly between 10 and 20 s after reward cue onset for 25 trials, ITI 90–130 s). This was followed by a single session of habituation (H) to the future fear cue and safety cue during a session in which reward cue training continued (25 reward trials, ITI 90–130 s). The future fear cue and safety cue were presented separately five times each for 20 s without reinforcement to allow subjects to habituate to their presentation thereby reducing any baseline freezing to these novel cues. Four sessions of discriminative conditioning followed (DC1–4): reward cue training continued



**FIGURE 1 | (A)** Summary of experimental design. S, surgical implantation of electrodes into the BA bilaterally followed by 10 d surgical recovery. R1–5, reward sessions in which the reward cue was paired with sucrose delivery. H, habituation in which, in addition to the reward cue-sucrose pairings, rats also received unreinforced presentations of the future fear and safety cues. DC1–4, discriminative conditioning in which reward cue-sucrose pairings continued as well as the addition of trials where the fear cue was paired with footshock, the fear cue was paired with the safety cue without footshock, or the safety cue was presented alone without footshock. E1–2, extinction in which the fear and reward cues were presented unreinforced. **(B)** Locations of each electrode tip from 14 rats. All 111 recorded neurons were in the BA. **(C)** Mean (±SEM) percentage of time spent freezing during each cue comparing early vs. late DC sessions (DC1 vs. DC3+4). During late DC, (Continued)

**FIGURE 1 | Continued**

rats freeze significantly more to the fear cue compared to the fear+safety cue, reward cue or safety-alone cue, demonstrating discriminatory fear behavior ( $p < 0.05$ ). **(D)** Mean ( $\pm$ SEM) percentage of time spent freezing for each fear cue trial during E1 and E2. Freezing was significantly suppressed compared to the first trial beginning at trial 7 and remained significantly suppressed for the remainder of trials during E1 and E2 ( $p < 0.05$ ). **(E)** Summary of fear cue unresponsive and responsive neurons before extinction and during late extinction. Above, neurons were assigned to one of four groups based on their response to the fear cue and fear+safety cue during late DC (DC3+4); i.e., before extinction. A neuron was considered responsive if there was a significant change in firing frequency during the first 200 ms of the cue compared to pre-cue baseline. Below, a summary of the subset of neurons from each of the four groups to switch their response to the fear cue during late extinction (trials 10–20 of E1 and trials 1–5 of E2 in which freezing behavior was significantly lowered). From left to right, before extinction, one group ( $n = 27$ ) showed no response to the fear cue but did show a significant change in firing frequency in response to the fear+safety cue. During late extinction, 11 of these neurons switched to being fear cue responsive. The next group ( $n = 48$ ) showed no response to either the fear or fear+safety cue before extinction. But during late extinction, 21 of these neurons became fear cue responsive. In contrast, the next group ( $n = 19$ ) showed a significant change in firing frequency in response to both the fear and fear+safety cue before extinction and nine of these neurons became fear cue unresponsive during late extinction. The last group ( $n = 17$ ) showed a significant change in firing frequency in response to the fear cue but not the fear+safety cue before extinction. Of these neurons, eight became fear cue unresponsive. **(F)** Comparison of the number of neurons that were fear cue responsive, irrespective to its responding to the other cues, before extinction (DC3+4) to late extinction. The number of neurons being fear cue responsive increased from 36 before extinction to 50 during late extinction, a 39% increase.

(3 s sucrose delivery commenced 18 s after reward cue onset; 15 trials), along with the additional presentation of the 20 s fear cue followed by a mild 0.5 s footshock at the offset of the fear cue (0.4 mA; four trials). On separate trials this same 20 s fear cue was simultaneously paired with a 20 s safety cue resulting in no footshock (fear+safety cue; 15 trials). Trials in which the 20 s safety cue was presented alone without any footshock were also included (safety-alone cue; 10 trials) to assess if any freezing developed to the safety cue as a result of being paired to the fear cue as well as providing the animal with additional trials that contained a safety cue-no shock contingency. Trials were presented pseudorandomly (ITI 100–140 s). Two sessions of extinction followed (E1–2), in which the fear and reward cues were presented unreinforced (E1: 20 trials each of the fear and reward cues, E2: five trials each of the fear and reward cues; trials were presented pseudorandomly, ITI 90–130 s).

## Behavioral Analyses

Fear behavior was assessed, offline from videos, by measuring freezing, defined as complete immobility with the exception of respiratory movements, which is an innate defensive behavior (Blanchard and Blanchard, 1969; Fendt and Fanselow, 1999). The total time spent freezing was quantified during the entire 20 s of each cue presentation and expressed as percent time spent freezing. Calculating the percent time spent in the port assessed reward-seeking behavior. Behavioral data were analyzed using one- or two-way repeated measures ANOVA, followed by Tukey's *post-hoc* test when indicated by significant ( $p < 0.05$ ) main effects or interactions.

## Surgery

Rats were anesthetized with isoflurane and stereotaxically implanted bilaterally with fixed eight-electrode arrays (NeuroBiological Laboratories) directed at the BA (relative to bregma: AP:  $-2.04$  to  $-2.92$  mm posterior, ML: 4.1–4.9 mm, DV: 6.6–7.5 mm ventral from brain surface (Paxinos and Watson, 2007) (Figure 1B). Rats were allowed 7–10 d to recover in which they had *ad libitum* access to food and water.

## In Vivo Single Unit Recordings

Neuronal activity was recorded with commercial hardware and software, including headstage amplifiers and programmable

amplifiers, filters (0.4 and 5 KHz), and multichannel spike-sorting software (Plexon). Implanted rats were connected to the recording apparatus via a swivel commutator. Discrimination of individual units was performed offline by using principal component analysis of waveform shape. Single cells were identified by constancy of waveform shape, cross-correlograms, and interspike intervals (Janak, 2002). In addition, quantitative J3 and Davies Bouldin validity index (DB) statistics were calculated. High J3 values and low DB values are indicative of good single unit isolation (Davies and Bouldin, 1979; Nicolelis et al., 2003; Herry et al., 2008; Sangha et al., 2013). Stability of units across sessions was assessed by calculating principal component space cylinders using WaveTracker (Plexon). In addition, linear correlation values between time-shifted average waveforms were calculated (Jackson and Fetzi, 2007; Herry et al., 2008; Sangha et al., 2013). As a control, the *r*-values from average waveforms of randomly paired neurons and sessions were computed. Only units deemed stable across sessions using these procedures were included in the analysis.

## Classification of Neurons

For each neuron, significance of cue-evoked firing rates was determined as previously published (Sangha et al., 2013), using a 10,000-round paired permutation test (Hesterberg et al., 2005) comparing the averaged 20 s pre-cue baseline period to the first 200 ms after cue onset during the last two DC sessions and during late extinction (trials 10–20 of E1 and trials 1–5 of E2). That is, the 20 s pre-cue baseline firing rates and the 200 ms post-cue firing rates for a given cue were shuffled and redistributed independently 10,000 times. The differences between the baseline and post-cue firing for the single real case and the 10,000 reshuffled cases were used to create a distribution. In accordance with the permutation test, if the actual mean difference was within  $<2.5\%$  of either tail, it was considered significant. *P*-values were then adjusted for multiple corrections using the Benjamini-Hochberg correction with a corrected cutoff of 0.05 (Benjamini and Hochberg, 1995). To avoid false positives, neurons that showed a significant cue-evoked inhibition using this permutation test were only included in the final analyses if the baseline firing frequency was  $>0.05$  Hz. Neurons were classified as “fear cue responsive” if there was a significant



increase or decrease in firing rate to the fear cue during late DC. These neurons were then segregated based on whether there was also a significant change in firing rate to the fear+safety cue. Neurons that did not show a significant change in firing rate to the fear cue during late DC (i.e., before extinction) were classified as “fear cue unresponsive.” A subset of these neurons did however show a significant increase or decrease in firing rate compared to baseline to the fear+safety cue and were analyzed separately. Similarly, neurons were classified as “reward cue responsive” if there was a significant increase or decrease in firing rate to the reward cue during late DC.

## Histology

Rats were deeply anesthetized with sodium pentobarbital. A 10 s 19  $\mu$ A DC current was passed through each wire to mark each electrode tip. Rats were then perfused with formalin containing 3% potassium ferrocyanide. Sections (50  $\mu$ M) were stained against acetylcholinesterase and only units recorded from electrode wires verified to be in the BA were included in the analyses.

## RESULTS

In a previous study neurons of the BA were tracked over the course of a discriminative conditioning task (Sangha et al., 2013). In this task rats learn to discriminate among fear, safety, and reward cues. In the present study, the same BA neurons were followed into fear and reward cue extinction to assess the plasticity of neurons that were fear cue responsive and fear cue unresponsive before extinction.

Recordings were made during each behavioral training session (Figure 1A, see Materials and Methods). A total of 111 single neurons located in the BA from 14 rats (Figure 1B) were isolated from recordings made during discriminative conditioning and extinction. Most neurons had low mean firing rates (Median = 0.83 Hz, Max = 20.35 Hz, Min = 0.06 Hz), suggesting the sample was predominantly putative projection neurons (Likhnik et al., 2006).

## Fear Behavior

### Discriminative Conditioning

The percent time spent freezing during each cue was averaged across early (first DC session, DC1) and late (final two DC sessions, DC3+4) discriminative conditioning (Figure 1C). A Two-way repeated-measures ANOVA on percent time spent freezing revealed a significant interaction between phase of training and cue type [ $F_{(3, 39)} = 8.575, p < 0.001$ ] and a main effect of phase of training [ $F_{(1, 13)} = 5.118, p < 0.05$ ] and cue type [ $F_{(3, 39)} = 29.331, p < 0.001$ ]. Freezing to the fear cue was significantly greater than the fear+safety cues, safety-alone cue, and reward cue during late DC (*post-hoc* Tukey's,  $p < 0.001$  each comparison), demonstrating discriminative fear behavior by these animals.

### Fear Extinction

The percent time spent freezing during each fear cue trial of E1 and E2 was averaged across animals (Figure 1D). A One-way

repeated-measures ANOVA on percent time spent freezing revealed a main effect of trial [ $F_{(24, 336)} = 6.35, p < 0.0001$ ] and Dunnett's multiple comparisons test showed freezing was significantly lower ( $p < 0.05$ ) during trial 7 and each subsequent trial compared to the first trial. Thus, freezing was significantly suppressed compared to the first trial beginning at trial 7 and remained significantly suppressed for the remainder of trials during E1 and E2.

## Neural Recordings of Fear and Safety Neurons

In order to compare neuronal responding during discriminative fear behavior to significant fear suppression during extinction, neuronal responding was analyzed during late DC (DC3+4) and compared to late extinction (Figure 1E). Late extinction consisted of the last 10 trials of E1 and the 5 trials of E2; freezing behavior during each of these trials was significantly lower than the beginning of extinction (trial 1 of E1, Figure 1D). Z-scores were calculated for each neuron's response to the first 200 ms of each cue (see Materials and Methods) and used to make comparisons among different neuronal populations.

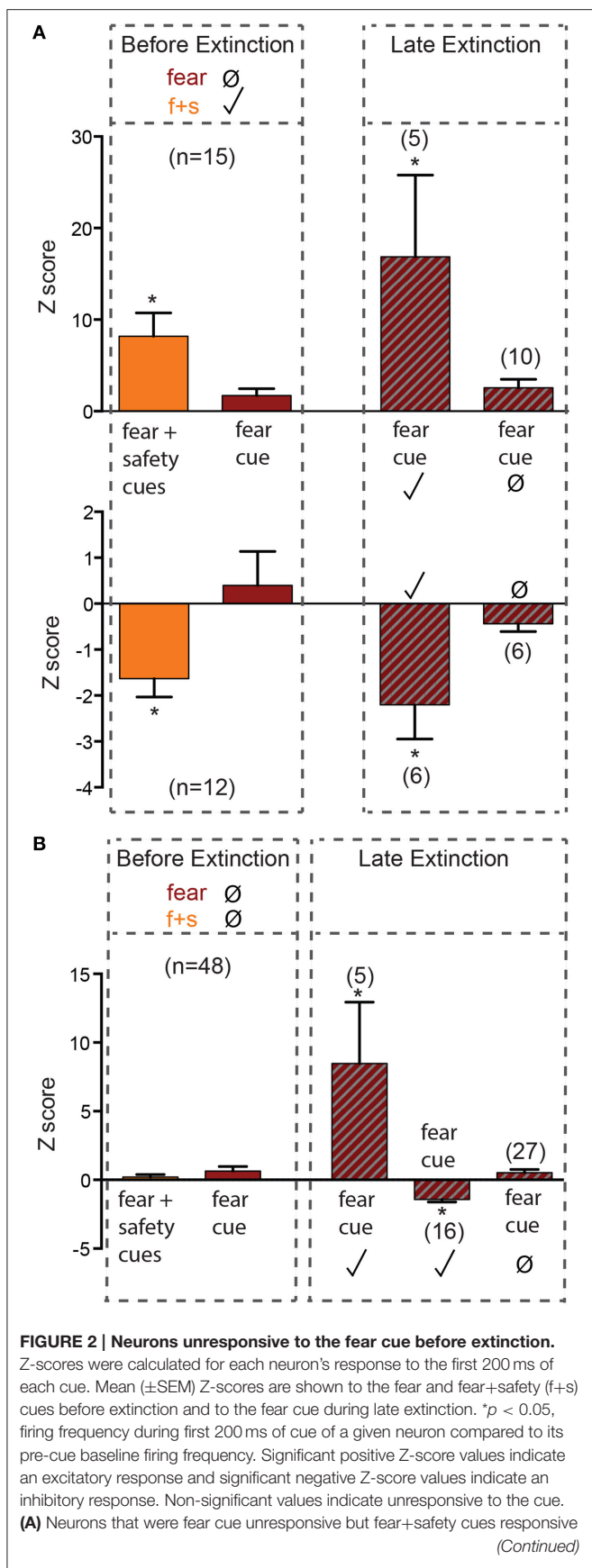
### Neurons Unresponsive to the Fear Cue Before Extinction

Neurons classified as “fear cue unresponsive” before extinction had no significant change in firing rates to the fear cue compared to baseline during DC3+4 (permutation tests,  $p > 0.05$ ). These neurons were then segregated based on whether or not they showed significant changes in firing rates to the fear+safety cue compared to baseline during DC3+4 (Figure 1E). This was done in an effort to assess if the “safety” neurons become “extinction” neurons. In other words, does one subpopulation preferentially switch to being fear cue responsive?

Before extinction, 27 neurons were fear cue unresponsive but fear+safety cue responsive (Figures 1E, 2A), showing a discriminative response to the fear+safety cue vs. fear cue. This subpopulation showed either an excitatory ( $n = 15$ , Figure 2A upper) or inhibitory ( $n = 12$ , Figure 2A lower) response to the fear+safety cue. Five of these fear cue unresponsive neurons developed an excitatory response to the fear cue in late extinction and six developed an inhibitory response (permutation tests,  $p < 0.05$ ). The remaining 16 neurons remained unresponsive to the fear cue (permutation tests,  $p > 0.05$ ).

Before extinction, 48 neurons were both fear cue and fear+safety cue unresponsive (Figures 1E, 2B). Of these 48 unresponsive neurons, five developed an excitatory response to the fear cue in late extinction and 16 developed an inhibitory response (permutation tests,  $p < 0.05$ ). The remaining 27 neurons remained unresponsive to the fear cue (permutation tests,  $p > 0.05$ ).

Together, of all the neurons that were unresponsive to the fear cue before extinction ( $n = 76$ ), 43% (32 of 76 neurons) switched to being responsive during late extinction (Figure 1E). Contrary to the hypothesis, neurons that responded to the fear+safety cue, but not the fear cue, before extinction did not appear to preferentially switch to being fear cue responsive during late extinction compared to neurons unresponsive to both cues.

**FIGURE 2 | Continued**

before extinction. Five of these fear cue unresponsive neurons developed an excitatory response to the fear cue in late extinction and six developed an inhibitory response. (B) Neurons that were both fear cue and fear+safety cue unresponsive before extinction. Of these 48 unresponsive neurons, five developed an excitatory response in late extinction and 16 developed an inhibitory response to the fear cue. The remaining 27 neurons remained unresponsive to the fear cue.

**Neurons Responsive to the Fear Cue Before Extinction**

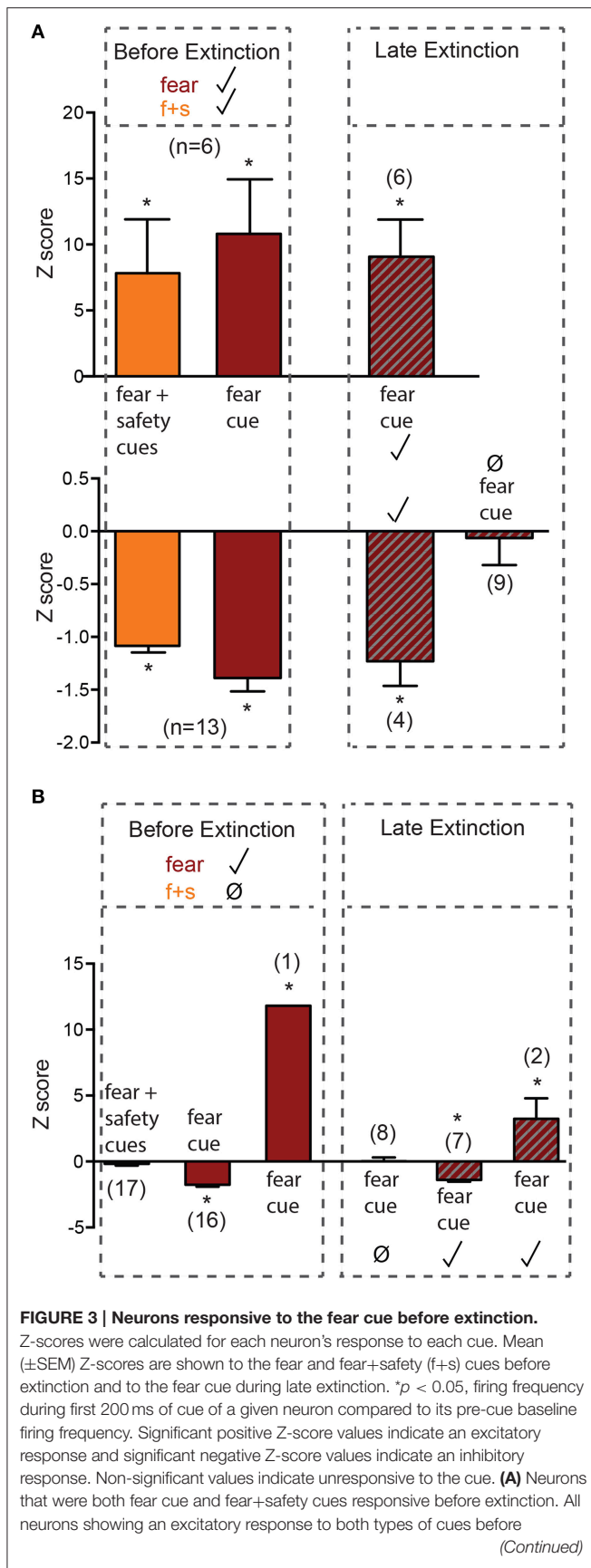
Neurons classified as “fear cue responsive” before extinction had significant increases or decreases in firing rates to the fear cue compared to baseline during DC3+4 (permutation tests,  $p < 0.05$ ). These neurons were then segregated based on whether or not they also showed significant changes in firing rates to the fear+safety cues compared to baseline during DC3+4 (Figure 1E). This was done to assess if one subpopulation preferentially switched to being fear cue unresponsive.

Before extinction, 19 neurons were both fear cue and fear+safety cue responsive (nondiscriminative; Figures 1E, 3A). This subpopulation showed either an excitatory ( $n = 6$ , Figure 3A upper) or inhibitory ( $n = 13$ , Figure 3A lower) response to the fear+safety cue. All neurons showing an excitatory response to both types of cues before extinction maintained their response through late extinction ( $n = 6$ ; permutation tests,  $p < 0.05$ ). Nine neurons showing an inhibitory response to both cues before extinction lost their inhibitory response in late extinction (permutation tests,  $p > 0.05$ ). The remaining four neurons maintained their inhibitory response through late extinction (permutation tests,  $p < 0.05$ ). That is, within this subpopulation of neurons, all excitation responses were maintained through extinction but the majority of inhibition responses were lost through extinction.

Before extinction, 17 neurons were fear cue responsive but fear+safety cue unresponsive (Figures 1E, 3B), showing a discriminative response to the fear cue vs. fear+safety cue. Only 1 neuron showed an excitatory response to the fear cue (Figure 3B) while the remaining 16 neurons showed an inhibitory response to the fear cue. Eight neurons that showed significant inhibition to the fear cue before extinction lost the inhibitory response in late extinction (permutation tests,  $p > 0.05$ ) and one neuron switched its inhibitory response to the fear cue before extinction to an excitatory response to the fear cue in late extinction. The remaining one excitatory response and seven inhibitory responses were maintained through late extinction (permutation tests,  $p < 0.05$ ).

Together, of all the neurons that were responsive to the fear cue before extinction ( $n = 36$ ), 47% (17 of 36 neurons) switched to being unresponsive during late extinction (Figure 1E).

In summary, extinction induced a gain in response to the fear cue in 43% of fear cue unresponsive neurons and a loss in response to the fear cue in 47% of fear cue responsive neurons. The number of fear cue responsive neurons before extinction was also compared to late extinction (Figure 1F) to determine if there was an overall increase or decrease in the absolute number of neurons being fear cue responsive as a result of extinction. Before

**FIGURE 3 | Continued**

extinction maintained their response through late extinction ( $n = 6$ ). Nine neurons showing an inhibitory response to both cues before extinction lost their inhibitory response in late extinction. The remaining 4 neurons maintained their inhibitory response through late extinction. (B) Neurons that were fear cue responsive but fear+safety cue unresponsive before extinction. Eight neurons that showed significant inhibition to the fear cue before extinction lost the inhibitory response in late extinction; one neuron switched its inhibitory response to the fear cue before extinction to an excitatory response to the fear cue in late extinction. The remaining one excitatory response and seven inhibitory response neurons maintained their responses through late extinction.

extinction, 75 neurons were fear cue unresponsive and 36 were fear cue responsive (Figure 1F, upper). During late extinction, 61 neurons were fear cue unresponsive and 50 were fear cue responsive (Figure 1F, lower). Thus, there was a 39% increase in the number of fear cue responsive neurons as a result of extinction. However, a Fisher's exact test revealed this increase was not significant ( $p = 0.073$ ).

## Reward Behavior and Neural Recordings of Reward Responsive Neurons

Since reward cue extinction occurred concurrently to fear cue extinction, neuronal responding to the reward cue before extinction (DC3+4) and during late extinction was also assessed.

### Discriminative Conditioning

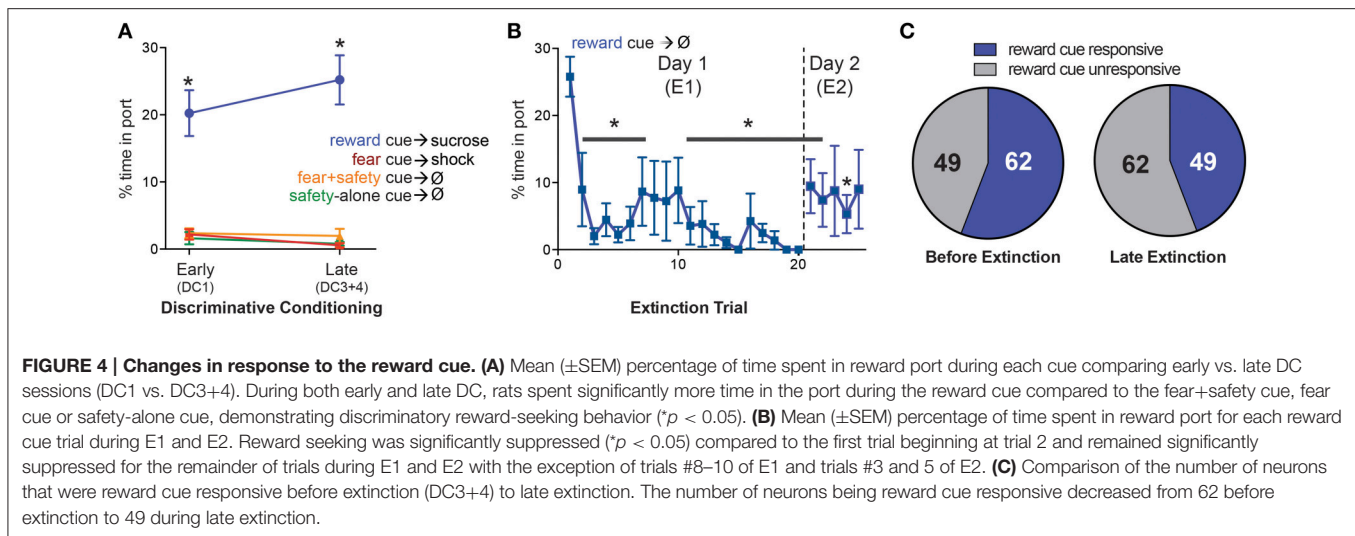
The percent time spent in the reward port during each cue was averaged across early (first DC session, DC1) and late (final two DC sessions, DC3+4) discriminative conditioning (Figure 4A). A Two-way repeated-measures ANOVA on percent time spent in port revealed a significant main effect of cue type [ $F_{(3, 39)} = 71.56, p < 0.0001$ ]. Reward seeking during the reward cue was significantly greater than the fear+safety cues, safety-alone cue, and fear cue during both early and late DC (*post-hoc* Tukey's,  $p < 0.001$  each comparison), demonstrating discriminatory reward seeking behavior by these animals.

### Reward Extinction

The percent time spent in the reward port during each reward cue trial of E1 and E2 was averaged across animals (Figure 4B). A One-way repeated-measures ANOVA on percent time spent in port revealed a main effect of trial [ $F_{(24, 336)} = 2.858, p < 0.05$ ] and Dunnett's multiple comparisons test showed reward seeking was significantly lower ( $p < 0.05$ ) during trial 2 and each subsequent trial compared to the first trial with the exception of trials #8, 9, and 10 of E1 and trials #3 and 5 of E2. Thus, compared to the first trial of E1, reward seeking was successfully extinguished by the end of E1 and maintained into E2.

### Neural Recording

Similar to the analyses completed for the fear responsive neurons, the number of reward cue responsive neurons before extinction was compared to late extinction (Figure 4C) to determine if there was an overall increase or decrease in the absolute number of neurons being reward cue responsive as a result of extinction. Before extinction, 62 neurons were reward cue



responsive and 49 were reward cue unresponsive. During late extinction, 49 neurons were fear cue responsive and 62 were fear cue unresponsive. This decrease in the number of reward cue responsive neurons as a result of extinction was not significant (Fisher's exact test,  $p > 0.05$ ).

## DISCUSSION

This study examined how neurons classified as discriminative, nondiscriminative or unresponsive during discriminative conditioning (DC), based on their responses to the fear and fear+safety cues, responded to the fear cue during extinction training and recall as fear behavior decreased. The hypothesis tested was that there is a bias for the neurons that were safety cue responsive during DC to become responsive to the fear cue as extinction progresses.

Although 41% of "safety" neurons became fear cue responsive as a result of extinction, the data revealed that there was no bias for these neurons to become preferentially responsive during fear extinction compared to the other identified subgroups. In addition to the plasticity seen in the "safety" neurons, 44% of neurons unresponsive to either the fear cue or fear+safety cue during DC became fear cue responsive during extinction. Together these emergent responses to the fear cue as a result of extinction support the hypothesis that new learning underlies extinction. The overall increase in fear cue responsive neurons in response to extinction also implies that these changes in neuronal responding during extinction are not a result of simple exposure to the sensory stimuli. If the shift were a result of repeated sensory exposures, one would expect the neurons across all groups to show decreased responding to sensory stimuli after multiple exposures as a result of sensory habituation. In contrast, 47% of neurons responsive to the fear cue during DC, regardless of its response to the fear+safety cue, became unresponsive to the fear cue during extinction. These findings are consistent with a suppression of neural responding mediated by inhibitory learning, or, potentially, by direct unlearning. Together, the data

support extinction as an active process involving both gains and losses of responses to the fear cue.

The prevalent view in the extinction field is that extinction is an active process, not a passive one (reviewed in Myers and Davis, 2002, 2007). There is ample evidence that extinction does not erase fear memories. In particular, it has been demonstrated by others (Repa et al., 2001; Herry et al., 2008; An et al., 2012), and here in this study (Figure 3A, upper), that amygdala neurons maintain increased responsiveness to the CS, even after extinction. However, there is also evidence that extinction reverses the changes induced by fear learning. For example, fear conditioning-induced potentiation is reversed with extinction in both the thalamo-lateral amygdala and cortico-lateral amygdala pathways (Kim et al., 2007; Hong et al., 2009). The data in the current study are in agreement with both views. There was both a gain of response to the fear cue (Figures 2A,B), which supports extinction as new learning, and a loss of response to the fear cue (Figure 3A, lower and Figure 3B), which may be due to unlearning.

Fear conditioning also induces synchronization at theta frequencies within the amygdala-hippocampal-prefrontal cortex (PFC) network (Sangha et al., 2009; Lesting et al., 2011). After extinction the synchronization between the amygdala and hippocampus is lost but theta synchronization is maintained between the amygdala and PFC, and between the hippocampus and PFC. A similar effect has been reported in the PFC-BA circuit in which fear extinction decreases excitatory transmission from PFC to BA while maintaining inhibitory transmission (Cho et al., 2013). These data demonstrate both reversal and maintenance of learning-induced network activity occurring in parallel during extinction.

This suggests that both new learning and unlearning mechanisms may occur in parallel during extinction. Both processes are active processes. During extinction, a new association regarding the CS is learned; i.e., a CS-no US association. And, similarly to learning the original association, long-term retention of extinction training requires both RNA and



protein synthesis across several learning paradigms and species (reviewed in Lattal et al., 2006). But, since extinction also involves reactivation of the original memory, the integrity of the original memory is vulnerable to disruption through reconsolidation mechanisms. When the original CS-US association is reactivated during extinction, it can be updated via reconsolidation mechanisms resulting in a weakening/reversal of the memory. Extinction and reconsolidation have been demonstrated to occur in parallel in the basolateral amygdala complex during reactivation of a fear memory that is no longer reinforced with shock (Duvarci et al., 2006), supporting a view that both new learning and unlearning mechanisms are at play during extinction. This view is also consistent with reports that briefly reactivating a fear memory before employing fear extinction training results in persistent attenuation of fear in both rats (Monfils et al., 2009) and humans (Schiller et al., 2010). In this case, the brief reactivation of the fear memory may induce unlearning via reconsolidation mechanisms and the extinction training results in the learning of a new CS-no US association.

The unlearning phenomena may be caused by reversal of learning-induced changes at the synapse and within the network, or it may be caused by suppression of neural responding mediated by increased inhibition. Several neurons reported here had decreased firing rates in response to the fear cue during extinction. It is not clear what the source of cue-evoked inhibition, nor its downstream effects, might be. However, it has been shown that the balance between excitation and inhibition in the PFC-BA pathway is shifted toward inhibition after extinction (Cho et al., 2013), suggesting that the upstream source for the inhibitions seen in the data presented here may be the PFC. This would be consistent with the requirement of the infralimbic region of the prefrontal cortex to discriminate between the fear and fear+safety cues in this task (Sangha et al., 2014), and to

successfully recall fear extinction (Quirk et al., 2000; Laurent and Westbrook, 2009; Chang and Maren, 2010; Fontanez-Nuin et al., 2011; Sierra-Mercado et al., 2011; Santini et al., 2012; Sangha et al., 2014, but see Do Monte et al., 2015).

In summary, the data implicate multiple levels of plasticity in response to fear extinction that most likely interact with multiple microcircuits within the BA. It also indicates that there may be a general remapping of these neuronal microcircuits within the BA in response to extinction. Together it suggests the final output of the integrated BA circuit to influence fear behavior is a balance of excitation and inhibition, and perhaps reversal of learning-induced changes. Further exploration of the intricacies of upregulating or downregulating these BA microcircuits on downstream targets and their effects on fear behavior will lead to greater understanding of the mechanisms contributing to successful fear inhibition which is compromised in individuals suffering from PTSD and similar disorders.

## AUTHOR CONTRIBUTIONS

SS: experimental design, performed research, analyzed data, wrote manuscript.

## ACKNOWLEDGMENTS

This work was supported by funding from the State of California for Medical Research on Alcohol and Substance Abuse through the University of California at San Francisco and NIH grant AA014925 to Dr. P. H. Janak. The author greatly appreciates Dr. P. H. Janak for discussion and comments on the manuscript. The author also thanks R. Reese and I. Grossrubatscher for technical assistance, Dr. T. M. Gill for programming and recording guidance and Dr. J. Z. Chadick for programming assistance.

## REFERENCES

- An, B., Hong, I., and Choi, S. (2012). Long-term neural correlates of reversible fear learning in the lateral amygdala. *J. Neurosci.* 32, 16845–16856. doi: 10.1523/JNEUROSCI.3017-12.2012
- Balleine, B. W., and Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci.* 29, 272–279. doi: 10.1016/j.tins.2006.03.002
- Belova, M. A., Paton, J. J., Morrison, S. E., and Salzman, C. D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* 55, 970–984. doi: 10.1016/j.neuron.2007.08.004
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57, 289–300.
- Blanchard, R. J., and Blanchard, D. C. (1969). Crouching as an index of fear. *J. Comp. Physiol. Psychol.* 67, 370–375. doi: 10.1037/h0026779
- Chang, C. H., and Maren, S. (2010). Strain difference in the effect of infralimbic cortex lesions on fear extinction in rats. *Behav. Neurosci.* 124, 391–397. doi: 10.1037/a0019479
- Cho, J. H., Deisseroth, K., and Bolshakov, V. Y. (2013). Synaptic encoding of fear extinction in mPFC-amygdala circuits. *Neuron* 80, 1491–1507. doi: 10.1016/j.neuron.2013.09.025
- Christianson, J. P., Fernando, A. B., Kazama, A. M., Jovanovic, T., Ostroff, L. E., and Sangha, S. (2012). Inhibition of fear by learned safety signals: a mini-symposium review. *J. Neurosci.* 32, 14118–14124. doi: 10.1523/JNEUROSCI.3340-12.2012
- Corbit, L. H., and Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrument transfer. *J. Neurosci.* 25, 962–970. doi: 10.1523/JNEUROSCI.4507-04.2005
- Davies, D. L., and Bouldin, D. W. (1979). A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* 1, 224–227. doi: 10.1109/TPAMI.1979.4766909
- Do Monte, F. H., Manzano-Nieves, G., Quinones-Laracuate, K., Ramos-Medina, L., and Quirk, G. J. (2015). Revisiting the role of infralimbic cortex in fear extinction with optogenetics. *J. Neurosci.* 35, 3607–3615. doi: 10.1523/JNEUROSCI.3137-14.2015
- Duvarci, S., ben Mamou, C., and Nader, K. (2006). Extinction is not a sufficient condition to prevent fear memories from undergoing reconsolidation in the basolateral amygdala. *Eur. J. Neurosci.* 24, 249–260. doi: 10.1111/j.1460-9568.2006.04907.x
- Fendt, M., and Fanselow, M. S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neurosci. Biobehav. Rev.* 23, 743–760. doi: 10.1016/S0149-7634(99)00016-0
- Fontanez-Nuin, D. E., Santini, E., Quirk, G. J., and Porter, J. T. (2011). Memory for fear extinction requires mGluR5-mediated activation of infralimbic neurons. *Cereb. Cortex* 21, 727–735. doi: 10.1093/cercor/bhq147
- Heldt, S. A., and Ressler, K. J. (2007). Training-induced changes in the expression of GABA-associated genes in the amygdala after the acquisition

- and extinction of Pavlovian fear. *Eur. J. Neurosci.* 26, 3631–3644. doi: 10.1111/j.1460-9568.2007.05970.x
- Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., and Lüthi, A. (2008). Switching on and off fear by distinct neuronal circuits. *Nature* 454, 600–606. doi: 10.1038/nature07166
- Hesterberg, T., Moore, D. S., Monaghan, S., Clipson, A., and Epstein, R. (2005). “Bootstrap methods and permutation tests,” in *Introduction to the Practice of Statistics, 5th Edn.*, eds D. S. Moore and G. P. McCabe (New York, NY: WH Freeman & Company), 14.1–14.70.
- Hong, I., Song, B., Lee, S., Kim, J., Kim, J., and Choi, S. (2009). Extinction of cued fear memory involves a distinct form of depotentiation at cortical input synapses onto the lateral amygdala. *Eur. J. Neurosci.* 30, 2089–2099. doi: 10.1111/j.1460-9568.2009.07004.x
- Jackson, A., and Fetz, E. E. (2007). Compact movable microwire array for long-term chronic unit recording in cerebral cortex of primates. *J. Neurophysiol.* 98, 3109–3118. doi: 10.1152/jn.00569.2007
- Janak, P. H. (2002). “Multichannel neural ensemble recording during alcohol self administration methods” in *Alcohol-Related Neuroscience Research*, eds Y. Liu and D. M. Lovinger (Boca Raton, FL: CRC Press), 241–258. doi: 10.1201/9781420042092.ch10
- Janak, P. H., and Tye, K. M. (2015). From circuits to behaviour in the amygdala. *Nature* 517, 284–292. doi: 10.1038/nature14188
- Jovanovic, T., Kazama, A., Bachevalier, J., and Davis, M. (2012). Impaired safety signal learning may be a biomarker of PTSD. *Neuropharmacology* 62, 695–704. doi: 10.1016/j.neuropharm.2011.02.023
- Kim, J., Lee, S., Park, K., Hong, I., Song, B., Son, G., et al. (2007). Amygdala depotentiation and fear extinction. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20955–20960. doi: 10.1073/pnas.0710548105
- Lattal, K. M., Radulovic, J., and Lukowiak, K. (2006). Extinction: does it or doesn't it? The requirement of altered gene activity and new protein synthesis. *Biol. Psychiatry* 60, 344–351. doi: 10.1016/j.biopsych.2006.05.038
- Laurent, V., and Westbrook, R. F. (2009). Inactivation of the infralimbic but not the prelimbic cortex impairs consolidation and retrieval of fear extinction. *Learn. Mem.* 16, 520–529. doi: 10.1101/lm.1474609
- Lesting, J., Narayanan, R. T., Kluge, C., Sangha, S., Seidenbecher, T., and Pape, H. C. (2011). Patterns of coupled theta activity in amygdala-hippocampal-prefrontal cortical circuits during fear extinction. *PLoS ONE* 6:e21714. doi: 10.1371/journal.pone.0021714
- Likhtik, E., Pelletier, J. G., Popescu, A. T., and Paré, D. (2006). Identification of basolateral amygdala projection cells and interneurons using extracellular recordings. *J. Neurophysiol.* 96, 3257–3265. doi: 10.1152/jn.00577.2006
- Lin, H. C., Mao, S. C., and Gean, P. W. (2009). Block of gamma-aminobutyric acid-A receptor insertion in the amygdala impairs extinction of conditioned fear. *Biol. Psychiatry* 66, 665–673. doi: 10.1016/j.biopsych.2009.04.003
- Málková, L., Gaffan, D., and Murray, E. A. (1997). Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforce devaluation effects in rhesus monkeys. *J. Neurosci.* 17, 6011–6020.
- Monfils, M. H., Cowansage, K. K., Klann, E., and LeDoux, J. E. (2009). Extinction-reconsolidation boundaries: key to persistent attenuation of fear memories. *Science* 324, 951–955. doi: 10.1126/science.1167975
- Myers, K. M., and Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron* 36, 567–584. doi: 10.1016/S0896-6273(02)01064-4
- Myers, K. M., and Davis, M. (2007). Mechanisms of fear extinction. *Mol. Psychiatry* 12, 120–150. doi: 10.1038/sj.mp.4001939
- Nicolelis, M. A., Dimitrov, D., Carmena, J. M., Crist, R., Lehew, G., Kralik, J. D., et al. (2003). Chronic, multisite, multielectrode recordings in macaque monkeys. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11041–11046. doi: 10.1073/pnas.1934665100
- Paton, J. J., Belova, M. A., Morrison, S. E., and Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870. doi: 10.1038/nature04490
- Paxinos, G., and Watson, C. (2007). “The rat brain,” in *Stereotaxic Coordinates, 6th Edn.*, eds G. Paxinos and C. Watson (Amsterdam: Elsevier), 93–100.
- Quirk, G. J., Russo, G. K., Barron, J. L., and Lebron, K. (2000). The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *J. Neurosci.* 20, 6225–6231.
- Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., and LeDoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nat. Neurosci.* 4, 724–731. doi: 10.1038/89512
- Sangha, S., Chadick, J. Z., and Janak, P. H. (2013). Safety encoding in the basal amygdala. *J. Neurosci.* 33, 3744–3751. doi: 10.1523/JNEUROSCI.3302-12.2013
- Sangha, S., Ilenseer, J., Sosulina, L., Lesting, J., and Pape, H. C. (2012). Differential regulation of glutamic acid decarboxylase gene expression after extinction of a recent memory vs. intermediate memory. *Learn. Mem.* 19, 194–200. doi: 10.1101/lm.025874.112
- Sangha, S., Narayanan, R. T., Bergado-Acosta, J. R., Stork, O., Seidenbecher, T., and Pape, H. C. (2009). Deficiency of the 65kDa isoform of glutamic acid decarboxylase impairs extinction of cued but not contextual fear memory. *J. Neurosci.* 29, 15713–15720. doi: 10.1523/JNEUROSCI.2620-09.2009
- Sangha, S., Robinson, P. D., Greba, Q., Davies, D. A., and Howland, J. G. (2014). Alterations in reward, fear and safety cue discrimination after inactivation of the rat prelimbic and infralimbic cortices. *Neuropsychopharmacology* 39, 2405–2413. doi: 10.1038/npp.2014.89
- Santini, E., Sepulveda-Orengo, M., and Porter, J. T. (2012). Muscarinic receptors modulate the intrinsic excitability of infralimbic neurons and consolidation of fear extinction. *Neuropsychopharmacology* 37, 2047–2056. doi: 10.1038/npp.2012.52
- Schiller, D., Monfils, M. H., Raio, C. M., Johnson, D. C., Ledoux, J. E., and Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature* 463, 49–53. doi: 10.1038/nature08637
- Schoenbaum, G., Chiba, A. A., and Gallagher, M. (1999). Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *J. Neurosci.* 19, 1876–1884.
- Shabel, S. J., and Janak, P. H. (2009). Substantial similarity in amygdala neural activity during conditioned appetitive and aversive emotional arousal. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15031–15036. doi: 10.1073/pnas.0905580106
- Sierra-Mercado, D., Padilla-Coreano, N., and Quirk, G. J. (2011). Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. *Neuropsychopharmacology* 36, 529–538. doi: 10.1038/npp.2010.184

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer Stefan Herlitze and handling Editor Denise Manahan-Vaughan declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2015 Sangha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.