



Encoding of conditioned inhibitors of fear in the infralimbic cortex

Ka H. Ng ¹, Susan Sangha ^{2,*}

¹Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, United States,

²Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN 46202, United States

*Corresponding author: Department of Psychiatry, Indiana University School of Medicine, 320 W. 15th Street, Indianapolis, IN 46202, United States.
Email: susangha@iu.edu

Cues in the environment signaling the absence of threat, i.e. safety, can influence both fear and reward-seeking behaviors. Heightened and maladaptive fear is associated with reduced activity in the medial prefrontal cortex. We have previously shown in male rats that the infralimbic (IL) prefrontal cortex is necessary for suppressing fear during a safety cue. The objective of the present study was to determine if there was safety cue-specific neural activity within the IL using a Pavlovian conditioning paradigm, where a fear cue was paired with shock, a safety cue was paired with no shock, and a reward cue was paired with sucrose. To investigate how safety cues can suppress fear, the fear and safety cues were presented together as a compound fear + safety cue. Single-unit activity showed a large proportion of neurons with excitatory responses to the fear + safety cue specifically, a separate group of neurons with excitatory responses to both the reward and fear + safety cues, and bidirectional neurons with excitation to the fear + safety cue and inhibition to the fear cue. Neural activity was also found to be negatively correlated with freezing during the fear + safety cue. Together, these data implicate the IL in encoding specific aspects of conditioned inhibitors when fear is being actively suppressed.

Key words: electrophysiology; fear discrimination; prefrontal cortex; safety.

Introduction

Accurate discrimination among cues predicting threat, safety, or reward appropriately guides motivated behaviors. Cues that were associated with threat, or are reminders of the threat, may elicit robust fear behaviors that persist maladaptively. Posttraumatic stress disorder represents a condition in which maladaptive fear persists through extinction (Morriss et al. 2015) and in response to explicit safety cues (Jovanovic et al. 2010). Learned safety cues are able to act as conditioned inhibitors such that, when presented in compound with a learned threat cue, it can inhibit fear responding (Bouton 2007; Christianson et al. 2012; Sangha et al. 2020). Thus, fear extinction and safety conditioning represent 2 methods to regulate learned fear.

Research on fear regulation has typically focused on a circuit comprising the medial prefrontal cortex (mPFC), amygdala, and hippocampus in both rodent and human studies (Sangha et al. 2020). Despite this translational relevance, the mechanisms of regulating fear in safe conditions remain unclear and the neural circuits continue to be mapped (Sangha et al. 2020). Reduced activity in the mPFC has been associated with heightened fear that is resistant to extinction in both humans and rodents (reviewed in (Patel et al. 2012; Goode et al. 2018)). In rodents, 2 subregions within the mPFC have been consistently linked with fear behavior, the infralimbic (IL) and prelimbic (PL) prefrontal cortical regions. More specifically, inactivation of the IL impairs the consolidation and expression of fear extinction memories (Laurent and Westbrook 2009; Sierra-Mercado et al. 2011; Sangha, Robinson, et al. 2014; Bukalo et al. 2015). We have previously shown that specifically inactivating the IL, and not the PL, impaired the ability of male rats to suppress fear in the presence

of a learned safety cue (Sangha, Robinson, et al. 2014) using a Pavlovian conditioning paradigm that we developed and validated, which consisted of a fear cue associated with shock, a safety cue with no shock, and a reward cue with sucrose. To investigate how safety cues can suppress fear elicited by cues associated with shock, we presented the fear and safety cues together as a compound cue. Our previous studies (Sangha et al. 2013; Sangha, Greba, et al. 2014; Sangha, Robinson, et al. 2014; Müller et al. 2018; Ng et al. 2018; Greiner et al. 2019; Woon et al. 2020) have shown male Long Evans rats readily learn to suppress fear responding to a learned fear cue if in the presence of a safety cue.

Despite the reduced fear levels in response to extinction and conditioned inhibition, the mechanisms mediating each are not the same. For example, we have recently shown that prior stress impaired fear extinction to a much greater extent than conditioned inhibition within the same animal, indicating that good extinction learning and conditioned inhibition do not always occur together (Woon et al. 2020). There have also been suggestions that using safety cues during exposure therapy, i.e. extinction, could potentially enhance treatment outcomes in reducing fear and anxiety (Odrizola and Gee 2021). Investigating explicit safety cues in conjunction with fear cues offers some unique advantages over typical fear extinction studies. The majority of fear extinction studies use a single cue throughout where a cue is first associated with threat and then later is no longer associated with threat. These studies do not typically assess additional cues in parallel, making it sometimes difficult to attribute how fear may be differentially regulated to other cues in the environment. By presenting safety, fear and reward cues within the same

sessions, we are better able to hone in on “discriminating” safety from fear and reward. Moreover, how does the brain encode the conflicting scenario of both a fear and safety cue presented at the same time?

The objective of the present study was to determine if there was safety cue-specific activity within the IL and how it may encode the conflicting fear + safety compound cue. Electrophysiological activity was collected in freely behaving rats as they were learning about safety, fear, and reward cues. Overall, we observed a large proportion of neurons with excitatory responses to the fear + safety cue specifically, a separate group of neurons with excitatory responses to both the reward and fear + safety cues, and bidirectional neurons with excitation to the fear + safety cue and inhibition to the fear cue.

Materials and methods

Subjects

Fifteen male Long Evans rats (Blue Spruce; Envigo, Indianapolis) weighing 300–350 g were single-housed under a 12-h light/dark cycle (lights on 09:00) and were handled for 1 week before commencing experiments. We have previously shown that female rats did not show fear suppression during the safety cue (Greiner et al. 2019), thus it was not possible to utilize a mixed-sex design. Since we have demonstrated the IL is necessary for fear suppression during the safety cue in this paradigm in male rats (Sangha, Robinson, et al. 2014), we elected to only include male subjects in the current study. All procedures were performed during the light cycle and were approved by the Purdue Animal Care and Use Committee. Rats had ad libitum access to food and water up until the first training session at which point they received 20–22 g of food per day after their daily training session for the remainder of the experiment.

Apparatus

The training chamber was a Med Associates Plexiglas box (28 cm length × 21 cm width × 35 cm height) encased in an electrically shielded sound-attenuating chamber (Med Associates, ST Albans, VT). The 10% liquid sucrose (100 μ L) was delivered through a recessed port located in the center of 1 wall, containing an infrared beam for detecting port entries and exits. There were 2 lights (28 V, 100 mA), 1 on each side of the port for delivering the 20-s continuous light cue, and a house light (28 V, 100 mA) located at the top of the wall opposite to the port for providing constant background illumination. Next to the house light was a “tweeter” speaker (ENV-224BM) for delivering auditory cues. Footshocks were delivered through the grid floor by a constant current aversive stimulator (ENV-414S). A side-view video camera located on the door of the sound-attenuating chamber recorded the rat’s behavior for offline video analyses.

Discriminative conditioning behavioral training procedure

Three stimuli were used as cues: a 20-s continuous 3-kHz tone (70 dB), a 20-s pulsing 11-kHz tone (200 ms on, 200 ms off; 70 dB), and a 20-s continuous light (28 V, 100 mA). The 20-s continuous 3-kHz tone (70 dB) was reserved for the reward cue for all animals. The remaining 2 cues were counterbalanced for fear and safety cues between 2 groups of animals ($n = 8$ had the light as the safety cue, $n = 7$ had the light as the fear cue).

Animals first received 5 sessions of reward training distributed across 5 days (R1–R5). Each session consisted of 25 pairings (ITI, 90–130 s) of the reward cue with a 3-s delivery of 10% liquid

sucrose (100 μ L pseudorandomly presented 10–20s after reward cue onset) into a port. Animals then received 1 session of habituation (HAB) training, which consisted of 25 trials of the reward cue paired with liquid sucrose (100 μ L pseudorandomly presented 10–20s after reward cue onset), 5 trials of the future fear cue presented alone, and 5 trials of the future safety cue presented alone (ITI, 90–130 s). This HAB procedure has been used in this task to reduce any baseline freezing that may be present to the novel cues with the number of trials presented not being sufficient to produce latent inhibition (Sangha et al. 2013). Animals then received 4 sessions of discriminative conditioning (DC1–4) across 4 days; i.e. 1 session per day for 4 days. Each session consisted of 15 trials of the reward cue paired with liquid sucrose (100 μ L pseudorandomly presented 10–20s after reward cue onset), 4 trials of the fear cue paired with footshock (0.5 s, 0.45 mA at cue offset), 15 trials of the safety cue and fear cue presented concurrently without footshock, and 10 trials of the safety cue presented alone without footshock (44 trials total, ITI 100–140 s).

In vivo electrophysiology

All electrode arrays were composed of 50- μ m stainless steel wires in a 2-by-4 arrangement and were spaced 250 μ m apart (Neuro-Biological Laboratories, Dallas, TX). During surgery, 2 sets of electrode arrays were implanted bilaterally to the IL (AP = +2.4 mm; ML = \pm 0.5 mm; DV = -3.9 mm; 8 wires per side), while the rats were deeply anesthetized with isoflurane. The animals had 7–10 days of surgical recovery with ad libitum access to food and water during recovery.

During training, the implanted electrode arrays were connected to a suspended headstage cable. The recorded neural activity passed through a headstage amplifier, a commutator, and a programmable amplifier (Plexon, Dallas, TX) to be amplified and filtered (0.4 and 5 kHz). Spike threshold was set to 2.5 standard deviations (SDs) from the mean of peak amplitude distribution that was customized to the individual channel before the start of each session (Omniplex and PlexControl, Plexon). Isolation of single-unit activity from background noise was performed manually with an offline multichannel spike sorter (Offline Sorter; Plexon) using peak amplitude, valley amplitude, and principal component analysis on waveform shapes.

Histology

After recordings were completed, rats were deeply anesthetized with sodium pentobarbital. A 15-s 20 μ A current was passed through each wire to mark each electrode tip. Rats were then perfused with PBS and 10% formalin containing 3% potassium ferrocyanide. Brains were soaked in 30% sucrose-containing formalin and were cryo-sectioned at 50 μ m. Sections were stained with cresyl violet and were examined under a light microscope to verify placements. Each wire was then matched to its corresponding track and lesion within the brain sections compared to the spatial configuration of the electrode array and was matched to its specific recording channel. Only wires confirmed to be in the IL were included in the analyses (Fig. 1).

Experimental design and statistical analysis

Behavioral analyses

Fear behavior was assessed manually offline from videos by measuring freezing, defined as complete immobility with the exception of respiratory movement, which is an innate defensive behavior (Blanchard and Blanchard 1969; Fendt and Fanselow 1999). The amount of time spent freezing within a 20-s interval

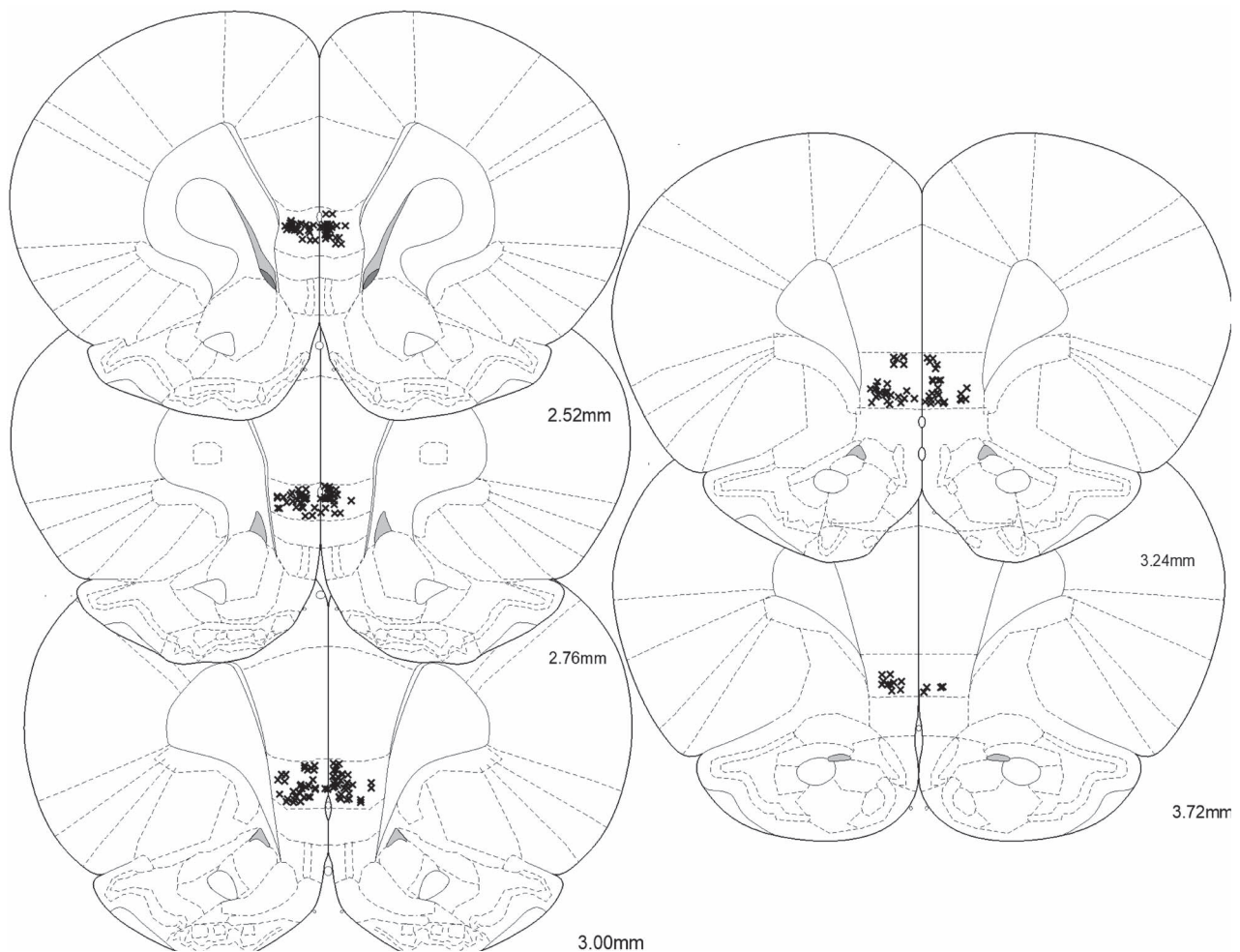


Fig. 1. Locations of each electrode tip from 15 rats verified to be in the IL.

during cue presentation was quantified and expressed as percentages. Reward behavior was assessed manually by quantifying the amount of time the animals spent inside the port or having their nose positioned at the port entrance and was expressed as percentages. The person performing the manual behavioral scoring had a Pearson's correlation of at least $r = 0.8$ with other scorers in the same laboratory for freezing and reward behaviors. The behavioral data were analyzed with 2-way repeated-measures ANOVAs with post hoc Dunnett's multiple comparisons in GraphPad Prism. Freezing to the fear cue was compared (Dunnett's) to each other cue and reward-seeking to the reward cue was compared (Dunnett's) to each other cue.

Single-unit analyses

Single-unit signals were isolated using Plexon's Offline Sorter, and peri-event histograms were generated for each of the 4 cues; 10,000 round permutation analyses were used to compare within each cue the averaged count per bin (50 ms/bin) during the 400-ms precue baseline epoch against the 400-ms postcue onset epoch to determine if there was a significant response to a given cue as we have done previously (Sangha et al. 2013; Sangha 2015). Units were categorized based on significant cue-related responses regardless of the expressed behavior.

For each isolated neuron, firing frequency was then Z-scored trial by trial by subtracting the averaged precue baseline across 400-ms bins spanning 2 s from the averaged spikes/s for each

400-ms bin and dividing by the SD across the 400-ms bins spanning the 2-s precue baseline. These values were then averaged per 400-ms bin across trials for a given cue type for the 2 s before and after cue onset. For each category identified, Z-scores for each neuron were combined, followed by a 1-way ANOVA to compare neural responding across cues during the 0–400 ms bin after cue onset. Post hoc Tukey's multiple comparisons test were used where appropriate to determine which cues were significantly different ($P < 0.05$). For the analysis of neural activity with freezing behavior, we calculated Pearson correlations between the Z-scored firing frequency during the 0–400-ms postcue onset with the percentage time spent freezing across each 20-s FS cue in DC4. This amounted to 15 trials and each trial was included in the analyses as separate data points.

Results

All rats ($n = 15$) underwent surgical implantation of electrode arrays into the IL before any behavioral training. Only rats with verified wire placements in the IL were included in this study; a total of 209 single units were isolated from these wires (Fig. 1). After surgical recovery, all rats received 5 sessions of reward conditioning (reward cue + sucrose), 1 HAB session (reward cue + sucrose, fear cue alone, and safety cue alone), and 4 sessions of DC (reward cue + sucrose, fear cue + shock, safety cue alone, and fear + safety cue with no shock).

cue, safety cue, and reward cue) across the 4 DC sessions (Fig. 2A). There was a significant cue by session interaction ($F(9,168)=30.40$, $P < 0.0001$) as well as main effects of cue ($F(3, 56)=112.5$, $P < 0.0001$) and session ($F(2.77, 155.2)=50.04$, $P < 0.0001$). Post hoc Dunnett's multiple comparisons to the fear cue showed that animals had significantly higher levels of freezing to the fear cue than the reward cue and safety cue from DC1 to DC4 ($P < 0.05$ to $P < 0.0001$), indicating that the animals learned to freeze to the fear cue beginning at DC1. Animals had significantly higher levels of freezing to the fear cue than the fear + safety cue from DC2 to DC4 ($P < 0.05$ to $P < 0.0001$), indicating that the animals learned to suppress freezing to the fear + safety cue beginning at DC2.

A 2-way repeated-measures ANOVA was performed to assess the percent time spent at the port to the same 4 cues across the same 4 DC sessions (Fig. 2B). There was no significant cue by session interaction ($F(9,168)=1.36$, $P = 0.21$) or main effect of session ($F(2.72, 152.5)=0.14$, $P = 0.92$). There was a significant main effect of cue ($F(3, 56)=43.98$, $P < 0.0001$) in which there was more time spent at the port during the reward cue overall. These data indicate that animals spent more time reward-seeking during the reward cue beginning at DC1, which remained stable throughout DC1–DC4.

Single units within the IL showed both cue-specific and cue-overlapping responses

Across the 15 rats, a total of 209 single units were isolated across both DC3 and DC4. Since these 209 units showed similar waveforms, firing frequency, and PCA of waveform shape across DC3 and DC4 within animal, we combined the trials from these 2 sessions for these 209 units to increase power for the number of trials per cue type. Tracking the same units from DC1 all the way to DC4 was not consistent, and thus we focused the following analyses to DC3 and DC4. The median firing rate across a 2-s precue baseline across all neurons was 9.53 Hz, with 1 unit at a baseline of 47.64 Hz, and the remaining 208 units with a baseline of <17 Hz, indicating a mix of neuron types. Of these 209 units, 46 single units did not show a significant response to any of the cues presented (unresponsive), while the remaining 163 showed either a significant increase or decrease in firing rate within the first 400 ms after cue onset to at least 1 cue type (permutation test, $P < 0.05$; Fig. 3A). The majority, 57% (119 neurons), showed a significant excitation to at least 1 cue, followed by 14% showing significant inhibition (30 neurons) and 7% showing mixed excitation and inhibition responses depending on the cue (14 neurons). Of the neurons showing a significant response, 64 were selective to 1 cue specifically (cue-specific neurons; Fig. 3B), 77 showed an overlapping response to >1 cue (cue-overlapping neurons; Fig. 3C), and 8 showed bidirectional responses to the fear + safety cue and fear cue (bidirectional neurons; Fig. 3D). Across all categories, the dominant response was excitation to the fear + safety (FS) cue, whether it was classified as cue-specific, cue-overlapping, or bidirectional.

Cue-specific neurons in the IL

Separate groups of neurons showing a specific excitation or inhibition to just 1 cue type were observed for all cue types (permutation test, $P < 0.05$; $n = 64$ neurons; Fig. 3B). The largest proportion (41%) showed a specific excitation to just the fear + safety (FS) cue and not the safety (S), fear (F), or reward (R) cues. The Z-scored neural activity was combined for neurons for each subcategory (Fig. 4) and is described in more detail below. For each subcategory, a 1-way ANOVA was used to compare neural responding across

cues during the 0–400-ms bin after cue onset followed by post hoc Tukey's multiple comparisons tests to determine which cues were significantly different (Table 1).

FS-specific neurons

Compared to other cue-specific categories, neurons showing an excitation to the FS cue specifically made up the largest subcategory ($n = 26$; Fig. 4A, left). Compared to the 400-ms precue baseline, these neurons showed a significant increase in firing rate in the 400-ms postcue to only the FS cue and not the S, F, or R cues. Comparing across cues during the 400-ms postcue period, the Z-scored neural activity was significantly higher during the FS cue compared to all other cues (Table 1). In addition, there was 1 neuron that showed a FS-specific inhibition during the 400-ms postcue (Fig. 4A, right).

S-specific neurons

For neurons classified as S-specific, there was either a significant increase ($n = 4$; Fig. 4B, left) or decrease ($n = 5$; Fig. 4B, right) in firing rate in the 400-ms postcue compared to the 400-ms precue baseline to only the S cue and not the FS, F, or R cues. For those showing increased firing, a 1-way ANOVA comparing the Z-scored activity across cues was significant, but the post hoc analyses yielded no significant differences between cues (Table 1). For those showing a significant inhibition postcue, a 1-way ANOVA across cues was significant, and post hoc analyses showed Z-scored neural activity was significantly lower during the S cue compared to the F cue but not other cues (Table 1).

F-specific neurons

Compared to the 400-ms precue baseline, these neurons showed a significant increase ($n = 8$; Fig. 4C, left) or decrease ($n = 9$; Fig. 4C, right) in firing rate in the 400-ms postcue to only the F cue and not the FS, S, or R cues. For those showing excitation, the Z-scored neural activity during the 400-ms postcue period was significantly higher during the F cue compared to all other cues. Similarly, the activity was significantly lower for those showing inhibition during the F cue compared to all other cues (Table 1).

R-specific neurons

For neurons classified as R-specific, there was either a significant increase ($n = 6$; Fig. 4D, left) or decrease ($n = 5$; Fig. 4D, right) in firing rate in the 400-ms postcue compared to the 400-ms precue baseline to only the R cue and not the FS, S, or F cues. For those showing increased firing, a 1-way ANOVA comparing activity across cues was significant, and post hoc analyses showed that activity was significantly higher during the R cue compared to the S cue but not other cues (Table 1). The 1-way ANOVA for neurons with a significant inhibition was not significant (Table 1).

Cue-overlapping neurons in the IL

Since the objective of this study was to correlate neural activity within the IL to conditioned inhibition specifically, we focus our description and analyses of cue-overlapping neurons on neurons that showed either an excitation or inhibition to the FS cue coupled with a similar excitation/inhibition to another cue (permutation test, $P < 0.05$). That is, neurons that may have shown an excitation to the F and R cues, for example, are not included here. The Z-scored neural activity was combined for neurons for each subcategory (Fig. 5) and is described in more detail below. For each subcategory, a 1-way ANOVA was used to compare neural responding across cues during the 0–400-ms bin after cue

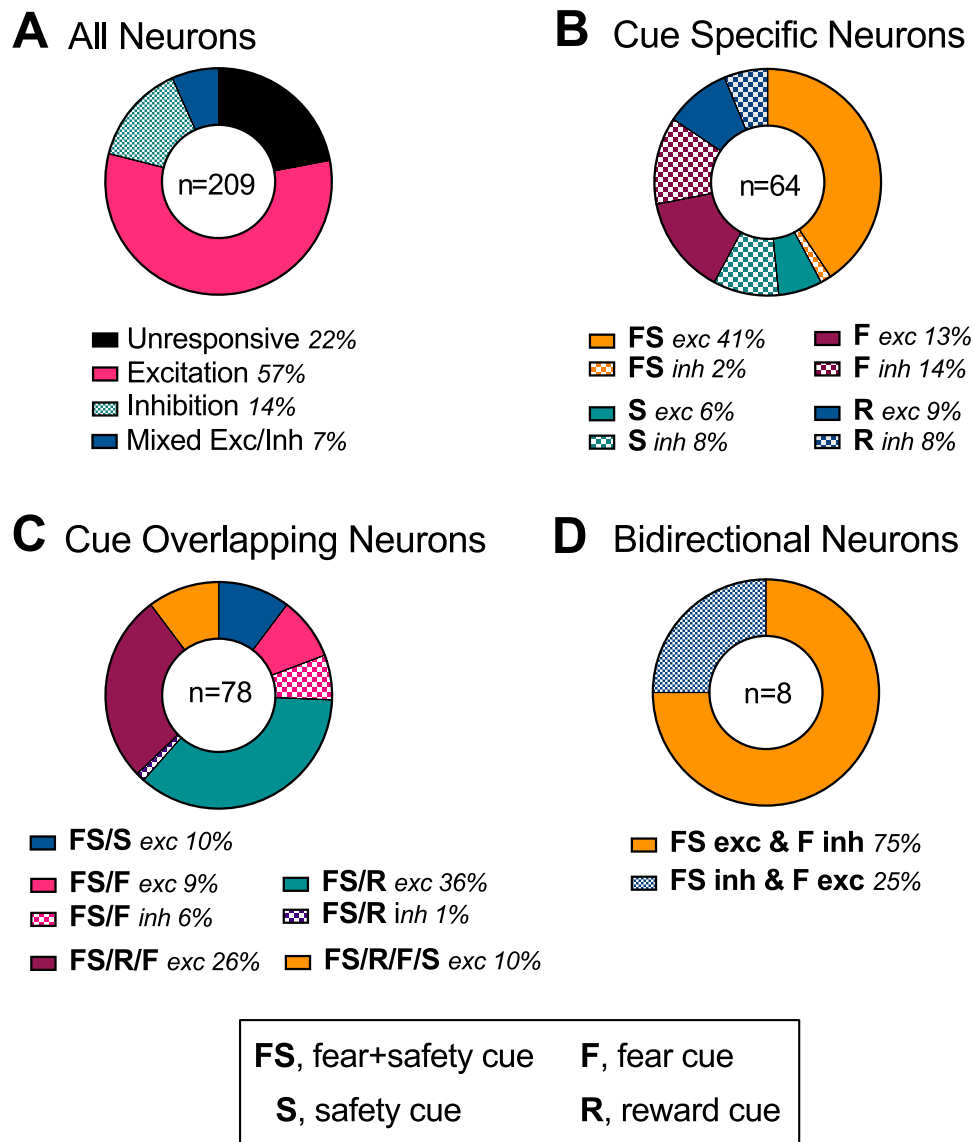


Fig. 3. Categorization of neurons. A) All neurons. A total of 209 neurons were isolated across DC3 and DC4. Unresponsive neurons did not show a significant response to any of the cues (22%); 57% of neurons showed a significant excitation to at least one cue, followed by 14% showing significant inhibition, and 7% showing mixed excitation and inhibition responses depending on the cue. B) Cue-specific neurons. A total of 64 neurons were selective to 1 cue specifically, with either a significant excitation (exc) or inhibition (inh). Neurons showing excitation to only the fear + safety cue (FS) made up the largest subpopulation (41%). C) Cue-overlapping neurons; 78 neurons showed a similar excitation (exc) or inhibition (inh) to >1 cue. Neurons showing excitation to the fear + safety cue and reward cue (FS/R) made up the largest subpopulation (36%). D) Bidirectional neurons; 8 neurons showed opposing responses to the fear + safety versus fear cues. Most of these showed an excitation to the fear + safety cue and inhibition to the fear cue (FS exc & F inh).

onset followed by post hoc Tukey's multiple comparisons tests to determine which cues were significantly different (Table 2).

FS/S-overlapping neurons

All neurons within this category showed a significant excitation to both the FS and S cues and not the F or R cues ($n = 8$; Fig. 5A). There were no neurons showing a significant inhibition to only the FS and S cues. Comparing across cues during the 400-ms postcue period, the Z-scored neural activity was significantly higher during the FS cue compared to all other cues, including the S cue (Table 2). Activity during the S cue was significantly higher than the F and R cues. That is, even though there was significant excitation to the FS and S cues compared to their precue baselines, the magnitude of this excitation was significantly higher to the FS cue compared to the S cue (Table 2).

FS/F-overlapping neurons

Compared to the 400-ms precue baseline, these neurons showed a significant increase ($n = 7$; Fig. 5B, left) or decrease ($n = 5$; Fig. 5B, right) in firing rate in the 400-ms postcue to the FS and F cues but not the S or R cues. For those showing excitation, the Z-scored neural activity during the 400-ms postcue period was significantly higher during both the FS and F cues compared to the S and R cues (Table 2). The 1-way ANOVA for neurons with a significant inhibition to the FS and F cues was not significant (Table 2).

FS/R-overlapping neurons

Compared to other cue-overlapping categories, neurons showing an excitation to the FS and R cues, and not the F or S cues, made up the largest subcategory ($n = 28$; Fig. 5C, left). Z-scored activity

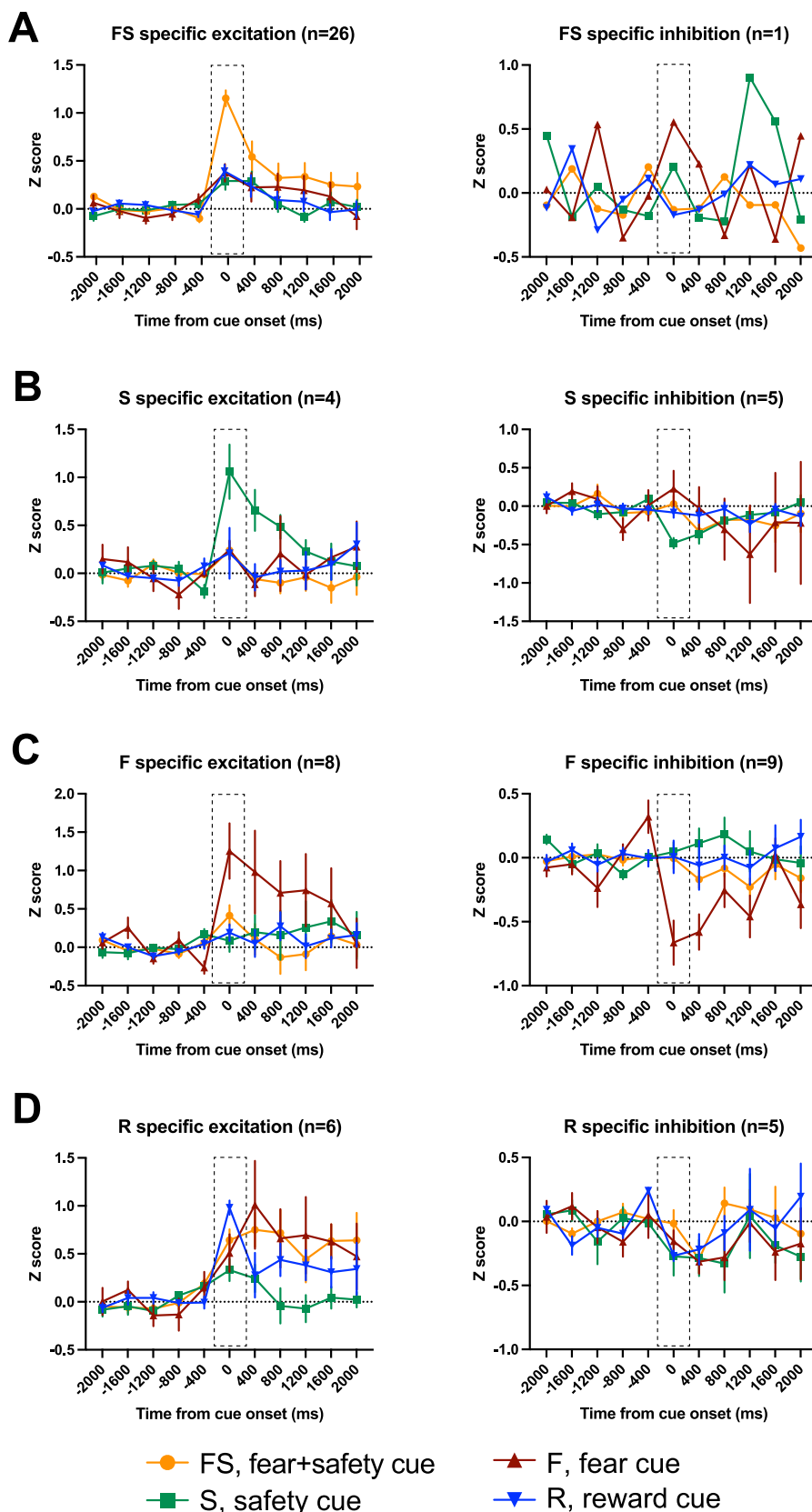


Fig. 4. Cue-specific neurons. Averaged Z-score activity across each subcategory of cue-specific neurons. Each of these neurons showed a significant change in activity in the first 400 ms after cue onset compared to 400 ms before cue onset to 1 cue specifically. Activity for the 2 s before and after cue onset are shown as time from cue onset (ms). Between-cue analyses were completed for the time bin 0–400 ms shown in the dashed boxes. A) FS specific excitation (left) and inhibition (right). B) S-specific excitation (left) and inhibition (right). C) F-specific excitation (left) and inhibition (right). D) R-specific excitation (left) and inhibition (right). Means \pm SEM.

Table 1. Cue-specific neurons. Summary of 1-way ANOVA results comparing neural activity across cues for the first 400 ms after cue onset.

| Neuron category | N | F(DFn, DFd) | P value | Tukey's multiple comparisons test |
|------------------------|----|-------------------|------------|--|
| FS-specific excitation | 26 | F(3, 100) = 23.01 | P < 0.0001 | FS > S, P < 0.0001 FS > F, P < 0.0001 FS > R, P < 0.0001 |
| FS-specific inhibition | 1 | N/A | N/A | N/A |
| S-specific excitation | 4 | F(3, 12) = 4.06 | P = 0.03 | ns |
| S-specific inhibition | 5 | F(3, 16) = 4.74 | P = 0.02 | S < F, P = 0.01 |
| F-specific excitation | 8 | F(3, 28) = 6.23 | P < 0.01 | F > FS, P = 0.04 F > S, P < 0.01 F > R, P < 0.01 |
| F-specific inhibition | 9 | F(3, 32) = 7.87 | P < 0.001 | F < FS, P < 0.01 F < S, P < 0.01 F < R, P < 0.01 |
| R-specific excitation | 6 | F(3, 20) = 4.33 | P = 0.02 | R > S, P = 0.01 |
| R-specific inhibition | 5 | F(3, 16) = 1.38 | P = 0.29 | N/A |

N/A, not applicable; ns, not significant.

for these neurons was significantly higher during both the FS and R cues compared to the F and S cues (Table 2). In addition, there was 1 neuron that showed inhibition during the FS and R cues (Fig. 5C, right).

FS/R/F-overlapping neurons

All neurons within this category showed a significant excitation to the FS, R, and F cues but not the S cue ($n = 21$; Fig. 5D, left). Averaged Z-scored activity during the S cue was significantly lower than all other cues (Table 2).

Neurons that showed excitation to all cues

All neurons within this final category showed a significant excitation to all cues ($n = 8$; Fig. 5D, right), with no neurons showing a similar inhibition across all cues. Comparing across cues during the 400-ms postcue period, the Z-scored neural activity was significantly higher during the FS cue compared to the S cue (Table 2).

Interestingly, in every single subcategory of cue-overlapping neurons that showed an excitation to the FS cue, the response to the FS cue was of significantly higher magnitude than the S cue even in those subcategories in which there was a significant excitation to the S cue compared to precue baseline (e.g. FS/S-overlapping neurons) (Table 2).

Bidirectional neurons in the IL

Even though not numerous, a small subset of neurons that showed a significant change in firing rate to the FS cue also showed the opposite response to the F cue ($n = 8$; Fig. 6). The majority of these bidirectional neurons showed an excitation to the FS cue and inhibition to the F cue and no response to the S or R cues ($n = 6$; Fig. 6A). Comparing across cues during the 400-ms postcue period (1-way ANOVA, $F(3, 20) = 24.86$, $P < 0.0001$), the Z-scored neural activity for the F cue was significantly lower than all other cues (FS, $P < 0.0001$; S, $P < 0.01$; R, $P < 0.0001$), while activity to the FS cue was higher than both the F ($P < 0.0001$) and S ($P < 0.01$) cues but not the R cue. Two neurons showed the opposite bidirectional response: inhibition to the FS cue and excitation to the F cue (Fig. 6B). The 1-way ANOVA for these neurons was not significant ($F(3, 4) = 6.01$, $P = 0.06$).

Correlating excitation to the FS cue with freezing behavior

To assess if the excitation to the FS cue was possibly negatively correlated with expressed freezing behavior during the FS cue, we calculated Pearson's correlations of the Z-scored neural activity during the 0–400-ms postcue onset with the amount of freezing expressed across the 20-s FS cue on a trial-by-trial basis. That is, during DC4, we separated the 15 FS trials to take advantage of the variability in freezing across trials; there was very little range in freezing across trials for the other cue types. Instead of calculating these correlations for every single subcategory reported, we focused our analyses to the more prominent subcategories: FS specific excitation (26 neurons; Fig. 7A), FS/R overlapping excitation (28 neurons; Fig. 7B), and FS/R/F overlapping excitation (21 neurons; Fig. 7D). We also included the bidirectional neurons showing FS excitation and F inhibition (6 neurons; Fig. 7C), as their neural activity profile appears to be particularly interesting for behavioral correlations. We found significant negative correlations between the neural activity and freezing behavior for neurons showing FS specific excitation ($r = -0.467$, $P < 0.0001$), FS/R overlapping excitation ($r = -0.502$, $P < 0.0001$), and bidirectional FS excitation/F inhibition ($r = -0.397$, $P < 0.0001$). There was no significant correlation found for neurons showing FS/R/F overlapping excitation ($r = +0.038$, $P = 0.51$).

Discussion

Given the demonstrated necessity of the IL in fear extinction recall (Laurent and Westbrook 2009; Sierra-Mercado et al. 2011; Sangha, Robinson, et al. 2014; Bukalo et al. 2015; but see Do-Monte et al. 2015), and conditioned inhibition of fear in response to a safety cue (Sangha, Robinson, et al. 2014), we hypothesized that IL activity would correlate with safety conditioning. Supporting our hypothesis, we observed a large proportion of neurons with excitatory responses to the fear + safety cue specifically, a separate group of neurons with excitatory responses to both the fear + safety and reward cues, and bidirectional neurons with excitation to the fear + safety cue and inhibition to the fear cue. Together, these data implicate the IL in encoding specific aspects of conditioned inhibitors and may be particularly engaged during the safety cue when presented in conflict with the fear cue when fear behavior is being actively suppressed.

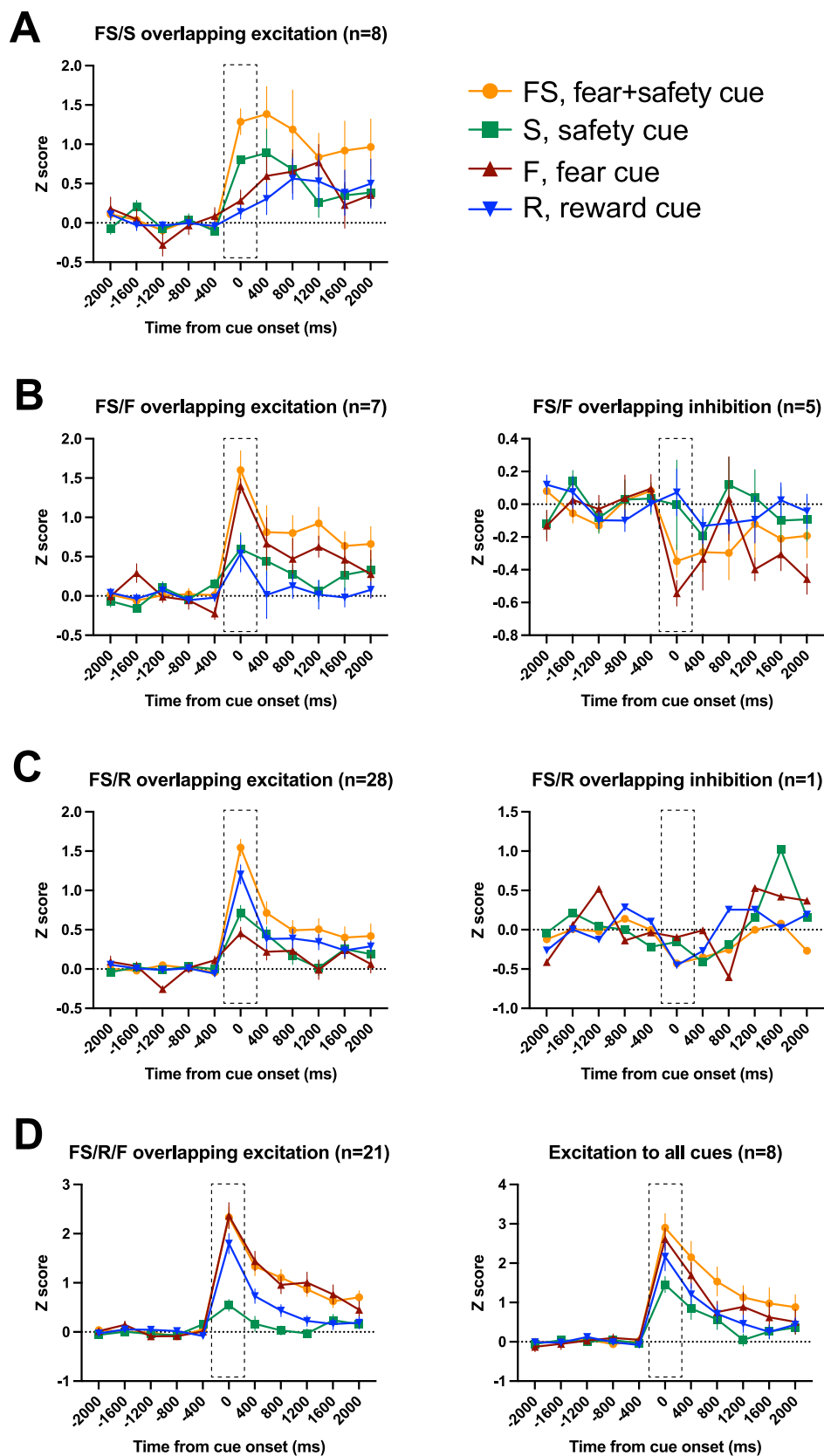


Fig. 5. Cue-overlapping neurons. Averaged Z-score activity across each subcategory of cue-overlapping neurons. Each of these neurons showed a significant change in activity in the first 400 ms after cue onset compared to 400 ms before cue onset to >1 cue. Activity for the 2 s before and after cue onset are shown as time from cue onset (ms). Between-cue analyses were completed for the time bin 0–400 ms shown in the dashed boxes. A) FS/S overlapping excitation (left). B) FS/F overlapping excitation (left) and inhibition (right). C) FS/R overlapping excitation (left) and inhibition (right). D) FS/R/F overlapping excitation (left) and excitation to all 4 cues (right). Means \pm SEM.

Table 2. Cue-overlapping neurons. Summary of 1-way ANOVA results comparing neural activity across cues for the first 400 ms after cue onset.

| Neuron category | N | F(DFn, DFd) | P value | Tukey's multiple comparisons test |
|-------------------------------|----|-------------------|------------|--|
| FS/S-overlapping excitation | 8 | F(3, 28) = 19.77 | P < 0.0001 | FS > S, P = 0.03 FS > F, P < 0.0001 FS > R, P < 0.0001 S > F, P = 0.02 S > R, P < 0.01 |
| FS/F-overlapping excitation | 7 | F(3, 24) = 7.19 | P < 0.01 | FS > S, P < 0.01 FS > R, P < 0.01 F > S, P = 0.047 F > R, P = 0.03 |
| FS/F-overlapping inhibition | 5 | F(3, 16) = 3.03 | P = 0.06 | N/A |
| FS/R-overlapping excitation | 28 | F(3, 108) = 22.88 | P < 0.0001 | FS > S, P < 0.0001 FS > F, P < 0.0001 R > S, P < 0.01 R > F, P < 0.0001 |
| FS/R-overlapping inhibition | 1 | N/A | N/A | N/A |
| FS/R/F-overlapping excitation | 21 | F(3, 76) = 17.14 | P < 0.0001 | FS > S, P < 0.0001 F > S, P < 0.0001 R > S, P < 0.001 FS > S, P = 0.01 |
| Excitation to all cues | 8 | F(3, 28) = 4.22 | P = 0.01 | |

N/A, not applicable.

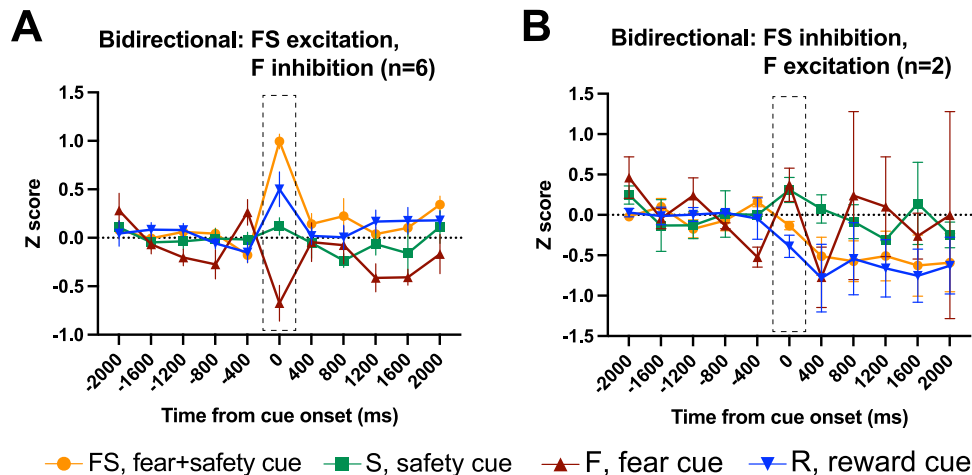


Fig. 6. Bidirectional neurons. Averaged Z-score activity across each subcategory of bidirectional neurons. Each of these neurons showed a significant change in activity in the first 400 ms after cue onset compared to 400 ms before cue onset to >1 cue. Activity for the 2 s before and after cue onset are shown as time from cue onset (ms). Between-cue analyses were completed for the time bin 0–400 ms shown in the dashed boxes. A) Neurons that showed a significant excitation to the FS cue and inhibition to the F cue. B) Neurons that showed a significant inhibition to the FS cue and excitation to the F cue. Means \pm SEM.

Overall, our results are consistent with previous studies. Here, we demonstrated that IL showed an increased neural activity in response to the combined fear + safety cue. Using the same training paradigm, our previous data have shown that IL was necessary during later discrimination sessions for suppressing learned fear in the presence of a learned safety signal (Sangha, Robinson, et al. 2014). Inactivating the IL after rats had learned the DC task, i.e. during DC4, our previous study presumably prevented the expression of the increased neural activity to the fear + safety cue, preventing learned safety behavior to be expressed. Our results showing increased activity within the first 400 ms of cue onset specific to the combined fear + safety cue is also consistent with a previous study showing increased IL activity to an extinguished fear cue during extinction recall (Milad and Quirk 2002). The IL may thus encode fear suppression during the early portion of a cue, with activity within the first second

being involved with initiating fear suppression. The more sustained increase we observed in many of our subcategories (e.g. FS/S overlapping neurons; Fig. 5A) may allow for the suppression of fear behavior to be maintained for a longer period of time. Our analyses were restricted to later sessions when reduced fear behavior to the fear + safety cue was quite consistent. Thus, the increased activity to the fear + safety cue during these later sessions may reflect the outcomes of the fear + safety cue versus fear cue becoming more certain, with the IL being part of a circuit that is estimating threat probability. Other regions that have been demonstrated to be engaged during threat uncertainty are the ventrolateral periaqueductal gray (Wright and McDannald 2019) and bed nucleus of the stria terminalis (Bjorn et al. 2020). Additionally, the posterior insular cortex has been shown to be necessary for conditioned inhibition of fear during a fear + safety cue (Foilb et al. 2016).

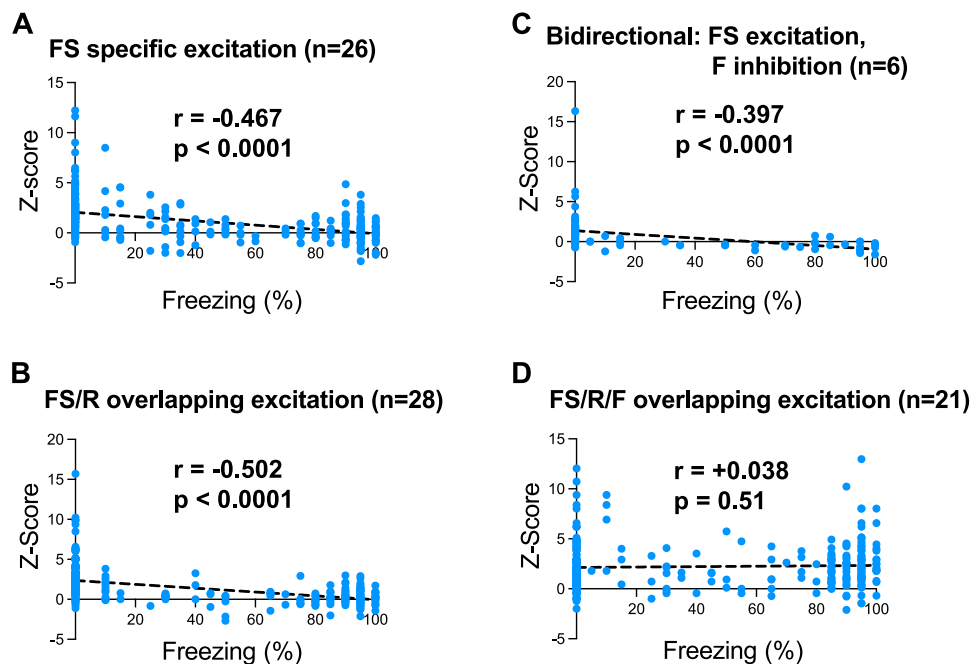


Fig. 7. Correlating neural activity with behavior. For each FS trial in DC4, Z-scored neural activity for the first 400 ms after cue onset was correlated with the percentage time spent freezing during the entire 20-s FS cue. A) Significant negative correlations were found for the neurons showing FS-specific excitation, B) FS/R overlapping excitation, C) and bidirectional neurons showing FS excitation/F inhibition. D) There was no significant correlation found for neurons showing FS/R/F excitation.

In an extinction of fear conditioning paradigm, [Giustino et al. \(2016\)](#) showed that, in male rats, the balance of neural activity between the IL and PL was correlated with the amount of fear that was expressed after extinction, with high freezing levels correlated with decreased IL activity relative to PL. Our results here are consistent with this as we found that neural activity to the FS cue was negatively correlated with freezing levels; i.e. low freezing levels were correlated with increased neural activity within the first 400 ms of the FS cue. Whether this is true for female subjects remains to be determined as our previous results have demonstrated that female Long Evans rats in our paradigm do not suppress freezing in the presence of a safety cue to the same extent as male rats ([Greiner et al. 2019](#)). We predict that the lack of fear suppression in female rats would correlate with decreased IL activity during the FS. Extinction has also repeatedly been shown to be context-specific (reviewed in [Bouton et al. 2021](#)). At a behavioral level, we would predict the fear suppressing ability of the safety cue would transfer to a new context, given all rats had reward training before the DC sessions in the same context and that we have not observed any background freezing or reward-seeking in any of our prior studies using the same paradigm ([Sangha, Greba, et al. 2014](#); [Ng et al. 2018](#); [Greiner et al. 2019](#); [Woon et al. 2020](#)). However, whether or not the neural activity within the IL to the FS cue would generalize to a new context is unclear and remains to be tested.

Interestingly, in every single subcategory showing an excitation to the FS cue, the response to the FS cue was of significantly higher magnitude as the S cue even in those subcategories in which there was a significant excitation to the S cue. This may indicate the importance of IL activity during conflict cues, here a compound fear + safety cue. Even though rats learned to significantly suppress freezing to the FS cue compared to the F cue, averaged freezing did not go to 0. Thus, the behavioral data suggest that there is still some conflict or uncertainty during the FS cue. When we correlated freezing levels to the FS cue on

a trial-by-trial basis to the neural activity in the first 400 ms of the FS cue, we found that low freezing levels were correlated with increased IL activity. These data appear to support the idea that the IL is being selectively engaged during instances where an adaptive behavior needs to be selected over a maladaptive behavior ([Nett and LaLumiere 2021](#)).

Compared to our previous work in the amygdala, it is interesting that there was a high number of IL neurons showing cue-evoked responding to the fear + safety cue but not the safety cue. This is in contrast to our previous work showing safety-related activity in the basal amygdala (BA), where significant changes in firing rates were seen consistently to both the safety cue and fear + safety cue ([Sangha et al. 2013](#)). In both studies, we counterbalanced the tone and light cues as the fear and safety cues. Perhaps the IL is not encoding the stimulus properties of the safety signal but, instead, the IL may only be involved during active fear suppression when the fear cue is also present. That is, there is something unique about the conflicting scenario of both the fear and safety cues being presented concurrently that the IL may be selectively engaged in. This again supports the idea that IL's function may be to select a particular behavior under conflict ([Nett and LaLumiere 2021](#)).

While the amygdala and hippocampus are also involved in safety processing, data looking at theta synchrony found that the firing of IL neurons leads the firing of neurons in the lateral amygdala and the CA1 region of the dorsal hippocampus during the recall of a successfully extinguished fear cue ([Lesting et al. 2013](#)). In a conditioned inhibition task, neurons projecting from the ventral hippocampus to the IL or basolateral amygdala (BLA) did not change in activity to a compound fear + safety cue, whereas neurons projecting from the ventral hippocampus to the PL prefrontal cortex did increase in activity to the fear + safety cue ([Meyer et al. 2019](#)). It is thus unlikely the increased activity we observed in the IL to the fear + safety cue was caused by increasing the input from the ventral hippocampus. Instead, IL may be

receiving an increased input from the BLA during safety learning. Reciprocal projections between the BLA and IL have been shown to be important in the acquisition and consolidation of fear extinction (Bukalo et al. 2015). IL neurons have excitatory projections to BLA pyramidal neurons, which are important for fear extinction (Strobel et al. 2015). Neurons in the BLA show decreased activity to the fear cue after extinction, which can be observed throughout the entire cue presentation (Hobin et al. 2003). Our previous data have shown both excitatory and inhibitory responses to the fear + safety cue and safety cue by neurons within the BA (Sangha et al. 2013). Optical activation of IL during the fear cue has been shown to facilitate fear extinction and subsequent retrieval of fear extinction, whereas optically inactivating the IL during the fear cue during retrieval impaired the retrieval of fear extinction (Do-Monte et al. 2015). Taken together, this reciprocal IL-BLA loop may be common during the learning phase for both fear extinction and conditioned inhibition to drive fear suppression (Sangha et al. 2020). Once learned though, expression of learned fear suppression via extinction does not require IL-BLA signaling (Bukalo et al. 2015) and may be instead mediated by IL's projections to the intercalated cells and/or central amygdala.

One disadvantage of our electrophysiological recordings in the present study is the inability to confirm the cell types responsible for fear suppression within the IL. It is possible that the cells that showed learning related changes in the present study were glutamatergic neurons since optically inactivating IL pyramidal neurons with a halorhodopsin-expressing virus using a CaMKII promoter impaired the retrieval of fear extinction (Do-Monte et al. 2015). The increased IL activity we observed may suppress fear behavior through BLA interneurons, in which there are two primary types. Typically, an auditory fear cue will excite parvalbumin-expressing interneurons which will, in turn, inhibit somatostatin-expressing (SOM) interneurons, leading to disinhibition of BLA principal neurons (Wolff et al. 2014). During fear suppression, we hypothesize that the excitatory IL input may decrease the BLA principal neuron activity by providing excitatory input to SOM interneurons directly, preventing the disinhibition in response to the fear cue.

Our present study has established that the IL encodes specific aspects of conditioned inhibitors when fear is being actively suppressed. We have identified changes in IL activity which were specific to when the fear cue was presented in conflict with the safety cue. This adds further support to the idea that the IL may be involved in downregulating fear across all types of inhibitory memories, as increased IL activity is associated with extinction (Laurent and Westbrook 2009; Sierra-Mercado et al. 2011; Bukalo et al. 2015), safety conditioning (Sangha, Robinson, et al. 2014) and latent inhibition (Lingawi et al. 2017) but not conditioned inhibition of appetitive behavior (Rhodes and Killcross 2007). These findings have advanced our understanding of IL's role in suppressing fear to a safety cue and begins to tease apart the neural dynamics involved in suppressing fear under conditions where it may be unclear if there is a viable threat.

Acknowledgements

We thank Yolanda Jonker and Signe Hobaugh for excellent animal care and Brent Bachman for technical assistance.

Funding

This work was supported by the National Institute of Mental Health at the National Institutes of Health (R01MH110425

to SS). KHN was supported by a Purdue University Bilsland Fellowship.

Conflict of interest statement: None declared.

References

- Bjorni M, Rovero NG, Yang ER, Holmes A, Halladay LR. Phasic signaling in the bed nucleus of the stria terminalis during fear learning predicts within- and across-session cued fear expression. *Learn Mem.* 2020;27(3):83–90.
- Blanchard RJ, Blanchard DC. Crouching as an index of fear. *J Comp Physiol Psychol.* 1969;67(3):370–375.
- Bouton ME. *Learning and behavior: a contemporary synthesis.* Sunderland (MA): Sinauer Associates, Inc; 2007
- Bouton ME, Maren S, McNally GP. Behavioral and neurobiological mechanisms of Pavlovian and instrumental extinction learning. *Physiol Rev.* 2021;101(2):611–681.
- Bukalo O, Pinard CR, Silverstein S, Brehm C, Hartley ND, Whittle N, Colacicco G, Busch E, Patel S, Singewald N, et al. Prefrontal inputs to the amygdala instruct fear extinction memory formation. *Sci Adv.* 2015;1(6):e1500251.
- Christianson JP, Fernando ABP, Kazama AM, Jovanovic T, Ostroff LE, Sangha S. Inhibition of fear by learned safety signals: a mini-symposium review. *J Neurosci.* 2012;32(41):14118–14124.
- Do-Monte FH, Manzano-Nieves G, Quiñones-Laracuente K, Ramos-Medina L, Quirk GJ. Revisiting the role of infralimbic cortex in fear extinction with optogenetics. *J Neurosci.* 2015;35(8):3607–3615.
- Fendt M, Fanselow MS. The neuroanatomical and neurochemical basis of conditioned fear. *Neurosci Biobehav Rev.* 1999;23(5):743–760.
- Foilib AR, Flyer-Adams JG, Maier SF, Christianson JP. Posterior insular cortex is necessary for conditioned inhibition of fear. *Neurobiol Learn Mem.* 2016;134 Pt B:317–327.
- Giustino TF, Fitzgerald PJ, Maren S. Fear expression suppresses medial prefrontal cortical firing in rats. *PLoS One.* 2016;11(10):e0165256.
- Goode TD, Jin J, Maren S. Neural circuits for fear relapse. In: Sangha S, Foti D, editors. *Neurobiology of abnormal emotion and motivated behaviors: integrating animal and human research.* 1st ed. London (UK): Academic Press; 2018. pp. 183–202.
- Greiner EM, Müller I, Norris MR, Ng KH, Sangha S. Sex differences in fear regulation and reward-seeking behaviors in a fear-safety-reward discrimination task. *Behav Brain Res.* 2019;368:111903.
- Hobin JA, Goosens KA, Maren S. Context-dependent neuronal activity in the lateral amygdala represents fear memories after extinction. *J Neurosci.* 2003;23(23):8410–8416.
- Jovanovic T, Norrholm SD, Blanding NQ, Davis M, Duncan E, Bradley B, Ressler KJ. Impaired fear inhibition is a biomarker of PTSD but not depression. *Depress Anxiety.* 2010;27(3):244–251.
- Laurent V, Westbrook RF. Inactivation of the infralimbic but not the prelimbic cortex impairs consolidation and retrieval of fear extinction. *Learn Mem.* 2009;16(9):520–529.
- Lesting J, Daldrup T, Narayanan V, Himpe C, Seidenbecher T, Pape H-C. Directional theta coherence in prefrontal cortex to amygdalo-hippocampal pathways signals fear extinction. *PLoS One.* 2013;8(10):e77707.
- Lingawi NW, Westbrook RF, Laurent V. Extinction and latent inhibition involve a similar form of inhibitory learning that is stored in and retrieved from the infralimbic cortex. *Cer Cortex.* 2017;27:5547–5556.
- Meyer HC, Odriozola P, Cohodes EM, Mandell JD, Li A, Yang R, Hall BS, Haberman JT, Zacharek SJ, Liston C, et al. Ventral hippocampus interacts with prelimbic cortex during inhibition of threat

- response via learned safety in both mice and humans. *Proc Natl Acad Sci U S A*. 2019;116(52):26970–26979.
- Milad MR, Quirk GJ. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature*. 2002;420(6911):70–74.
- Morriss J, Christakou A, van Reekum CM. Intolerance of uncertainty predicts fear extinction in amygdala-ventromedial prefrontal cortical circuitry. *Biol Mood Anxiety Disord*. 2015;5(1):4.
- Müller I, Brinkman AL, Sowinski EM, Sangha S. Adolescent conditioning affects rate of adult fear, safety and reward learning during discriminative conditioning. *Sci Rep*. 2018;8(1):17315.
- Nett KE, LaLumiere RT. Infralimbic cortex functioning across motivated behaviors: Can the differences be reconciled? *Neurosci Biobehav Rev*. 2021;131:704–721.
- Ng KH, Pollock MW, Urbanczyk PJ, Sangha S. Altering D1 receptor activity in the basolateral amygdala impairs fear suppression during a safety cue. *Neurobiol Learn Mem*. 2018;147:26–34.
- Odriozola P, Gee DG. Learning about safety: Conditioned inhibition as a novel approach to fear reduction targeting the developing brain. *Am J Psychiatry*. 2021;178(2):136–155.
- Patel R, Spreng RN, Shin LM, Girard TA. Neurocircuitry models of posttraumatic stress disorder and beyond: a meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev*. 2012;36(9):2130–2142.
- Rhodes SEV, Killcross AS. Lesions of rat infralimbic cortex result in disrupted retardation but normal summation test performance following training on a Pavlovian conditioned inhibition procedure. *Eur J Neurosci*. 2007;26(9):2654–2660.
- Sangha S. Plasticity of fear and safety neurons of the amygdala in response to fear extinction. *Front Behav Neurosci*. 2015;9:354.
- Sangha S, Chadick JZ, Janak PH. Safety encoding in the basal amygdala. *J Neurosci*. 2013;33(9):3744–3751.
- Sangha S, Greba Q, Robinson PD, Ballendine SA, Howland JG. Heightened fear in response to a safety cue and extinguished fear cue in a rat model of maternal immune activation. *Front Behav Neurosci*. 2014;8:168.
- Sangha S, Robinson PD, Greba Q, Davies DA, Howland JG. Alterations in reward, fear and safety cue discrimination after inactivation of the rat prelimbic and infralimbic cortices. *Neuropsychopharmacology*. 2014;39(10):2405–2413.
- Sangha S, Diehl MM, Bergstrom HC, Drew MR. Know safety, no fear. *Neurosci Biobehav Rev*. 2020;108:218–230.
- Sierra-Mercado D, Padilla-Coreano N, Quirk GJ. Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. *Neuropsychopharmacology*. 2011;36(2):529–538.
- Strobel C, Marek R, Gooch HM, Sullivan RKP, Sah P. Prefrontal and auditory input to intercalated neurons of the amygdala. *Cell Rep*. 2015;10(9):1435–1442.
- Wolff SBE, Gründemann J, Tovote P, Krabbe S, Jacobson GA, Müller C, Herry C, Ehrlich I, Friedrich RW, Letzkus JJ, et al. Amygdala interneuron subtypes control fear learning through disinhibition. *Nature*. 2014;509(7501):453–458.
- Woon EP, Seibert TA, Urbanczyk PJ, Ng KH, Sangha S. Differential effects of prior stress on conditioned inhibition of fear and fear extinction. *Behav Brain Res*. 2020;381:112414.
- Wright KM, McDannald MA. Ventrolateral periaqueductal gray neurons prioritize threat probability over fear output. *elife*. 2019;8:e45013.