Neurobiology of Learning and Memory 180 (2021) 107405

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme

Elemental and configural threat learning bias extinction generalization

Elizabeth V. Goldfarb[a,*], Tahj Blow[b], Joseph E. Dunsmoor[c], Elizabeth A. Phelps[d]

[a] Departments of Psychiatry and Psychology, Yale University, New Haven, CT, USA
[b] Weill Cornell College of Medicine, New York, NY, USA
[c] Department of Psychiatry, University of Texas at Austin, Austin, TX, USA
[d] Department of Psychology, Harvard University, Cambridge, MA, USA

ARTICLE INFO

Keywords:
Fear
Conditioning
Extinction
Configural
Elemental

ABSTRACT

Emotional experiences often contain a multitude of details that may be represented in memory as individual elements or integrated into a single representation. How details associated with a negative emotional event are represented in memory can have important implications for extinction strategies designed to reduce emotional responses. For example, is extinguishing one cue associated with an aversive outcome sufficient to reduce learned behavior to other cues present at the time of learning that were not directly extinguished? Here, we used a between-subjects multi-day threat conditioning and extinction task to assess whether participants generalize extinction from one cue to unextinguished cues. On Day 1, one group of participants learned that a compound conditioned stimulus, composed of a tone and colored square, predicted an uncomfortable shock to the wrist (Compound group). A second group learned that the tone and square separately predicted shock (Separate group). On Day 2, participants in both groups were exposed to the tone in the absence of shocks (cue extinction). On Day 3, we tested whether extinction generalized from the extinguished to the unextinguished cue, as well as to a compound composed of both cues. Results showed that configural and elemental learning had unique and opposite effects on extinction generalization. Subjects who initially learned that a compound cue predicted shock successfully generalized extinction learning from the tone to the square, but exhibited threat relapse to the compound cue. In contrast, subjects who initially learned that each cue individually predicted shock did not generalize extinction learning from the tone to the square, but threat responses to the compound were low. These results highlight the importance of whether details of an aversive event are represented as integrated or separated memories, as these representations affect the success or limits of extinction generalization.

1. Introduction

Aversive experiences are complex, containing many cues that can be remembered in different ways (Brewin, 2014; Rudy, Huff, & Matus-Amat, 2004). How these events are remembered may have important consequences for the spreading, or generalization, of later threat responses (Maren, Phan, & Liberzon, 2013). However, the effects of these different threat memories on the generalization of later extinction learning – a crucial aspect of exposure therapy (Kredlow, de Voogd, & Phelps, 2020) – is not yet clear.

Complex aversive events can be learned as configural representations, in which multiple cues are integrated into a coherent context, or elemental representations, in which separate associations are formed between each cue and the aversive outcome (Rudy et al., 2004; Stout et al., 2018). An imbalance between these representations, with an overreliance on elemental rather than configural learning, may contribute to the etiology and maintenance of posttraumatic stress disorder (Acheson, Gresack, & Risbrough, 2012; Bisby, Burgess, & Brewin, 2020; Gilbertson et al., 2007). For example, after elemental learning (in which each individual feature is separately associated with the aversive outcome), any individual feature – regardless of the context in which it occurs – could evoke a threat response, leading threat responses to generalize to inappropriate contexts. In contrast, configural learning (in which the conjunction of all present features is encoded as a unified whole) would require the environment to fully match the encoded representation in order to evoke a threat response (Acheson et al., 2012; Maren et al., 2013).

These different threat representations may also have implications for later extinction learning. During extinction, a previously threatening stimulus is repeatedly presented without reinforcement, enabling

* Corresponding author.
E-mail address: elizabeth.goldfarb@yale.edu (E.V. Goldfarb).

https://doi.org/10.1016/j.nlm.2021.107405
Received 13 November 2020; Received in revised form 8 February 2021; Accepted 11 February 2021
Available online 18 February 2021
1074-7427/

individuals to learn that this stimulus is now "safe" and attenuate their threat responses. Extinction learning can spread, changing responses to other stimuli in a process known as extinction generalization (Dubin & Levis, 1973; Vervliet, Vansteenwegen, & Eelen, 2004). This process is clinically relevant, as exposure therapy relies on extinction of individual items with the goal that this new learning will generalize to complex real-world scenarios (Kredlow et al., 2020; Rowe & Craske, 1998). This clinical importance has motivated laboratory investigations of strategies to promote extinction generalization (Dunsmoor, Niv, Daw, & Phelps, 2015; Fitzgerald, Seemann, & Maren, 2014; Hennings, Bibb, Lewis-Peacock, & Dunsmoor, 2020). A recent report showed that threat learning (inferred belief that linear vs similarity rules explained the occurrence of mild shocks) shaped later extinction generalization in humans (Wong, Glück, Boschet, & Engelke, 2020). However, the influence of elemental and configural threat learning on extinction generalization in humans remains unclear. Here we asked whether initial threat associations for multi-cue aversive experiences modulate extinction generalization from one component cue to the rest of the aversive event.

1.1. Generalizing extinction between cues

To date, most research examining the influence of threat representations on extinction generalization has been conducted in rodents, and has focused on generalization between individual cues. In these studies, animals are first conditioned to associate cues A and X with an electric shock. Next, cue A is presented without shock reinforcement (extinction) and responses to cue X are measured. These studies show that, if the initial threat learning was elemental (shock separately associated with cues A and X), extinguishing A did not generalize to X (Kasprow, Schachtman, Cacheiro, & Miller, 1984). In contrast, if the shock was paired with an AX compound, extinguishing component cue A did generalize and attenuate responses to X (Debiec, Diaz-Mataix, Bush, Doyere, & LeDoux, 2013; Durlach & Rescorla, 1980; Nakajima & Kawai, 1997; Pineno, 2007; Schnelker & Batsell, 2006). Notably, a recent study in humans differed from this pattern, providing evidence that forming separate cue/threat associations did enable extinction generalization between cues when considering explicit expectations of shock delivery; however, compound threat conditioning was not examined (Mertens et al., 2019; see also Vurbic & Bouton, 2011). These findings largely support the hypothesis that human participants who form separate threat representations will show impaired extinction generalization between cues relative to those who form compound representations.

This hypothesis is also supported by theoretical learning models. Several models posit that associations are formed during configural learning that are not present during elemental learning (see Harris, 2006 for review). For example, configural learning may involve the formation of links between component cues A and X (Durlach & Rescorla, 1980; Rescorla & Cunningham, 1978). These links would enable changes in the value of one cue during extinction (A) to update the value of the other unseen cue (X), enabling extinction to generalize between cues following configural learning. However, as these links between cues are not formed during elemental threat learning, the value of the unseen cue would not change during extinction, thus preventing extinction generalization.

1.2. Generalizing extinction from cues to compounds

Fewer studies have examined extinction generalization from component cues to the combination of cues that were part of the initial aversive event. That is, after learning to associate an AX compound with the delivery of shock, would extinction of cue A generalize and attenuate responses to the AX compound? These limited studies have yielded mixed results. Although some work has shown that some cue extinction procedures do generalize to the compound (Jones, Ringuet, & Monfils, 2013), other studies suggest weak generalization (Shanks, Darby, & Charles, 1998; Troisi, Dooley, & Craig, 2013) or indicate a potential for increased responding to the compound (Bouton, Doyle-Burr, & Vurbic, 2012). Furthermore, it is not yet clear whether extinction generalization from cues to a compound is modulated by whether the threat association was initially formed with separate cues or a full compound.

Here too, learning models provide clues regarding how different types of threat learning may influence extinction generalization. These models posit that, with configural learning, the full AX compound is represented as an indivisible unit (Melchers, Shanks, & Lachnit, 2008; Pearce, 2002), as a unique new cue that represents the AX combination ("added elements" model), and even as altering the representation of components A and X because they are presented together ("replaced elements" model, both discussed in Brandon, Vogel, & Wagner, 2000). Because extinction of individual component cues would not alter these configural threat associations, the AX compound could still elicit a threat response. In contrast, if threat learning was elemental, extinction of all component cues would fully address learned threat associations, and the AX compound would not trigger a threat response (Soto, Gershman, & Niv, 2014). This indicates a hypothesis that is distinct from generalization between cues. That is, associating a threat with a compound may limit extinction generalization from cues to a compound, whereas elemental learning would enable extinction to generalize.

1.3. Current study

In this study, we assessed whether the type of representation formed for a multi-cue aversive event would modulate how broadly human participants generalize extinction of component cues. After training participants to associate cues individually (Separate group) or in combination (Compound group) with an aversive outcome, we extinguished the aversive association with one cue. We then tested whether extinction learning generalized to the unseen cue. Next, we extinguished aversive associations with both cues individually, then tested whether extinction learning about both component cues would generalized to the combined presentation of both cues. Each phase of the experiment was separated by a 24-hour break to allow for the consolidation of acquisition and extinction memories (consistent with Debiec et al., 2013).

2. Material and methods

2.1. Participants

A total of 115 participants were enrolled in the experiments (Compound: N = 56, Separate: N = 59). Of these, 16 failed to complete all three days, 1 violated the protocol (had recently completed another shock study), and 2 had technical issues that made their data unusable. The final sample included 96 participants with 48 participants in each group. The groups did not differ in age or sex, or measures of anxiety (Spielberger, 1983), intolerance of uncertainty (Buhr & Dugas, 2002), or stress (Cohen, Kamarck, & Mermelstein, 1983) measured on the first day of the experiment (Table 1). All participants provided written informed consent before participation. Procedures were approved by the New York University Committee on Activities Involving Human Subjects.

Table 1
Participant demographics.

	Compound Group (N = 48)	Separate Group (N = 48)	Different?
Sex (% female)	54.1%	62.5%	$p > .25$
Age	23 [4.01]	22.25 [4.14]	$p > .25$
State anxiety (STAI-S)	33.35 [8.9]	34.23 [7.96]	$p > .25$
Trait anxiety (STAI-T)	41.6 [9.59]	41.94 [8.01]	$p > .25$
Intolerance of uncertainty (IUS)	63.77 [16.60]	63.04 [14.75]	$p > .25$
Perceived stress (PSS)	16.21 [6.91]	17.69 [7.30]	$p > .25$

2.2. Task design

All participants completed a 3-session threat conditioning and extinction paradigm (Fig. 1). Each session was separated by ~24 h. Stimuli were presented using Eprime 2.0. Trial sequences were pseudorandomized such that participants could not experience three of any trial type in a row, with participants randomly assigned to complete one of two pre-defined trial sequences.

Day 1: Acquisition. Participants viewed colored squares (blue or green) and listened to tones delivered through headphones (500 or 800 Hz). Each trial lasted 6 s followed by a jittered 8–10 s inter-trial interval (ITI) during which participants viewed a fixation cross on a blank screen. A subset of trials co-terminated with a mild electric shock to the wrist. These shocks were calibrated prior to the start of the conditioning session to a level that the participants considered "highly annoying but not painful" using a visual analogue scale. The shock was administered through pregelled snap electrodes (BIOPAC EL508) attached to the wrist of the participant's dominant hand and connected to a Grass Medical Instruments stimulator (Model SD9; West Warwick, RI).

In the Compound group, tones and squares were presented concurrently on each trial. Participants completed 22 trials (11 per tone/square pair), with one tone/square pair sometimes co-terminating with a mild shock (compound CS+) on 8 trials (73% reinforcement) and the other tone/square pair never co-terminating with shock (compound CS-).

In the Separate group, tones and squares were presented on separate trials. Participants completed 44 trials (11 per cue), with one tone (tone CS+) and one square (square CS+) each co-terminating with a mild shock at the same reinforcement rate (73%, 16 trials total reinforced) and the other tone (tone CS-) and square (square CS-) never co-terminating with shock.

Day 2: Extinction. Both Compound and Separate groups completed the same protocol. As on Day 1, pre-gelled snap electrodes were attached to the participant's dominant wrist and to the stimulator, which was switched on. However, no shocks were delivered during this session. Participants were presented with 24 trials (interleaved tone CS+ and tone CS-) on every trial and did not view any squares during this session.

Day 3: Test/Re-Extinction. Both Compound and Separate groups completed the same protocol. As on the previous days, pre-gelled snap electrodes were attached to the participant's dominant wrist and the stimulator. No shocks were delivered during this session. The first trial was always a tone CS- trial used to capture the initial orienting response and discarded from analysis (de Voogd & Phelps, 2020; Keller & Dunsmoor, 2020).

First, participants were exposed to each cue 5x with no reinforcement (interleaved tone CS+, tone CS-, square CS+, square CS-). This enabled us to assess extinction generalization between the tone and the square cues, and then extinguish associations with both cues.

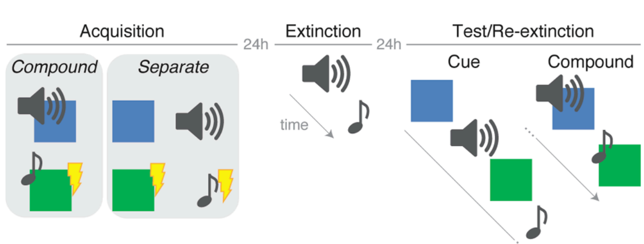

Fig. 1. Experimental procedure. All participants completed a three-day threat conditioning and extinction protocol. During acquisition, participants in the Compound group learned to associate a tone/colored square compound with shock. Participants in the Separate group learned that a colored square and a tone were each separately associated with shock. Following acquisition, all participants in both groups completed the same extinction (tones with no shock) and test/re-extinction (all stimuli presented with no shock) procedures. During the test/re-extinction period, participants were first exposed to individual cues (tones and squares), followed by compound stimuli.

Next, participants were exposed to each tone/square compound 5x with no reinforcement (interleaved compound CS+, compound CS-). This enabled us to assess extinction generalization from component cues (both of which were now extinguished) to the compound.

2.3. Measures of learning

Skin conductance responses (SCR). Anticipatory physiological responses were measured using skin conductance (SCR). Levels were assessed using two Ag/AgCL electrodes with NaCL gel attached to the left palm.

Shock expectancy. Throughout the experiment, participants rated how much they expected a shock on every trial using a three-alternative forced choice response (+ = shock anticipated, ? = not sure, - = no shock anticipated).

2.4. Analyses

Analyses were conducted in R (3.6.2). As participants could complete the experiments in one of two pre-defined trial sequences, we included trial order as a covariate in all models.

SCR. Response magnitude (de Voogd & Phelps, 2020) was scored automatically using Autonomate (Green, Kragel, Fecteau, & LaBar, 2014) with Matlab 2019b (MathWorks). Responses were considered valid if the trough-to-peak response occurred between 0.5 and 6 s after stimulus onset with a minimum amplitude of 0.02 microsiemens and a maximum duration of 5 s. The largest response during this window was counted. Trials that did not meet these criteria were scored as zero. We then computed the square root of the response magnitude on each trial.

Responses were analyzed using linear mixed effects models (*nlme* package) with participant as a random effect. Effect sizes were estimated using *sjstats*. Significant interactions were followed by pairwise comparisons of estimated marginal means (*emmeans* package).

Shock expectancy. To facilitate interpretation, responses that shock was anticipated (+) were coded as "Yes", and not sure (?) and no (-) were coded as "Other". Ratings were analyzed per trial using binomial general linear mixed effects models with participant as a random effect, followed by pairwise comparisons of estimated marginal means.

3. Results

3.1. Threat acquisition and cue extinction

To determine whether extinction would generalize, we first needed to ascertain that threat responses were successfully acquired and extinguished.

3.1.1. Threat acquisition

Trial-by-trial SCR responses throughout the experiment are shown in Fig. 2A. To test whether participants successfully learned to associate cues with shock, we averaged SCR during the first and second half of acquisition, separately for shock-paired (CS+) and unpaired (CS-) trials. The difference between responses to CS+ and CS- trials will hereafter be referred to as the conditioned response. One participant had an error with SCR data collection during acquisition and was not included in these analyses.

Across groups, we observed a significant conditioned response (main effect Shock Reinforcement: $F(1,279) = 126.38$, $p < .001$, $\eta_p^2 = 0.31$), which did not differ between Separate and Compound groups (Group × Shock Reinforcement: $p > .25$). There was also a significant main effect of time ($F(1,279) = 75.41$, $p < .001$, $\eta_p^2 = 0.21$), and an interaction between time and shock reinforcement ($F(1,279) = 11.83$, $p < .001$, $\eta_p^2 = 0.04$), with a stronger conditioned response during the second half of acquisition (CS+ vs CS-: $b = 0.27$ [0.03], $p < .001$) compared to the first half of acquisition ($b = 0.14$ [0.03], $p < .001$). Within the Separate group, we separated responses by cue modality (tone vs. square). We

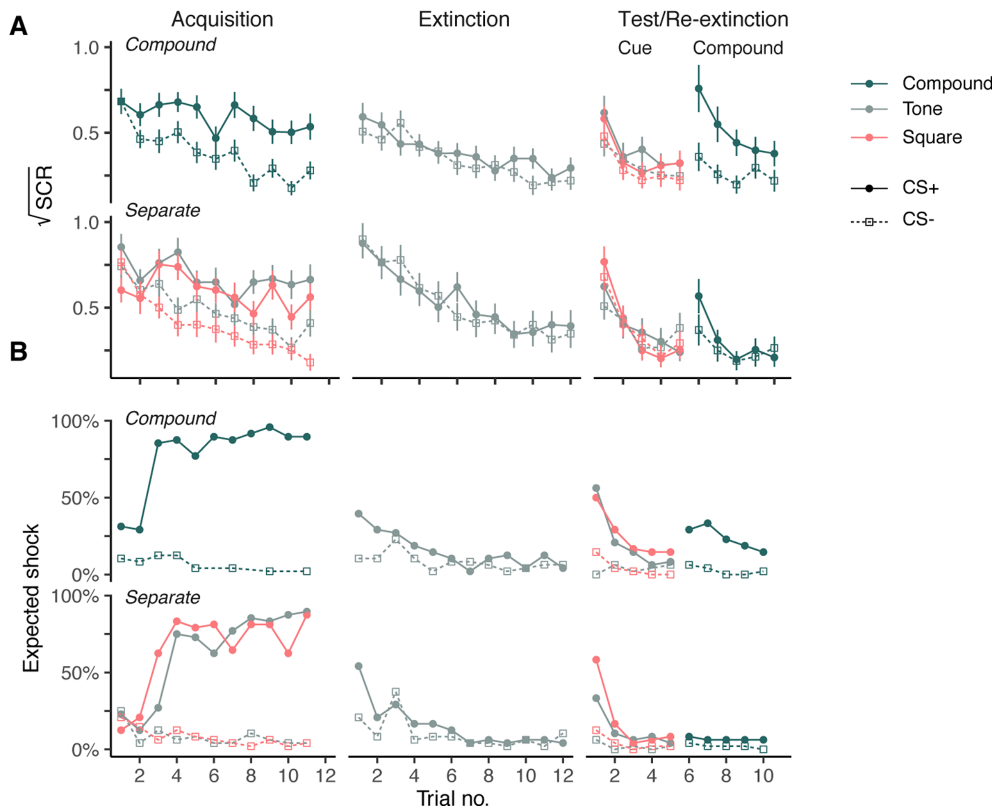

Fig. 2. Time course of trial-by-trial responses for Compound and Separate groups. (A) Skin conductance response (SCR; processed as the square root of trough-to-peak responses) per trial and cue. CS+ = stimulus associated with shock during acquisition; CS- = stimulus not associated with shock during acquisition. Error bars = ± 1 SE. (B) Proportion of participants rating that they expected a shock to occur on that trial.

confirmed that, although participants were overall more reactive to tone than square cues ($F(1,329) = 15.18$, $p < .001$, $\eta_p^2 = 0.044$), conditioned responses did not significantly differ between modalities, indicative of comparable conditioning across cues (Cue Type × Shock Reinforcement: $p > .25$).

Trial-by-trial ratings of shock expectancy (Fig. 2B) also provided evidence for successful acquisition of threat associations. In the second half of acquisition, there was a significant main effect of shock reinforcement across groups ($b = 6.88$ [SD = 0.58], $p < .001$), with higher expectation of shock from the CS+ than the CS-. Unlike SCR, these ratings differed between Separate and Compound groups (Group × Shock Reinforcement: $b = 1.84$ [0.62], $p = .003$). Although both groups learned to expect shock on CS+ trials (CS+ v CS-, Separate: $b = 5.04$ [0.28], $p < .001$; Compound: $b = 6.88$ [0.58], $p < .001$), the Separate group were more likely than the Compound group to expect shock from the CS- ($b = 1.43$ [0.57], $p = .012$; CS+: $b = -0.41$ [0.31], $p = .18$), perhaps because the Separate group had more cues to learn. Finally, consistent with SCR responses, participants in the Separate group were more likely to expect shock from tone than square cues ($b = 0.58$ [2.8], $p = .035$), but conditioned responses did not differ significantly between cue modalities (Cue Type × Shock Reinforcement: $p > .25$). Together, these analyses confirm that both the Compound and Separate groups successfully acquired threat associations.

3.1.2. *Threat extinction: Single cue*

As with acquisition, we averaged SCR during the first and second half of extinction, separately for tone cues that had been previously associated (CS+) or not associated (CS-) with shock. One participant had an error with SCR data collection and was not included in these analyses. We did not observe a significant conditioned response during extinction (main effect Shock Reinforcement: $p > .25$) or a difference in conditioned responses between groups ($F(1,279) = 1.5$, $p = .22$). As there was a trend-level change in conditioned responses over time (Time × Shock Reinforcement: $F(1,279) = 2.96$, $p = .086$, $\eta_p^2 = 0.01$) we ran an additional analysis on the last trial of extinction, confirming that there was no significant difference in responses to the tone CS+ and CS- at the end of the extinction session ($p > .25$), indicative of successful extinction of the tone/shock association.

We note that the lack of conditioned response in SCR at the beginning of the extinction session is unusual. However, we did observe significant differences in ratings of shock expectancy. Participants were more likely to expect shock from the tone CS+ vs. CS- during the first half of extinction ($b = 1.23$ [0.28], $p < .001$) and this did not differ between groups (Group × Shock Reinforcement: $p > .25$). By the second half of extinction, this differential response was no longer evident (main effect Shock Reinforcement: $b = 0.57$ [0.46], $p = .22$) and did not significantly differ between groups (Group × Shock Reinforcement: $p > .25$). Thus, both SCR and expectancy metrics indicate that participants in both groups successfully extinguished the threat association with the tone.

3.2. *Extinction generalization between component cues*

We first tested whether extinction of the threat association with one cue (tone) would generalize and attenuate responses to the other cue (square). We hypothesized that extinction generalization would differ based on how the initial threat association was formed. Specifically, participants in the Separate group would not generalize extinction between cues, but participants in the Compound group would. We tested this hypothesis by analyzing responses at the beginning of the test phase and, in the Separate group, probing the change in response from the end of acquisition (pre-extinction) to the test phase (post-extinction).

3.2.1. Cue responses at the beginning of the test phase

We examined responses to individual cues at the beginning (first two trials) of the test phase, separating trials by group (Compound vs. Separate), cue type (tone vs. square) and whether they had previously been paired with shock (or, in the case of the Compound group, whether they had been part of a compound paired with shock). Confirming that threat associations from acquisition were remembered, we found a persistent conditioned response (main effect of past shock reinforcement: $F(1,282) = 7.2$, $p = .008$, $\eta_p^2 = 0.024$). This did not differ between Compound and Separate groups (Group × Shock Reinforcement: $p > .25$). No other main effects were observed.

Our main hypothesis concerned group differences in responses to the extinguished tone and the non-extinguished square. Specifically, we hypothesized that the Separate group would not generalize extinction (leading to greater reactivity to the non-extinguished square than the extinguished tone), whereas the Compound group would generalize extinction (leading to comparable responses to the tone and the square).

Consistent with this hypothesis, we found that Compound and Separate groups had significantly different SCR to the tone and square cues (Group × Cue Type: $F(1,282) = 4.22$, $p = .04$, $\eta_p^2 = 0.014$; Fig. 3A). Crucially, participants in the Separate group were significantly more reactive to the square than the tone ($b = 0.08$ [0.04], $p = .022$), indicating that extinction did not generalize. In fact, the Separate group was significantly more reactive to the square than the Compound group ($b = 0.2$ [0.096], $p = .039$). In contrast, the Compound group did not significantly differ in their responses to the tone and square ($p > .25$), indicating extinction generalization.

It is worth noting that the Separate group was more reactive to both square cues (CS+ and CS-), rather than specifically reactive to the square previously paired with shock (Group × Cue Type × Shock Reinforcement: $p > .25$). This broad reactivity is especially striking because these participants had shown the opposite pattern during acquisition, with stronger responses to the tone cues than the square cues (3.1.1).

Rather than limited extinction generalization, it was possible that higher reactivity to square cues in the Separate group was instead associated with failure to extinguish the tone/shock association. If this were the case, we would expect that participants who had less successful tone extinction would also be most reactive to the square cues during test. Accordingly, we examined SCR to the CS+ cues at the start of the Day 3 test phase as a function of cue type (tone vs square) and tone extinction success (tone CS+ vs tone CS- at the end of the Day 2 extinction phase). In the Separate group, we found a significant Cue Type × Extinction Success interaction ($F(1,46) = 9.4$, $p = .004$, $\eta_p^2 = 0.15$). Follow-up correlations showed that, while participants who had less successful tone extinction also had stronger responses to the tone CS+ at test ($r(46) = 0.36$, $p = .01$), there was no such correlation between tone extinction and responses to the square CS+ ($p > .25$). Thus, persistent responses to the non-extinguished cue were not driven by poor cue extinction. Indeed, a median split showed that participants who had better cue extinction actually showed stronger evidence for impaired extinction generalization (better extinction, square vs tone: $b = 0.17$ [0.06], $p = .014$; worse extinction: $p > .25$). These analyses demonstrate that failure to generalize extinction between cues in the Separate group was not the result of poor cue extinction.

In addition to SCR, the Compound and Separate groups also differed in their explicit ratings of how much they expected to receive shock at the beginning of the test phase (Fig. 3B). However, unlike SCR (in which the groups broadly differed in their response to square and tone cues regardless of reinforcement history), expectancy ratings were different for CS+ cues (Group × Cue Type: $b = 1.94$ [0.86], $p = .025$) but not CS- cues ($p > .25$). Consistent with SCR, the Separate group was more likely to expect shock from the non-extinguished square than the extinguished tone ($b = 1.81$ [0.65], $p = .004$), whereas the Compound group did not show this distinction ($p > .25$). Together, these analyses provide evidence that extinction generalized between cues following compound but not separate threat learning.

3.2.2. Change in cue responses from acquisition to test (Separate group)

In addition to differential responses to tone and square cues during test, limited extinction generalization in the Separate group can also be observed through the change in responses to CS+ trials from the end of acquisition (pre-extinction of the tone cue) to test (post-extinction; Rashid et al., 2016). If extinction did not generalize, responses to the tone CS+ would decrease (as a result of extinction), but responses to the square CS+ would not be attenuated. To test this, we modeled SCR at the end of acquisition and beginning of test as a function of phase and cue type (tone CS+ vs square CS+).

Consistent with our hypotheses, we found that SCR changed significantly from acquisition to test within the Separate group, and these effects differed for the tone CS+ and square CS+ (Phase × Cue Type: $F(1,141) = 4.59$, $p = .034$, $\eta_p^2 = 0.029$; Fig. 4A). Although these participants showed a trend-level decrease in responses to the tone following extinction ($b = -0.14$ [0.08], $p = .08$), responses to the square did not significantly change ($b = 0.1$ [0.08], $p = .2$).

We observed a similar pattern in ratings of shock expectancy (Fig. 4B). Although there was an overall decrease in explicit ratings

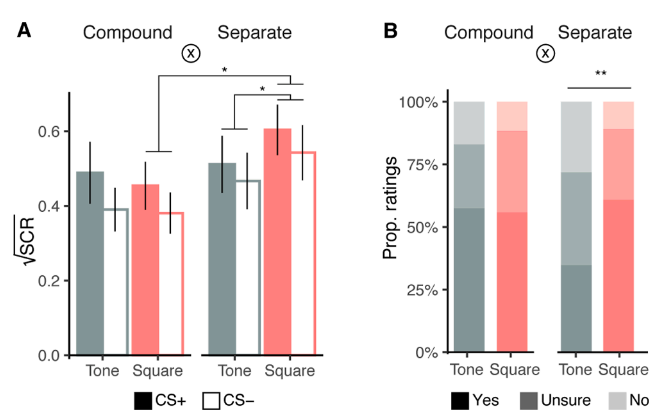

Fig. 3. Extinction generalization between cues varies as a function of separate vs compound threat learning. The Separate group showed higher skin conductance response (SCR, panel A) and expectation of receiving shock (panel B) to the non-extinguished square cues than the extinguished tone cues at the beginning of the test phase. In contrast, the Compound group had comparable responses across cues. *$p < .05$; **$p < .01$.

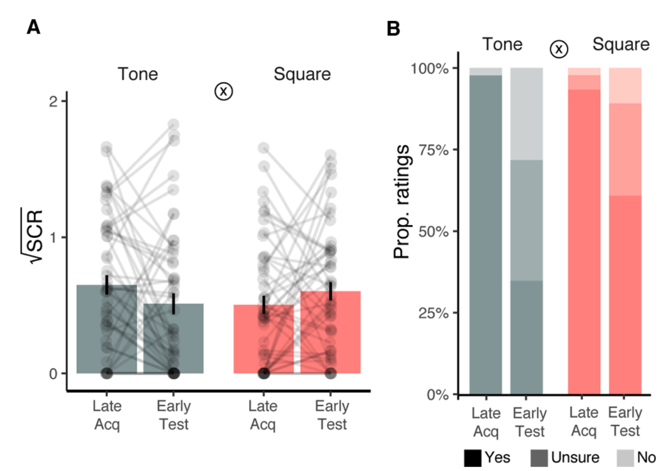

Fig. 4. Learning threat associations with separate cues limits extinction generalization between cues. Between the end of acquisition (pre-extinction) and the beginning of the test phase (post-tone extinction), the Separate group showed a decrease in skin conductance response (SCR; A) and shock expectation (B) to the extinguished tone cue. However, this pattern was not observed for the non-extinguished square cue.

(likely because participants did not receive any shocks during the extinction session; main effect Phase: $b = -2.86$ [0.85], $p < .001$), this drop was larger for the extinguished tone than the non-extinguished square (Phase × Cue Type: $b = 2.83$ [1.38], $p = .04$). Combined with the between-group analyses described above, these results provide further support for impaired extinction generalization between cues following separate threat learning.

3.3. Extinction generalization from component cues to compound

After testing extinction generalization between cues, we next examined whether extinguishing both cues would generalize to the compound as follows. On Day 3, participants were presented with individual tone and square cues with no reinforcement (analyzed above). Then, participants were repeatedly presented with the tone and square cues, constituting "re-extinction" for the tone (as these had been presented without reinforcement during the Day 2 extinction session), and extinction for the square. Finally, participants were presented with the tone/square compounds without reinforcement (tone CS+ and square CS+ together; tone CS- and square CS- together). These were the same compounds that the Compound group experienced during threat acquisition.

We first confirmed that conditioned responses to individual cues were successfully extinguished. To test this, we again separated trials by cue type (tone vs. square) and reinforcement history (CS+ vs CS-), but instead focused on the last trial of Day 3 cue extinction. Unlike the beginning of Day 3 (see 3.2.1), there was no longer a significant main effect of prior shock reinforcement ($p > .25$). However, this conditioned response significantly differed between groups (Group × Shock Reinforcement: $F(1,282) = 6.995$, $p = .009$, $\eta_p^2 = 0.023$). This was driven by differences in responses to the CS-: the Separate group had significantly stronger responses to the CS- cues relative to the Compound group ($b = 0.17$ [0.08], $p = .046$), and had marginally higher responses to CS- than CS+ cues ($b = -0.09$ [0.048], $p = .057$). In contrast, the Compound group had marginally higher responses to the CS+ than CS- cues ($b = 0.09$ [0.048], $p = .068$). Follow-up analyses examining tone and square cues separately showed that, within the Compound group, responses to CS+ and CS- cues were not significantly different (square CS+ vs CS-: $b = 0.1$ [0.06], $p = .11$; tone CS+ vs CS-: $p > .25$), indicating that the individual cue associations were successfully extinguished. Nevertheless, we ran tests to determine whether extinction generalization from component cues to compounds was influenced by variability in cue extinction (see 3.3.1 below).

3.3.1. Responses to compounds during test phase

As with the analysis of individual cues, we examined responses to the first two compound presentations on Day 3, separating trials by group and whether they had been paired with shock (or, in the case of the Separate group, whether they contained cues that had been paired with shock). Due to a technical error, one participant was not exposed to compound cues and is not included in these analyses.

We hypothesized that extinction generalization from individual cues to a compound would show a different pattern to that observed for extinction generalization between cues. Drawing on learning models positing that compound learning involves additional associations beyond those between individual cues and the aversive outcome, we hypothesized that extinction generalization from cues to compound would be attenuated following compound threat learning. In contrast,

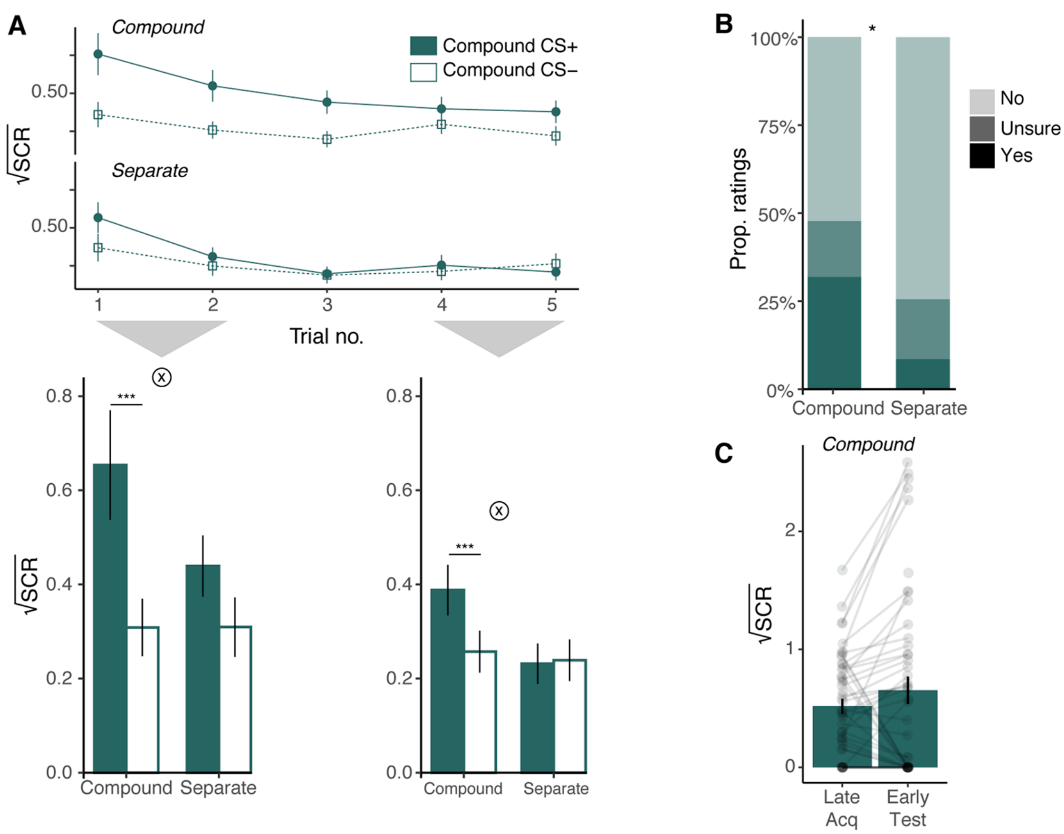

Fig. 5. Extinction generalization from cues to compound varies as a function of separate or compound threat learning. (A) Skin conductance responses (SCR) to the compound previously associated with shock (CS+) and the compound never associated with shock (CS-). Even after extinction of both component cues, the Compound group showed a higher response to the compound CS+ than the compound CS- that persisted over repeated exposures. (B) On the first presentation of the compound CS+, the Compound group was more likely to expect shock than the Separate group. (C) For participants in the Compound group, responses to the previously reinforced compound did not decrease between the end of acquisition (pre-extinction) and the test phase (post-extinction of both component cues; compare to Fig. 4A). *$p < .05$, ***$p < .001$.

we hypothesized that extinction would successfully generalize from cues to compound following separate threat learning.

Consistent with our hypothesis, we found that participants in the Separate and Compound groups showed significantly different responses to the previously reinforced and nonreinforced compound (Group × Shock Reinforcement: $F(1,93) = 4.24$, $p = .042$, $\eta_p^2 = 0.035$; Fig. 5A). Participants in the Compound group reacted significantly more strongly to the compound CS+ vs. CS- ($b = 0.35$ [0.07], $p < .001$), whereas this was attenuated in the Separate group ($b = 0.13$ [0.07], $p = .08$). Compared to the Separate group, the Compound group was also more likely to expect shock on the first presentation of the previously reinforced compound ($b = 1.6$ [0.79], $p = .043$; Fig. 5B).

The difference in SCR between the Separate and Compound groups persisted throughout extinction of the compound (Group × Shock Reinforcement: $F(1,849) = 13.43$, $p < .001$, $\eta_p^2 = 0.015$; Fig. 5A). Indeed, the Compound group appeared not to successfully extinguish the compound, continuing to show an elevated response to the compound previously paired with shock throughout repeated presentations ($b = 0.24$ [0.04], $p < .001$) whereas the Separate group did not show this ($b = 0.05$ [0.04], $p = .16$). Even in the last two trials of compound presentation, the difference between groups persisted (Group × Shock Reinforcement: $F(1,93) = 4.62$, $p = .034$, $\eta_p^2 = 0.039$), with the Compound group showing a persistently higher response to the compound that had been paired with shock ($b = 0.13$ [0.05], $p = .005$). These results show that, following compound threat learning, extinction of component cues did not generalize to the compound.

Follow-up analyses demonstrated that limitations in extinction generalization in the Compound group were not driven by impairments in extinction of the two component cues (average of conditioned responses to tone and square at end of Day 3 extinction). One participant was an outlier for cue extinction (>3SD outside mean for tone and > 2SD outside mean for square) and was not included in these analyses. Extinction success did not explain levels of response to the compound previously paired with shock (Shock Reinforcement × Extinction Success: $p > .25$; including outlier: $p = .12$). A median split showed that, regardless of extinction success, participants in the Compound group showed significantly higher responses to the compound CS+ relative to the compound CS- (better cue extinction: $b = 0.27$ [0.11], $p = .02$; worse cue extinction: $b = 0.4$ [0.16], $p = .02$).

3.3.2. Changes in compound responses from acquisition to test (Compound group)

We next examined whether responses to the compound changed from acquisition (prior to extinction of component cues) to test (post-extinction) within the Compound group. We found no significant change in response to the shock-paired compound, even following extinction of both elements (Fig. 5C; $F(1,45) = 2.34$, $p = .13$). In fact, the majority of participants (69.6%) showed either no change or an increase in their response. These data further support the interpretation that extinction of component cues did not generalize to the compound following compound threat learning.

3.4. Comparing responses to cues and compounds across learning groups

The above analyses separately examined extinction generalization between cues and from component cues to a full compound. They revealed impaired extinction generalization between cues following separate learning, and impaired extinction generalization from cues to compound following compound learning. Another way to ask whether threat learning had opposite effects on these two forms of extinction generalization would be to directly compare generalization to the non-extinguished cue (square) and the non-extinguished compound. To conduct this confirmatory analysis, we examined SCR to the first presentations of the square CS+ and the first presentations of the compound CS+ during Day 3 for the Separate and Compound groups.

Consistent with prior analyses, this model revealed a significant Cue Type by Group interaction ($F(1,93) = 13.87$, $p < .001$, $\eta_p^2 = 0.11$, Fig. 6). There was a marginal difference in responses between groups, with higher reactivity to the compound in the Compound than Separate groups ($b = 0.22$ [0.12], $p = .075$). In the Separate group, responses to the compound were significantly lower than the square, consistent with better extinction generalization from cues to compound than from cue to cue ($b = -0.17$ [0.07], $p = .02$). In contrast, for the Compound group, responses to the compound were significantly higher than the square ($b = 0.2$ [0.07], $p = .005$), consistent with worse extinction generalization from cues to compound than from cue to cue.

4. Discussion

In this study, we demonstrated that learning to associate threat with a compound or with individual cues had opposite effects on the success of later extinction generalization. We assessed extinction generalization in two ways. First, we tested whether extinguishing threat associations with one cue (tone) would generalize and attenuate responses to another cue (square). This phenomena of extinction generalization between cues has been shown to vary in rodent models based on whether threats were initially paired with a compound or separate cues (Debiec et al., 2013). Next, we tested whether extinguishing associations with both cues separately would generalize to a combination of these cues (tone + square compound). We found that participants who formed a compound threat association showed evidence of successful extinction generalization between component cues, but not between component cues and the full compound. In contrast, participants who formed separate threat associations did not show extinction generalization between cues, but did generalize from component cues to the compound. These patterns were consistent across conditioned responses (measured using skin conductance) and explicit ratings of whether shocks were expected on each trial.

4.1. Compound, but not separate, threat learning enabled extinction generalization between component cues

Consistent with past work in nonhuman animals, we found that elemental threat learning blocked extinction generalization between cues, with persistently higher responses to the non-extinguished cue. Participants who associated threat with a compound showed a significantly different pattern, with comparable responses across both non-extinguished and extinguished cues (Debiec et al., 2013; Durlach & Rescorla, 1980; Nakajima & Kawai, 1997; Pineno, 2007; Schnelker & Batsell, 2006).

Compared to the Compound group, participants in the Separate

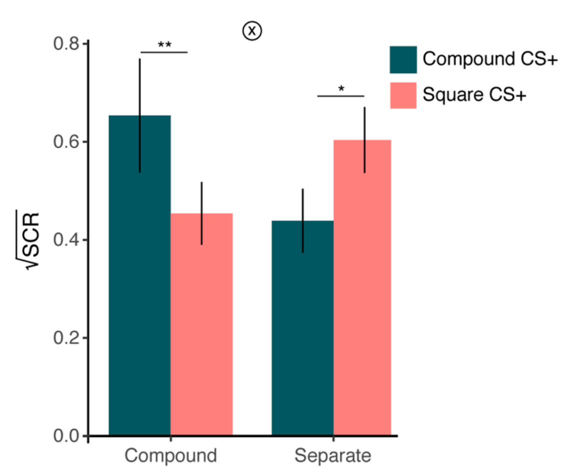

Fig. 6. Comparing magnitude of extinction generalization from one cue to another cue (square CS+) and from two component cues to a combination of cues (compound CS+). *$p < .05$, **$p < .01$.

group showed significantly higher responses to the non-extinguished (square) cue. However, this was not limited to the square cue that had previously been paired with shock, perhaps indicative of sensitization. While this non-associative process is often associated with fear-relevant stimuli (Dunsmoor, Mitroff, & LaBar, 2009; Ohman & Mineka, 2001), as opposed to the fear-irrelevant stimuli of squares and tones in the current study, it is possible that the Separate group became sensitized as a result of receiving more shock (US) exposure than the Compound group. This process would result in broadly elevated SCR to any threat or safety stimulus (discussed in Haddad, Pritchett, Lissek, & Lau, 2012). However, participants in the Separate group were not indiscriminately reactive – instead, although they responded to both the square CS+ and CS-, these were significantly higher than their own responses to tone or compound cues (CS+ or CS-), arguing against sensitization. Furthermore, analyses of the change in responses from acquisition to test in the Separate group demonstrated a decrease in responding to the (extinguished) tone CS+ but no change in response to the square CS+, which also indicates specific associative learning changes. Previous studies using similar levels of shock exposure also provide evidence that this level of shock exposure is unlikely to result in sensitization. For example, Willems and Vervliet found that, far from broadly elevated reactivity, participants showed distinct responses based on expectation of shock (Willems & Vervliet, 2021). Nevertheless, the difference in US exposure between groups (designed to create equivalent learning of cue/shock associations in the Separate group and compound/shock associations in the Compound group) is a limitation of this design and the precedent in rodents (Debiec et al., 2013). Further work is needed to determine whether this distinct US history has an impact on the observed between-group differences in responses to the non-extinguished cues at test.

It is worth noting that the success of extinction generalization between cues in the Compound group was limited, as these participants still showed significant conditioned responses (i.e., responded more strongly to the CS+ than CS-) after extinction of the tone cue. Such spontaneous recovery is commonly observed in traditional extinction procedures (Dunsmoor et al., 2015) and demonstrates that compound learners did acquire some threat association with the tone and square cues individually (Jones et al., 2013), rather than only associating a unique tone/square compound with shock (Pearce, 2002). However, it also highlights limitations in extinction. Thus, it is possible that improved cue extinction procedures could enhance extinction generalization between cues following compound threat learning, in line with rodent studies showing complete attenuation of conditioned responses to both extinguished and non-extinguished cues (e.g., Debiec et al., 2013).

Our findings of impaired extinction generalization following elemental threat learning are inconsistent with some recent reports. These studies demonstrated that, when separate threat associations were learned close in time, they were integrated into a combined representation (Cai et al., 2016; Rashid et al., 2016). Accordingly, extinction generalized between cues (Rashid et al., 2016). Although participants in our study also learned separate associations close in time, one important difference is that we used cues from different modalities, auditory and visual. The studies demonstrating integration across associations leveraged cues from the same modality (two contexts or two tones). In contrast, intermixing auditory and visual modalities blocked extinction generalization (Debiec et al., 2013), although interactions with temporal context are complex (Vurbic & Bouton, 2011). Thus, it is possible that, in our study, participants may have shown greater extinction generalization if they had been conditioned to two tones rather than a tone and a square. Using stimuli from different modalities also raises the possibility of differing salience between the cues, which can influence the extent of extinction generalization (Jones et al., 2013). Further work is needed to determine whether extinguishing the square would generalize differently compared to extinguishing the tone (Mertens et al., 2019; Trost & Batsell, 2004).

4.2. *Separate, but not compound, threat learning enables extinction generalization from cues to compound*

Among participants who formed threat associations with a compound, there was a striking deficit in generalization of extinction from cues to the full compound. Although this has been studied less than extinction generalization between cues, this finding is consistent with prior reports in rodents showing weak extinction generalization from cues to compounds (Jones et al., 2013; Shanks et al., 1998; Troisi et al., 2013).

It is possible that the fear relapse to the compound observed in the Compound, but not Separate group, is simply due to the fact that the Compound group had a history of reinforcement with this combined stimulus whereas the Separate group did not – that is, the difference was driven by conditioning rather than extinction generalization. This limitation has been noted in other recent work examining the influences of different types of threat learning on extinction generalization: further studies are needed to determine whether group-level differences in extinction generalization are purely the result of generalization of extinction, differences in threat learning, or both (Wong et al., 2020). Notably, findings within the Compound group provide evidence supporting limited extinction generalization. As all participants in this group underwent the same threat learning procedure, this suggests that the observed effects are not solely attributable to differences in threat learning. Even after extinguishing all component elements (one of which was extinguished twice), participants in the Compound group showed a persistently elevated response to the compound previously paired with shock. In fact, responses to the compound changed relatively little between the end of acquisition (when the compound was actively being reinforced with shock) and test, showing that the intervening extinction phase did not significantly alter responses. Furthermore, by examining how successfully participants extinguished individual cues, we found that even strong extinction of individual cues was inadequate to attenuate responses to the compound. Finally, far from showing attenuated responses that would be consistent with extinction generalization, responses to the compound remained significantly elevated even after repeated nonreinforced presentations.

This failure to generalize from cues to compound after learning compound/shock associations is consistent with learning models postulating that unique associations are formed between the compound and the outcome which are not addressed by extinction of individual cues. Some models suggest further that extinction of an individual cue engages a new configural unit, leaving the association attached to the full compound intact (Bouton et al., 2012). However, this explanation was used to account for limited generalization of extinction from one cue to a compound; it is not clear why extinction of both component cues would seemingly fail to generalize. Indeed, both configural (Pearce, 2002) and latent cause models (Gershman, Norman, & Niv, 2015) suggest that the similarity between the extinguished cues and the conditioned compound should facilitate some generalization. That is, as the tone and the square are each a 50% match to the tone/square compound, some of the extinction learning should spread to the compound. Further work is needed to replicate the current empirical findings demonstrating poor cue/compound extinction generalization and incorporate them into a learning framework.

As noted above for generalization between cues, it is possible that enhanced extinction procedures (e.g., sequential extinction) or using stimuli of equivalent salience may improve extinction generalization from cues to compounds after learning to associate threat with a compound (Jones et al., 2013). Interestingly, Jones et al. (2013) found that an enhanced extinction procedure (retrieval + extinction) generalized more strongly from tone to compound than from light to compound in rodents. Based on this finding, it is noteworthy that the current study, leveraging two rounds of tone extinction (Days 2 and 3), did not generalize to the compound. Furthermore, creating a compound composed of stimuli from different modalities may have led to inference of different latent causes (Gershman et al., 2015) and facilitated retrospective revaluation (Liljeholm & Balleine, 2009; Miller & Witnauer, 2016). Further work is needed to test whether extinction generalization within the Compound group would be facilitated if the compound was composed of two stimuli from the same modality.

Despite deficits in extinction generalization between cues, these findings highlight a novel advantage for elemental threat learning in later extinction generalization to combinations of cues. Representing complex aversive events as amygdala-dependent elemental threat associations is often considered to be maladaptive, enabling threat responses to occur in contexts that are not actually dangerous (Acheson et al., 2012). Here we show that, after extinction of all component cues, participants who formed separate threat associations actually showed more adaptive responses to a multi-cue event than participants who formed compound threat associations. As elemental threat learning has been proposed as a potential contributor to PTSD symptomatology, these findings support the efficacy of clinical approaches that expose patients to many components of aversive experiences.

5. Conclusion

The results presented here demonstrate the significant – and opposite – effects of separate and compound threat representations on later extinction generalization. These findings provide further evidence of distinct properties for learning about individual cues and compounds (Urushihara, Stout, & Miller, 2004), and suggest that patterns of extinction generalization may provide a useful tool to gain insight into what threat associations were formed. Behavioral dissociations of elemental and configural learning in nonhuman animals have enabled important neurobiological insights into the mechanisms underlying these processes (Honey, Iordanova, & Good, 2014). Thus, examining extinction generalization may delineate the contributions of factors, such as attentional scope (Byrom & Murphy, 2016) and stress (Simon-Kutscher, Wanke, Hiller, & Schwabe, 2019), to the relative formation of configural and elemental memories in humans. More broadly, these results highlight limits of extinction generalization that raise important questions for therapeutic interventions targeting complex aversive experiences.

CRediT authorship contribution statement

Elizabeth V. Goldfarb: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Visualization. **Tahj Blow:** Investigation, Data curation. **Joseph E. Dunsmoor:** Conceptualization, Resources, Writing - review & editing. **Elizabeth A. Phelps:** Conceptualization, Methodology, Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Institutes of Health (R01 MH097085 to E.A.P., KL2 TR001862 and K01 AA027832 to E.V.G.) and the James S. McDonnell Foundation (E.A.P.).

References

Acheson, D. T., Gresack, J. E., & Risbrough, V. B. (2012). Hippocampal dysfunction effects on context memory: Possible etiology for posttraumatic stress disorder. *Neuropharmacology, 62*(2), 674–685. https://doi.org/10.1016/j.neuropharm.2011.04.029.

Bisby, J. A., Burgess, N., & Brewin, C. R. (2020). Reduced memory coherence for negative events and its relationship to posttraumatic stress disorder. *Current Directions in Psychological Science, 29*(3), 267–272.

Bouton, M. E., Doyle-Burr, C., & Vurbic, D. (2012). Asymmetrical generalization of conditioning and extinction from compound to element and element to compound. *Journal of Experimental Psychology: Animal Behavior Processes, 38*(4), 381–393. https://doi.org/10.1037/a0029726.

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. Behavioural Brain Research, 110, 67–72.

Brewin, C. R. (2014). Episodic memory, perceptual memory, and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychological Bulletin, 140*(1), 69–97. https://doi.org/10.1037/a0033722.

Buhr, K., & Dugas, M. J. (2002). The Intolerance of Uncertainty Scale: Psychometric properties of the English version. *Behaviour Research and Therapy, 40*(8), 931–945. https://doi.org/10.1016/s0005-7967(01)00092-4.

Byrom, N. C., & Murphy, R. A. (2016). Individual difference in configural associative learning. *Journal of Experimental Psychology: Animal Learning and Cognition, 42*(4), 325–335. https://doi.org/10.1037/xan0000111.

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., … Silva, A. J. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature, 534*(7605), 115–118. https://doi.org/10.1038/nature17955.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. Retrieved from *Journal of Health and Social Behavior, 24*(4), 385–396 https://www.ncbi.nlm.nih.gov/pubmed/6668417.

de Voogd, L. D., & Phelps, E. A. (2020). A cognitively demanding working-memory intervention enhances extinction Research article. *Scientific Reports, 10*(1), 7020. https://doi.org/10.1038/s41598-020-63811-0.

Debiec, J., Diaz-Mataix, L., Bush, D. E., Doyere, V., & LeDoux, J. E. (2013). The selectivity of aversive memory reconsolidation and extinction processes depends on the initial encoding of the Pavlovian association. *Learning & Memory, 20*(12), 695–699. https://doi.org/10.1101/lm.031609.113.

Dubin, W. J., & Levis, D. J. (1973). Generalization of extinction gradients: A systematic analysis. *Journal of Experimental Psychology: General, 100*(2), 403–412.

Dunsmoor, J. E., Mitroff, S. R., & LaBar, K. S. (2009). Generalization of conditioned fear along a dimension of increasing fear intensity. *Learning & Memory, 16*(7), 460–469. https://doi.org/10.1101/lm.1431609.

Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking Extinction. *Neuron, 88*(1), 47–63. https://doi.org/10.1016/j.neuron.2015.09.028.

Durlach, P. J., & Rescorla, R. A. (1980). Potentiation rather than overshadowing in flavor-aversion learning: An analysis in terms of within-compound associations. *Journal of Experimental Psychology: Animal Behavior Processes, 6*(2), 175–187.

Fitzgerald, P. J., Seemann, J. R., & Maren, S. (2014). Can fear extinction be enhanced? A review of pharmacological and behavioral findings. *Brain Research Bulletin, 105*, 46–60. https://doi.org/10.1016/j.brainresbull.2013.12.007.

Gershman, S. J., Norman, K. A., & Niv, Y. (2015). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences, 5*, 43–50.

Gilbertson, M. W., Williston, S. K., Paulus, L. A., Lasko, N. B., Gurvits, T. V., Shenton, M. E., … Orr, S. P. (2007). Configural cue performance in identical twins discordant for posttraumatic stress disorder: Theoretical implications for the role of hippocampal function. *Biological Psychiatry, 62*(5), 513–520. https://doi.org/10.1016/j.biopsych.2006.12.023.

Green, S. R., Kragel, P. A., Fecteau, M. E., & LaBar, K. S. (2014). Development and validation of an unsupervised scoring system (Autonomate) for skin conductance response analysis. *International Journal of Psychophysiology, 91*(3), 186–193. https://doi.org/10.1016/j.ijpsycho.2013.10.015.

Haddad, A. D. M., Pritchett, D., Lissek, S., & Lau, J. Y. F. (2012). Trait Anxiety and Fear Responses to Safety Cues: Stimulus Generalization or Sensitization? *Journal of Psychopathology and Behavioral Assessment, 34*(3), 323–331. https://doi.org/10.1007/s10862-012-9284-7.

Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review, 113*(3), 584–605. https://doi.org/10.1037/0033-295X.113.3.584.

Hennings, A. C., Bibb, S. A., Lewis-Peacock, J. A., & Dunsmoor, J. E. (2020). Thought suppression inhibits the generalization of fear extinction. *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2020.112931.

Honey, R. C., Iordanova, M. D., & Good, M. (2014). Associative structures in animal learning: Dissociating elemental and configural processes. *Neurobiology of Learning and Memory, 108*, 96–103. https://doi.org/10.1016/j.nlm.2013.06.002.

Jones, C. E., Ringuet, S., & Monfils, M. H. (2013). Learned together, extinguished apart: Reducing fear to complex stimuli. *Learning & Memory, 20*(12), 674–685. https://doi.org/10.1101/lm.031740.113.

Kasprow, W. J., Schachtman, T. R., Cacheiro, H., & Miller, R. R. (1984). Extinction does not depend upon degradation of event memories. *Bulletin of the Psychonomic Society, 22*(2), 95–98.

Keller, N. E., & Dunsmoor, J. E. (2020). The effects of aversive-to-appetitive counterconditioning on implicit and explicit fear memory. *Learning & Memory, 27*(1), 12–19. https://doi.org/10.1101/lm.050740.119.

Kredlow, M. A., de Voogd, L., & Phelps, E. A. (2020). Laboratory Analogues and Therapy Procedures: A Case for Translation from the Clinic to the Laboratory. *PsyArXiv*. https://doi.org/10.31234/osf.io/7r6p2.

Liljeholm, M., & Balleine, B. W. (2009). Mediated conditioning versus retrospective revaluation in humans: The influence of physical and functional similarity of cues. *Quarterly Journal of Experimental Psychology (Hove), 62*(3), 470–482. https://doi.org/10.1080/17470210802008805.

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: Implications for fear conditioning, extinction and psychopathology. *Nature Reviews Neuroscience, 14*(6), 417–428. https://doi.org/10.1038/nrn3492.

Melchers, K. G., Shanks, D. R., & Lachnit, H. (2008). Stimulus coding in human associative learning: flexible representations of parts and wholes. Behaviour Processes, 77(3), 413–427; discussion 451–413. http://doi.org/10.1016/j.beproc.2007.09.013.

Mertens, G., Leer, A., van Dis, E. A. M., Vermeer, L., Steenhuizen, A., van der Veen, L., & Engelhard, I. M. (2019). Secondary extinction reduces reinstatement of threat expectancy and conditioned skin conductance responses in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry, 62*, 103–111. https://doi.org/10.1016/j.jbtep.2018.09.007.

Miller, R. R., & Witnauer, J. E. (2016). Retrospective revaluation: The phenomenon and its theoretical implications. *Behavioural Processes, 123*, 15–25. https://doi.org/10.1016/j.beproc.2015.09.001.

Nakajima, S., & Kawai, N. (1997). Failure of retrospective inference in rats' taste aversion. *Japanese Psychological Research, 39*(2), 87–97.

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. https://doi.org/10.1037/0033-295x.108.3.483.

Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. Animal Learning Behaviour, 30(2), 73–95.

Pineno, O. (2007). An examination of the effectiveness of inflation and deflation treatments in detecting within-compound learning of a taste aversion. *Behavioural Processes, 75*(1), 33–39. https://doi.org/10.1016/j.beproc.2007.01.004.

Rashid, A. J., Yan, C., Mercaldo, V., Hsiang, H. L., Park, S., Cole, C. J., … Josselyn, S. A. (2016). Competition between engrams influences fear memory formation and recall. *Science, 353*(6297), 383–387. https://doi.org/10.1126/science.aaf0594.

Rescorla, R. A., & Cunningham, C. L. (1978). Within-compound flavor associations. *Journal of Experimental Psychology: Animal Behavior Processes, 4*(3), 267–275. https://doi.org/10.1037//0097-7403.4.3.267.

Rowe, M. K., & Craske, M. G. (1998). Effects of varied-stimulus exposure training on fear reduction and return of fear. *Behaviour Research and Therapy, 36*(7–8), 719–734. https://doi.org/10.1016/s0005-7967(97)10017-1.

Rudy, J. W., Huff, N. C., & Matus-Amat, P. (2004). Understanding contextual fear conditioning: Insights from a two-process model. *Neuroscience & Biobehavioral Reviews, 28*(7), 675–685. https://doi.org/10.1016/j.neubiorev.2004.09.004.

Schnelker, J., & Batsell, W. R., Jr. (2006). Within-compound associations are not sufficient to produce taste-mediated odor potentiation. *Behavioural Processes, 73*(2), 142–148. https://doi.org/10.1016/j.beproc.2006.04.008.

Shanks, D. R., Darby, R. J., & Charles, D. (1998). Resistance to Interference in Human Associative Learning: Evidence of Configural Processing. *Journal of Experimental Psychology: Animal Behavior Processes, 24*(2), 136–150.

Simon-Kutscher, K., Wanke, N., Hiller, C., & Schwabe, L. (2019). Fear Without Context: Acute Stress Modulates the Balance of Cue-Dependent and Contextual Fear Learning. *Psychological Science, 30*(8), 1123–1135. https://doi.org/10.1177/0956797619852027.

Soto, F. A., Gershman, S. J., & Niv, Y. (2014). Explaining compound generalization in associative and causal learning through rational principles of dimensional generalization. *Psychological Review, 121*(3), 526–558. https://doi.org/10.1037/a0037018.

Spielberger, C. D. (1983). *State-Trait Anxiety Inventory for Adults*. Palo Alto, CA: Consulting Psychologists Press.

Stout, D. M., Glenn, D. E., Acheson, D. T., Spadoni, A. D., Risbrough, V. B., & Simmons, A. N. (2018). Neural measures associated with configural threat acquisition. *Neurobiology of Learning and Memory, 150*, 99–106. https://doi.org/10.1016/j.nlm.2018.03.012.

Troisi, J. R., 2nd, Dooley, T. F., 2nd, & Craig, E. M. (2013). The discriminative stimulus effects of a nicotine-ethanol compound in rats: Extinction with the parts differs from the whole. *Behavioral Neuroscience, 127*(6), 899–912. https://doi.org/10.1037/a0034824.

Trost, C. A., & Batsell, W. R., Jr. (2004). Taste + odor interactions in compound aversion conditioning. *Learning & Behavior, 32*(4), 440–453.

Urushihara, K., Stout, S. C., & Miller, R. R. (2004). The basic laws of conditioning differ for elemental cues and cues trained in compound. *Psychological Science, 15*(4), 268–271. https://doi.org/10.1111/j.0956-7976.2004.00664.x.

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory, 11*(5), 555–558. https://doi.org/10.1101/lm.77404.

Vurbic, D., & Bouton, M. E. (2011). Secondary extinction in Pavlovian fear conditioning. *Learning & Behavior, 39*(3), 202–211. https://doi.org/10.3758/s13420-011-0017-7.

Willems, A. L., & Vervliet, B. (2021). When nothing matters: Assessing markers of expectancy violation during omissions of threat. *Behaviour Research and Therapy, 136*, Article 103764. https://doi.org/10.1016/j.brat.2020.103764.

Wong, A. H. K., Glück, V. M., Boschet, J. M., & Engelke, P. (2020). Generalization of extinction with a generalization stimulus is determined by learnt threat beliefs. *Behaviour Research and Therapy*. https://doi.org/10.1016/j.brat.2020.103755.