Enhancing extinction learning: Occasional presentations of the unconditioned stimulus during extinction eliminate spontaneous recovery, but not necessarily reacquisition of fear

Alina Thompson, [1] Peter M. McEvoy, [1,2] & Ottmar V. Lipp[1,3]

[1]School of Psychology, Curtin University, Kent Street, Bentley, WA 6102, Australia

[2] Centre for Clinical Interventions, 223 James Street, Northbridge, WA 6003, Australia

[3] ARC-SRI: Science of Learning Research Centre, University of Queensland, Brisbane, QLD 4072, Australia

Author Notes

Address for correspondence:

Alina Thompson, School of Psychology, Curtin University, GPO Box U1987 Perth WA 6845, Australia. Email: alina.thompson@curtin.edu.au

Words: 9172

Abstract

Background: Fears underlying anxiety disorders are commonly treated with exposure-based therapies, which are based on the principles of extinction learning. While these treatments are efficacious, fears may return after successful treatment. Past research suggested that post-extinction recovery of fear could be reduced through extinction training that involves occasional presentations of the aversive unconditioned stimulus (US), paired with the conditioned stimulus (CS). Here, we examined whether extinction training with occasionally paired or unpaired US presentations is superior in the reduction of fear recovery to non-reinforced extinction. **Method**: Following differential fear conditioning to neutral cues, participants (*N*=72; *M age*=21.61 years, *SD*=3.95) underwent either non-reinforced, partially reinforced, or unpaired extinction training. **Results**: Extinction involving paired or unpaired US presentations, but not non-reinforced extinction, eliminated spontaneous recovery of differential skin conductance responses (SCRs). Results further suggested that unpaired, but not paired, US presentations may guard against rapid reacquisition of differential SCRs. No benefits of US presentations during extinction were found on the reinstatement of SCRs or recovery of differential negative CS+ valence. **Conclusion**: Presenting USs during extinction training was more effective than non-reinforced extinction in the reduction of fear recovery, as indexed by SCRs, with unpaired extinction being more effective than partially reinforced extinction.

Key words: fear conditioning; occasionally reinforced extinction; partially reinforced extinction; unpaired extinction; reacquisition; spontaneous recovery; reinstatement; return of fear

Introduction

Past research has provided us with a good understanding of mechanisms underlying the development and reduction of fears, phobias, and anxiety disorders. Fears are acquired through association of neutral cues (conditioned stimuli, CSs), such as animals, with aversive outcomes (unconditioned stimuli [USs]; Davey, 1992), such as an animal bite. Through CS-US pairings we learn to predict which cues signal the arrival of aversive and potentially threatening events. While learning to fear cues which may pose a threat to our survival is an important adaptive mechanism that can protect us from harm and facilitate survival (Öhman & Mineka, 2001), fears may also become maladaptive and contribute to the development of anxiety and stress disorders, which can interfere with daily functioning (Foa & McLean, 2016). The current global prevalence rate of anxiety disorders is estimated at 7.3%, with approximately 11.6% of the population experiencing an anxiety disorder in a given year (Baxter, Scott, Vos, & Whiteford, 2013; Craske & Stein, 2016). Anxiety disorders are commonly treated with exposure-based therapies, which are based on the principles of extinction learning (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). In its basic form, extinction training involves the repeated presentation of the CS, in the absence of the US, until a reduction of fear is achieved (Bouton, 2000).

While the efficacy of exposure therapies is well established, not all individuals respond to these treatments, while others experience a return of fear after successful treatment (Craske & Mystkowski, 2006; Weisman & Rodebaugh, 2018). Research suggests that extinguished fear may return, because extinction training does not result in the unlearning or persistent elimination of the original fear learning (i.e. the CS-US association), but creates a new, inhibitory association (CS-no US) that co-exists with the fear association (Bouton, 1993). As such, future CS presentations may activate the CS-no US or CS-US association, whereby the latter would allow for a return of fear (Bouton, 1993). Recovery from extinction phenomena are well-documented in the conditioning literature and include recovery of extinguished responding in a new context (renewal), after the passage of time (spontaneous recovery), after the unsignaled presentation of the US (reinstatement), or after additional post-extinction CS-US pairings (reacquisition; Bouton, 2002). Findings from animal research indicate that reacquisition after extinction may occur at a faster rate than de novo conditioning (Napier, Macrae, & Kehoe, 1992; but see Ricker & Bouton, 1996), suggesting that the original fear learning is preserved during extinction and, thereby, may be retrieved through future cue encounters. Taken together, research has identified several pathways that may result in the return of fear following successful extinction training or the successful completion of exposure therapy.

When applied to an example of relapse in the clinical setting, for instance the return of social anxiety, which is characterized by fear of social situations in which individuals may be exposed to rejection, embarrassment, or negative evaluations by others (American Psychiatric Association, 2013), fear may recover when individuals are re-exposed to previously feared social situations (CS) or through exposure to additional CS-US pairings (reacquisition), such as receiving negative feedback (US) during a meeting at work (CS). Given the frequency with which feared cues and outcomes may be encountered in daily life, whether in a paired (CS-US) or unpaired manner (CS or US), the likelihood of fear recovery appears high - this may seem discouraging from a clinical point of view. However, recent evidence suggests that exposure therapy may be optimized in a way that would minimize recovery of extinguished fear, even in light of occasional post-extinction CS-US pairings.

A method of exposure therapy for reducing the return of fear proposed by Craske et al. (2014) involves *occasionally reinforced extinction*, meaning intentionally exposing clients to occasional presentations of the feared event (US) during exposure therapy. In the case of social anxiety, this may involve the delivery of rejection or “shame attacks” during exposure to social situations (Craske et al., 2014). While this idea appears counterintuitive as CS-US pairings are implicated in fear acquisition (Davey, 1992), extant literature suggests that occasional presentations of the US during extinction training may be superior to conventional, non-reinforced extinction in preventing recovery of extinguished responding (e.g. Bouton, Woods, & Pineño, 2004; Culver, Stevens, Fanselow, & Craske, 2018).

Specifically, experiments conducted with animal subjects demonstrated that partially reinforced extinction training, involving occasional delivery of CS-US pairings, interfered with the reacquisition of extinguished responding in appetitive (Bouton et al., 2004) and operant conditioning preparations (Woods & Bouton, 2007). Of particular interest was the observation that partially reinforced extinction slowed the reduction of responding during extinction, as would be expected from reinforced training, but protected against rapid reacquisition, relative to non-reinforced extinction. Additionally, compared to partially reinforced training, an unpaired extinction procedure, whereby reinforcers were not paired with the CS, but instead delivered in the inter-trial interval, further reduced the rate of reacquisition (Bouton et al., 2004 [Experiment 2]). Replication attempts with humans, however, have yielded mixed results in an appetitive conditioning study (van den Akker, Havermans, & Jansen, 2015), showing reduced reacquisition of US expectancies, but not self-rated *conditioned desires* for chocolate mousse, subsequent to partially reinforced and unpaired extinction training. That being said, the authors also reported group differences at baseline and differential effects of acquisition training on verbal (e.g. US expectancy) and physiological indices of conditioned responding (i.e. participants' rate of salivation in anticipation of food), making the overall interpretation of findings difficult.

An extension of Bouton et al.'s (2004) findings to human fear conditioning, on the other hand, has yielded more promising results, suggesting that partially reinforced extinction may successfully reduce the reacquisition of extinguished fear responses (Culver et al., 2018). Following differential fear conditioning to neutral cues, participants underwent either non-reinforced or partially reinforced extinction training. Similar to Bouton and colleagues' work, a 2:8 reinforcement schedule was used during extinction in the partially reinforced group, translating to six reinforced and 18 non-reinforced CS+ trials and 24 non-reinforced CS- trials. Tests of fear recovery showed that partially reinforced extinction training, relative to non-reinforced extinction, interfered with subsequent reacquisition of conditioned fear, as indexed by electrodermal responding. An aspect requiring further investigation, however, is the effect of partially reinforced extinction on the spontaneous recovery of extinguished fear. While Culver and colleagues observed reduced recovery of electrodermal responding to the CS+ after partially reinforced extinction, relative to non-reinforced extinction, these results must be interpreted with caution, as conditioned responding failed to extinguish during partially reinforced extinction training and, consequently, could not "recover." Nevertheless, the results of the reacquisition test provide evidence for cross-species applicability of partially reinforced extinction. The aim of the present study was to replicate and extend previous findings (Bouton et al., 2004; Culver et al., 2018) to the spontaneous recovery, reinstatement, and reacquisition of extinguished conditioned responding in human fear conditioning, employing partially reinforced, unpaired, and non-reinforced extinction training. Furthermore, a direct comparison of occasionally paired and unpaired US presentations during extinction would also allow for examination of underlying mechanisms, which may differ across different types of reinforced extinction training (e.g. Bouton et al., 2004; Rauhut, Thomas, & Ayres, 2001; Rescorla & Skucy, 1969).

Several mechanisms have been proposed to account for the superior protection from fear recovery effects subsequent to reinforced and unpaired extinction training, compared to non-reinforced extinction, including: Weakening of the CS-US contingency through unpaired US presentations (Frey & Butler, 1977; Rescorla & Skucy, 1969; Vervliet, Vansteenwegen, & Hermans, 2010); US habituation (Rauhut et al., 2001; but see Thomas, Longo, & Ayres, 2005); sequential learning (Bouton et al., 2004; Capaldi, 1966, 1994); and enhanced extinction learning through violation of expectancies, also referred to as prediction errors (Craske et al., 2014; Culver et al., 2018; Rescorla & Wagner, 1972; Vurbic & Bouton, 2014).

Prediction errors are implicated in the acquisition and extinction of conditioned responding (e.g. Pearce & Hall, 1980; Rescorla & Wagner, 1972; Vurbic & Bouton, 2014), whereby learning is proposed to cease when the CS reliably predicts the delivery of the US (or its absence, in the case of extinction learning). Extinction learning may be enhanced through the occasional presentation of the US during extinction training, due to the violation of expectancies regarding the frequency of US presentations or changes to the CS-US relationship (e.g. Craske et al., 2014). For instance, the omission of the US at the onset of extinction provides an opportunity for new learning due to the mismatch between current information (CS-no US) and past learning (CS-US), while the presentation of occasionally paired and unpaired USs on later trials would sustain learning through the presentation of novel information that needs to be reconciled with prior learning. Hence, the occasional presentation of USs during extinction would allow participants to learn about the likelihood of future threat encounters, such as the frequency of US presentations, relative to CS-only trials, or the relationship between the CS and the US (i.e. occasionally paired or unpaired). Subsequent fear recovery could be reduced because participants learned that the CS predicts the absence of the US (unpaired extinction; Rescorla & Skucy, 1969; Vervliet et al., 2010) or that occasional CS-US trials occur in the presence of many CS-no US trials (partially reinforced extinction). This proposition is also supported by Bouton et al.'s (2004) adaptation of sequential theory (Capaldi, 1966, 1994).

Bouton et al. (2004) proposed that the key aspect learned during partially reinforced extinction training is that CS-US trials do not occur exclusively in the "context" of other CS-US trials (i.e. acquisition), but may also occur in the context of extinction trials (i.e. a CS-US trial is followed by several CS-no US trials). Due to the association of reinforced trials with non-reinforced trials, reacquisition of extinguished responding may occur at a slower rate, as participants may expect a CS-US trial to be followed by further CS-no US trials. Bouton et al. further proposed that similar learning would occur during unpaired extinction, whereby unpaired US presentations would weaken the US's exclusive association with the acquisition context. Conversely, reacquisition subsequent to non-reinforced extinction is proposed to occur at a faster rate, as the omission of the US during training maintains the US's exclusive association with the acquisition context (i.e. CS-US trials occurring in the context of other CS-US trials). Hence, post-extinction presentations of reinforced trials would signal delivery of further reinforced trials and lead to rapid reacquisition. It should be noted that Bouton et al.'s model would also predict reduced reinstatement of fear, particularly in participants who received unpaired USs during extinction, but it is not readily applicable to spontaneous recovery, unless additional assumptions are made. Spontaneous recovery may depend, in part, on how easily the extinction memory can be retrieved, meaning the memory that a CS-only trial is more likely to signal further non-reinforced than reinforced trials. To summarize, there are several mechanisms that may account for reduction of fear recovery following reinforced extinction training, although, at present, they are still poorly understood and require further examination.

In the present study, we investigated whether occasional presentations of paired and unpaired USs during extinction training would result in superior reduction of spontaneous recovery, reinstatement, and reacquisition of fear, compared to non-reinforced extinction training. The fear association was established through differential Pavlovian conditioning (Culver et al., 2018; Lipp, 2006a), whereby one neutral cue (CS+) was continuously paired with an aversive electrotactile stimulus (US), while another cue (CS-) was presented by itself. Conditioned fear in differential paradigms is reflected in larger responding to the CS+, relative to the CS- (Lipp, 2006a). In line with

Culver et al. (2018), electrodermal responding and CS valence ratings were recorded as primary and secondary dependent measures of conditioned responding, respectively. Based on the reviewed literature (Bouton et al., 2004; Craske et al., 2014; Culver et al., 2018), we predicted that extinction learning would be enhanced through occasional presentations of the US during extinction training, which would be reflected in reduced fear recovery and reacquisition of fear, relative to non-reinforced extinction. In line with Bouton et al.'s findings, we also predicted that unpaired extinction would result in slower reacquisition of extinguished fear than partially reinforced extinction.

Materials and Methods

Participants

University students who met inclusion criteria (i.e. no cardiovascular disease, seizure disorder, or pregnancy) participated in exchange for partial course credit or a financial compensation of 30 AUD. After exclusion of one participant who failed to verbalize the CS-US contingency, data from 72 participants (44 females, 28 males; female-male ratio per group: 15:9 [non-reinforced and partially reinforced extinction group], 14:10 [unpaired extinction group]) were included in the analyses. The age range of participants was 18 to 38 years (M = 21.61 years, SD = 3.95). Ethical approval for this study was obtained from the Curtin University Human Research Ethics Committee.

Apparatus and Materials

Stimuli. In line with previous research (Culver et al., 2018), non-fear relevant stimuli have been employed in the present study. Conditioned stimuli (CS) included four color images of animals, two birds and two fish (sourced from the internet). Each participant was presented with a subset of two pictures, comprising one bird and one fish picture; stimulus sets were counterbalanced across participants. The pictures measured between 700 x 467 pixels and 700 x 541 pixels and were presented for 6 s, in the center of a 17-inch color LCD screen, over a black background, with an inter-trial interval of 13 to 17 s. To control for order effects, the assignment of bird and fish pictures as CS+ or CS- and the presentation order (whether the first trial of each phase was a CS+ or CS-) were counterbalanced across participants. Stimuli were presented in a pseudo-randomized order, whereby each CS was presented four times within blocks of eight trials, with the restriction that no more than two consecutive CS+ or CS- trials were presented. The unconditioned stimulus (US) consisted of a mild electric shock, which was generated with a Grass SD9 stimulator (Grass Technologies, Middleton, WI) and was delivered to the wrist of the dominant hand via a concentric electrode. The shock was presented for 200 ms (pulsed at 50 Hz) and coincided with the CS+ offset (unless otherwise indicated); the CS- was never paired with the US. The delivery of the US and CSs was controlled with DMDX 5.0.5 software (Forster & Forster, 2003).

Electrodermal activity (skin conductance responses, SCRs). Electrodermal activity was recorded through two self-adhesive isotonic gel electrodes (Biopac Systems EL507), attached to the thenar and hypothenar eminences of the non-dominant hand. Electrodermal activity was DC amplified at a gain of 5 micro Siemens (µS) per volt and recorded with a Biopac MP150 system at a sampling frequency of 1000 Hz, using AcqKnowledge 4 (Biopac Systems, Goleta, CA). A Biopac respiration belt was fitted around each participant's waist to control for respiration-induced artefacts in SCRs.

Subjective evaluation of stimulus valence. Participants provided post-test CS and US valence ratings on a 9-point scale (from 1 [unpleasant] to 9 [pleasant]) at baseline, after acquisition, extinction, spontaneous recovery (CS ratings only), reinstatement test, and reacquisition. US valence ratings were not obtained after spontaneous recovery to prevent potential interference with subsequent reinstatement, as reinstatement requires the unexpected presentation of the US (Haaker, Golkar, Hermans, & Lonsdorf, 2014). Valence ratings were obtained electronically through a custom-made Microsoft Access application, whereby the stimuli were presented on the computer screen in randomized order and participants were instructed to rate stimulus valence on the scale located below the picture.

Self-report questionnaires. To ensure groups did not differ on variables known to affect conditioned responding (Dunsmoor, Campese, Ceceli, LeDoux, & Phelps, 2015; Lonsdorf & Merz, 2017), participants were asked to complete the short version of the Intolerance of Uncertainty Scale (IUS-12; Carleton, Norton, & Asmundson, 2007). The 12-item IUS measures beliefs about and reactions to uncertainty, ambiguous situations, and the future (e.g. "Unforeseen events upset me greatly.") and comprises two subscales: prospective IU (measures anxiety about future events) and inhibitory IU (indicative of behavioral inhibition or avoidance). The scale has good psychometric properties (Carleton et al., 2007) and demonstrated excellent internal consistency in the current sample (α = .90). For exploratory purposes, participants also completed the short version of the Depression, Anxiety, and Stress Scales (DASS-21; Henry & Crawford, 2005; Lovibond & Lovibond, 1995). The questionnaires were completed electronically at the start of the experiment.

Manipulation checks. Following acquisition, participants were presented with a CS-US contingency questionnaire, containing the four stimulus pictures used in this study, and were asked to indicate which pictures had been paired with the US. As inability to verbalize the correct contingency may reflect a genuine failure to learn the CS-US relationship (Lipp, 2006a), data from one participant who failed this test were excluded from statistical analyses.

Procedure

A schematic representation of the experimental paradigm and the reinforcement schedule employed during extinction training is presented in Figure 1. Upon arrival in the laboratory, participants were informed about the experimental procedures and had the opportunity to ask questions, before providing information about current medication use and medical history. Participants were assigned to groups in the order they presented for testing, with the restriction that an approximately equal number of females and males were assigned to each group. Individuals who met inclusion criteria provided written consent, were seated in front of the computer screen, and were fitted with the skin conductance electrodes, respiration belt, and the shock electrode. After completing the self-report measures, participants were asked to relax and look at the blank computer screen while a 2-min baseline of their electrodermal activity was recorded. Subsequently, participants provided baseline CS valence ratings, set the US intensity to a level which was perceived as “unpleasant, but not painful,” and rated US valence.

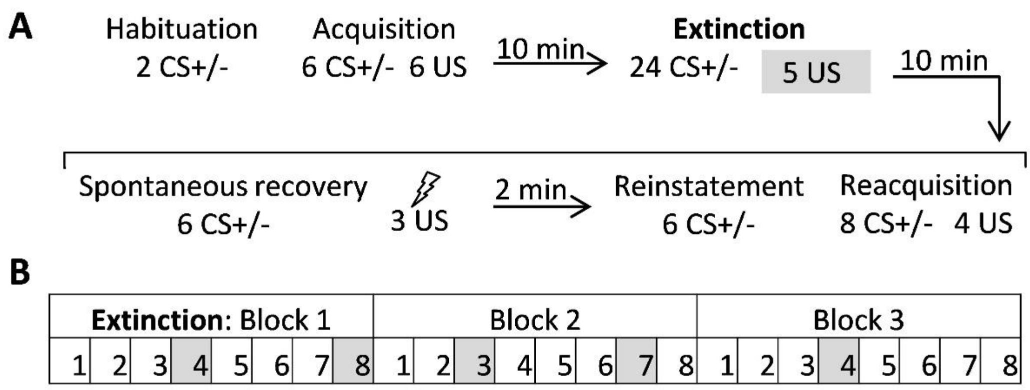

Figure 1. Schematic representation of the experimental paradigm (**A**) and reinforcement schedule during extinction training (**B**). Panel A illustrates the sequential order of experimental stages with the respective number of conditioned (CS) and unconditioned stimulus (US) presentations. Boxes in panel B denote delivery of the previously conditioned CS+ (CS- is not depicted here). The shaded boxes represent delivery of the US, which either coincided with the CS+ offset (partially reinforced extinction group) or was presented in the middle of the inter-trial interval (unpaired extinction group), either after CS+ offset or before CS+ onset.

Acquisition. Participants were asked to pay attention to the computer screen and to learn which CSs were followed by the US. Conditioning commenced with a habituation phase, consisting of two non-reinforced presentations of each CS, and was immediately followed by acquisition, which involved six presentations of the CS+ and the CS-. The US was presented on all CS+ trials. Thereafter, participants completed the CS-US contingency questionnaire and rated CS and US valence. The shock electrode was removed and a 10-min rest period was inserted during which participants were offered magazines to read. The break was included to allow electrodermal activity to return to a stable baseline before the start of extinction training.

Extinction. The shock electrode was reattached and participants were informed that they would be presented with several stimuli and asked to pay attention to the computer screen at all times. No information was provided about the types of stimuli to be presented, their frequency, duration, or contingencies. Extinction training consisted of 24 CS+/- presentations; the CSs+ were non-reinforced in the non-reinforced extinction (EXT) and unpaired extinction (UNP) group and partially reinforced (5 CS-US pairings; see Figure 1B) in the partially reinforced extinction (PRE) group. In the UNP group, five US presentations were delivered in the middle of the inter-trial interval (ITI). Within each block of eight trials, the US was delivered on trial 4 and 8 (block 1), 3 and 7 (block 2), and on trial 4 (block 3). The US coincided with the CS+ offset in the PRE group. In the UNP group, the US was presented in the middle of the ITI, either after CS+ offset or in the ITI before CS+ onset, with the restriction that the US was never presented between two CSs+. As previous reports indicated that a 2:8 reinforcement ratio maintained differential responding at the end of extinction training (Culver et al., 2018), we decreased the reinforcement ratio on the last block of training to a 1:8 ratio. This decrease was in line with past research (Bouton et al., 2004) and served to facilitate loss of conditioned responding. Following extinction training, participants were asked to rate CS and US valence. Subsequently, the shock electrode was removed and a 10-min break was inserted (identical to the post-acquisition break).

Test of fear recovery. The shock electrode was reattached and participants were informed that they would be presented with several stimuli and were asked to pay attention to the computer screen at all times (instructions were identical to those presented at the start of extinction training). Assessment of spontaneous recovery involved six non-reinforced presentations of the CS+/-, which were followed by post-test CS valence ratings. Subsequently, participants received three unsignaled presentations of the US, with a duration of 200 ms and an ITI of 6 s. The computer screen remained switched on and displayed a black background (in line with previous training). After a 2-min delay, reinstatement of extinguished responding was tested through six non-reinforced presentations of the CS+/-. After rating CS and US valence, participants underwent partially reinforced reacquisition, comprising eight presentation of the CS+/-. The US coincided with the CS+ offset on 50% of the trials (1^{st}, 3^{rd}, 5^{th} and 6^{th} CS); the remaining trials were not reinforced. The physical intensity of the US during reinstatement and reacquisition was identical to that employed during acquisition. We employed a partial reinforcement schedule (in line with Bouton et al., 2004) to permit emergence of group differences (Lissek, Pine, & Grillon, 2006). The partial reinforcement schedule further served to enhance the ecological validity of the reacquisition test, as individuals would be more likely to encounter occasional, than continuous, CS-US pairings in real life. At the conclusion of reacquisition, participants provided the final CS and US valence ratings.

Scoring and Response Definition

Electrodermal responses were scored offline in AcqKnowledge 4. Participants' baseline electrodermal activity was determined by counting all spontaneous responses that occurred during a 2-min rest period (Dawson, Schell, & Filion, 2007). A visual inspection of data was conducted to identify movement- or respiration-induced artefacts in SCRs. Eight SCRs (across groups) were discarded due to the presence of artefacts. In accordance with past research (Culver et al., 2018; Pineles, Orr, & Orr, 2009), SCRs elicited by the CSs were calculated by subtracting the mean skin conductance level during the 2 s baseline preceding CS onset from the largest skin conductance level occurring 1 to 6 s after CS onset. SCRs were range corrected to control for individual differences in electrodermal activity (Lykken, 1972) and then square root transformed to reduce the skew of the distribution (Dawson et al., 2007). The range correction was obtained by dividing each response by the largest response displayed by the participant. Electrodermal responses were averaged into blocks of two consecutive trials, to reduce the influence of trial by trial variability.

Statistical Analyses

Electrodermal responding during habituation was analyzed through a repeated measures analysis of variance (ANOVA) with group (EXT, PRE, UNP) as a between-groups factor and CS type (CS+, CS-) as a within-groups factor. Analysis of acquisition, extinction, and reacquisition data was conducted through mixed ANOVAs for repeated measures, with group as a between-groups factor and CS type (CS+ vs. CS-) and block/time as within-groups factors (acquisition: block 1-3; extinction: early vs. late phase; reacquisition: block 1-3). Extinction of conditioned responding was assessed with data from the early (block 1-2) and late phase (block 11-12) of extinction training, during which no USs were presented in any of the groups (the US presented at the end of trial 4/block 2 in the PRE and UNP group would affect responding on the subsequent trials, but not on trial 4). In line with past research (Dunsmoor et al., 2015), recovery of extinguished responding was assessed during the early phase (block 1) of spontaneous recovery and reinstatement tests, through separate repeated measures ANOVAs, with CS type (CS+ vs. CS-) as a within-groups factor and group as a between-groups factor. CS valence ratings were analyzed through a series of mixed ANOVAs for repeated measures, with group as a between-groups factor and CS type (CS+ vs. CS-) and time as within-groups factors (acquisition: baseline vs. acquisition; extinction: acquisition vs. extinction; spontaneous recovery: extinction vs. spontaneous recovery; reinstatement: extinction vs. reinstatement). Reacquisition was assessed by means of a repeated measures ANOVA with group as a between-groups factor and CS type (CS+ vs. CS-) as a within-groups factor. Multivariate *F* values (Pillai's Trace) and partial eta squared values are reported for all main effects and interactions. Statistical significance was assessed at $\alpha = .05$; Bonferroni corrections were used for follow-up analyses to guard against the accumulation of a Type 1 error.

Results

Preliminary Analyses

The groups did not differ in age, selected US intensity, baseline CS or US valence ratings, baseline electrodermal activity, IUS-12 scores, or DASS-21 scores (Table 1). Selected US intensities ranged from 34 to 80 Volt, with a mean of 66.19 Volt (*SD* = 12.90).

Table 1

Means (M) and Standard Deviations (SD) for Age, US Intensity, Baseline Valence Ratings (VR), Baseline Electrodermal Activity (EDA), and Self-Report Questionnaires.

	Non-reinforced extinction		Partially reinforced extinction		Unpaired extinction		
	M	*SD*	*M*	*SD*	*M*	*SD*	Test
Age	20.37	2.43	21.79	3.50	22.67	5.21	*F*(2,69)=2.13, *p*=.127
US intensity	63.38	13.39	68.17	11.59	67.04	13.67	*F*(2,69)=0.90, *p*=.410
Baseline VR							
CS+	7.25	1.51	6.71	2.05	7.38	1.35	*F*(2,69)=1.09, *p*=.342
CS-	6.96	1.57	6.88	2.15	7.54	1.56	*F*(2,69)=1.00, *p*=.375
US	3.88	1.92	3.83	1.88	3.42	1.74	*F*(2,69)=0.45, *p*=.639
Baseline EDA	12.42	9.60	11.50	8.89	10.13	8.35	*F*(2,69)=0.40, *p*=.673
IUS-12							
Prospective	18.46	5.12	19.62	5.45	18.96	5.49	*F*(2,69)=0.29, *p*=.752
Inhibitory	9.54	3.56	11.38	4.75	10.92	4.69	*F*(2,69)=1.14, *p*=.324
DASS-21							
Depression	8.67	10.05	7.92	7.99	9.00	7.55	*F*(2,69)=0.10, *p*=.905
Anxiety	6.17	5.04	7.50	8.18	7.75	7.65	*F*(2,69)=0.35, *p*=.709
Stress	10.75	7.73	13.75	9.57	14.25	9.35	*F*(2,69)=1.08, *p*=.345

Note. CS = conditioned stimulus, US = unconditioned stimulus, IUS-12 = Intolerance of Uncertainty Scale (short version), DASS-21 = Depression, Anxiety, and Stress Scales (short version). Baseline EDA refers to the number of spontaneous responses that occurred during a 2-min rest period. US intensity is reported in Volt.

Electrodermal Responding

Habituation. Electrodermal responding across groups is presented in Figure 2. Analysis of SCRs during habituation indicated that SCRs to CS+/- did not differ across groups during habituation, as reflected in the non-significant main effect of CS type, $F(1, 69) < 1$, and the CS type x group interaction, $F(2, 69) = 2.93$, $p = .060$, $\eta p^2 = .08$. Follow-up comparisons conducted for the trend towards significance in the interaction revealed larger SCRs to the CS- ($M = 0.57$, $SD = 0.05$) than to the CS+ ($M = 0.47$, $SD = 0.05$) in the EXT group (Figure 2A), $F(1, 69) = 4.10$, $p = .047$, $\eta p^2 = .06$, but not in the PRE or UNP group, both $F(1, 69) \leq 1.94$, $p \geq .168$, $\eta p^2 \leq .03$. As the interaction indicates that groups may have differed on their level of electrodermal responding to CS+/- at the start of acquisition, we conducted a separate 3 (group) x 2 (CS+, CS-) repeated measures ANOVA with data from the first trial of acquisition. The analysis revealed no significant group differences at the start of acquisition, as reflected in the non-significant main effect of CS type, $F(1, 69) = 0.53$, $p = .471$, $\eta p^2 = .01$ and the non-significant CS type x group interaction, $F(2, 69) = 0.35$, $p = .709$, $\eta p^2 = .01$.

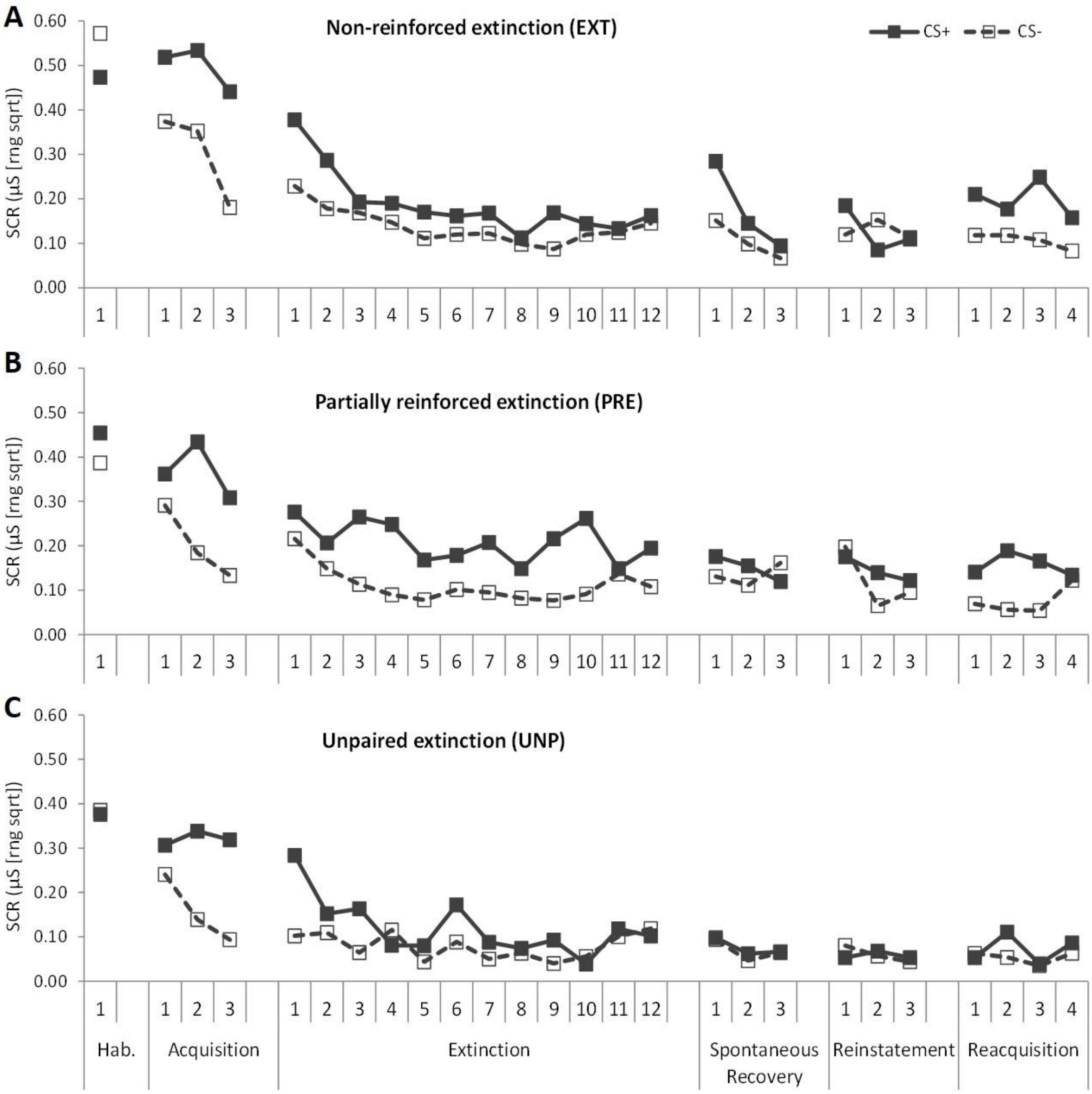

Figure 2. Mean skin conductance responses (SCRs) to reinforced (CS+) and non-reinforced (CS-) conditioned stimuli in the EXT (**A**), PRE (**B**), and UNP (**C**) group. SCRs are presented in blocks of two consecutive trials.

Acquisition. Differential SCRs were evident across groups during acquisition, as reflected in main effects of CS type, $F(1, 69) = 61.91$, $p < .001$, $\eta p^2 = .47$; block, $F(2, 68) = 12.18$, $p < .001$, $\eta p^2 = .26$, and a CS type x block interaction, $F(2, 68) = 6.85$, $p = .002$, $\eta p^2 = .17$. The interaction reflects an increase in differential SCRs across blocks of acquisition, all $F(1, 69) \geq 13.16$, $p \leq .001$, $\eta p^2 \geq .16$. The remaining interactions did not attain significance, with the largest of the non-significant effects suggesting that acquisition of differential SCRs did not significantly differ across groups, (CS type x block x group), $F(4, 138) = 0.94$, $p = .443$, $\eta p^2 = .03$.

Extinction. A visual inspection of Figure 2 indicates that differential responding was larger during partially reinforced than during non-reinforced extinction. However, statistical analyses showed that differential responding during the reinforced stage of training did not interfere with the extinction of conditioned responding. Results of the ANOVA examining SCRs between the early (block 1-2) and late phase (block 11-12) of training revealed main effects of CS type, $F(1, 69) = 18.84$, $p < .001$, $\eta p^2 = .21$, and block, $F(3, 67) = 7.96$, $p < .001$, $\eta p^2 = .26$, which were qualified by a CS type x block interaction, $F(3, 67) = 4.69$, $p = .005$, $\eta p^2 = .17$. The interaction reflects differential electrodermal responding on block 1 and 2, both $F(1, 69) \geq 6.80$, $p \leq .011$, $\eta p^2 \geq .09$, but not on block 11 or 12, both $F(1, 69) \leq 1.27$, $p \geq .265$, $\eta p^2 \leq .02$, indicating that differential responding was extinguished in all groups. The remaining interactions did not attain significance, largest effect (CS type x block x group), $F(6, 136) = 1.17$, $p = .324$, $\eta p^2 = .05$. As Figure 2 indicates that differential responding may have been present on the last block of training in the PRE group, we subjected the mean SCRs of the CS+ and the CS- to a *t* test. In line with the between-groups comparisons, results confirmed that there were no significant differences between SCRs to the CS+ ($M = 0.20$, $SD = 0.22$) and the CS- ($M = 0.11$, $SD = 0.17$), $t(23) = 1.68$, $p = .106$, $d = 0.05$.

Spontaneous recovery. Our primary prediction was that occasional US presentations during extinction training would result in less recovery of fear than non-reinforced extinction training. Inspection of spontaneous recovery data in Figure 2 suggests that differential responding recovered following non-reinforced extinction, but not after extinction conducted with occasionally paired or unpaired US presentations. This observation was confirmed by the results of statistical analyses, revealing a main effect of CS type, $F(1, 69) = 8.22$, $p = .005$, $\eta p^2 = .11$, which was qualified by a CS type x group interaction, $F(2, 69) = 3.25$, $p = .045$, $\eta p^2 = .09$. The interaction reflects differential electrodermal responding in the EXT group, $F(1, 69) = 13.19$, $p = .001$, $\eta p^2 = .16$, but not in the PRE group, $F(1, 69) = 1.51$, $p = .223$, $\eta p^2 = .02$, or UNP group, $F(1, 69) = 0.01$, $p = .917$, $\eta p^2 < .01$.

Reinstatement. A visual inspection of differential responding on block 1 of reinstatement (Figure 2) suggests that differential SCRs were reinstated in the EXT group, but not in the PRE or

UNP group. However, these differences did not attain significance in the omnibus analysis. The results neither yielded a significant main effect of CS type, $F(1, 69) = 0.04$, $p = .837$, $\eta p^2 = .01$, nor a CS type x group interaction, $F(2, 69) = 1.48$, $p = .234$, $\eta p^2 = .04$, showing that differential SCRs did not differ across groups on the first block of reinstatement testing.

Reacquisition. Our first prediction was that occasional presentations of the US during extinction training would reduce reacquisition of extinguished responding, compared to non-reinforced extinction. To test this prediction, we conducted an ANOVA using data from block 1-3, which was the reinforced stage of reacquisition during which group differences would be expected to emerge (see also Figure 2). The results revealed a main effect of CS type, $F(1, 69) = 17.29$, $p < .001$, $\eta p^2 = .20$, as well as a trend towards significance in the CS type x group interaction, $F(2, 69) = 2.57$, $p = .084$, $\eta p^2 = .07$. The interaction reflects increased differential responding to the CS+, relative to the CS-, in the EXT group, $F(1, 69) = 10.18$, $p = .002$, $\eta p^2 = .13$, as well as in the PRE group, $F(1, 69) = 11.94$, $p = .001$, $\eta p^2 = .15$, but not in the UNP group, $F(1, 69) = 0.31$, $p = .580$, $\eta p^2 = .01$. The main effect of block and remaining interactions did not attain significance, largest effect (block x group interaction), $F(4, 138) = 1.09$, $p = .362$, $\eta p^2 = .03$.

To follow up on Bouton et al.'s (2004) findings, we also tested whether the rate of reacquisition would be slower after unpaired than partially reinforced extinction. The ANOVA conducted with data from the PRE and UNP group yielded a main effect of CS type, $F(1, 46) = 9.87$, $p = .003$, $\eta p^2 = .18$, as well as a CS type x group interaction, $F(1, 46) = 5.15$, $p = .028$, $\eta p^2 = .10$, reflecting larger differential SCRs in the PRE group ($M = 0.11$, $SD = 0.17$) than in the UNP group ($M = 0.02$, $SD = 0.08$), $t(46) = 2.23$, $p = .031$, $d = 0.68$. The main effect of block and remaining interactions did not attain significance, largest effect (CS type x block interaction), $F(2, 45) = 1.30$, $p = .282$, $\eta p^2 = .06$. These results show that the rate of reacquisition differed between the PRE and UNP group; although, contrasting Bouton et al.'s findings, we did not observe reduced reacquisition, but an absence of reacquisition in the UNP group.

Examination of underlying mechanisms: US habituation. The lack of reacquisition in the UNP group may indicate that the unpaired presentations of the US during extinction training resulted in US habituation. The reduced aversiveness of the US could have attenuated subsequent responding to the CS+ and slowed the rate of reacquisition (Rescorla, 1973). If US habituation occurred during extinction training in the UNP group, this would be reflected in smaller unconditioned responses (URs) in the UNP than in the PRE group.[1] The results of a 2 (group) x 5 (US presentations) ANOVA did not support the US habituation hypothesis. Results revealed a main effect of trial, $F(4, 43) = 6.76$, $p < .001$, $\eta p^2 = .39$, but no significant trial x group interaction, $F(4, 43) = 0.50$, $p = .736$, $\eta p^2 = .04$. The main effect reflects decreased URs across trials in both groups. However, the mean UR to the final US presentation in the PRE ($M = 0.74$, $SD = 0.34$) and UNP group ($M = 0.67$, $SD = 0.29$) resembled that on the last block of acquisition (PRE: $M = 0.58$, $SD = 0.25$; UNP: $M = 0.65$, $SD = 0.30$). Similarly, analysis of post-extinction US valence ratings showed that the US was rated as equally unpleasant in both groups (PRE: $M = 3.38$, $SD = 1.66$; UNP: $M = 3.96$, $SD = 1.73$), $t(46) = 1.19$, $p = .240$. The combined results, therefore, indicate that US habituation did not occur in either group.

CS Valence Ratings

Acquisition. Mean ratings of CS+ and CS- valence as well as negative evaluations of the CS+, relative to the CS-, are presented in Figure 3. Assessment of acquisition (baseline vs. post-acquisition ratings) showed increased negative evaluations of the CS+, relative to the CS-, in all groups. The repeated measures ANOVA yielded main effects of CS type, $F(1, 69) = 25.24$, $p < .001$, $\eta p^2 = .27$, and time, $F(1, 69) = 46.98$, $p < .001$, $\eta p^2 = .41$, which were qualified by a CS type x time interaction, $F(1, 69) = 53.67$, $p < .001$, $\eta p^2 = .44$. The interaction reflects a significant difference between evaluations of the CS+ and the CS- in all groups, after acquisition, $F(1, 69) = 48.31$, $p < .001$, $\eta p^2 = .41$, but not at baseline, $F(1, 69) = 0.01$, $p = .944$, $\eta p^2 < .01$.

[1] Respective unconditioned responses (URs) in the partially reinforced and unpaired extinction group were scored as the largest response starting in the 1 to 3 s window following US offset (Prokasy & Ebel, 1967). URs were range corrected and square root transformed prior to analysis.

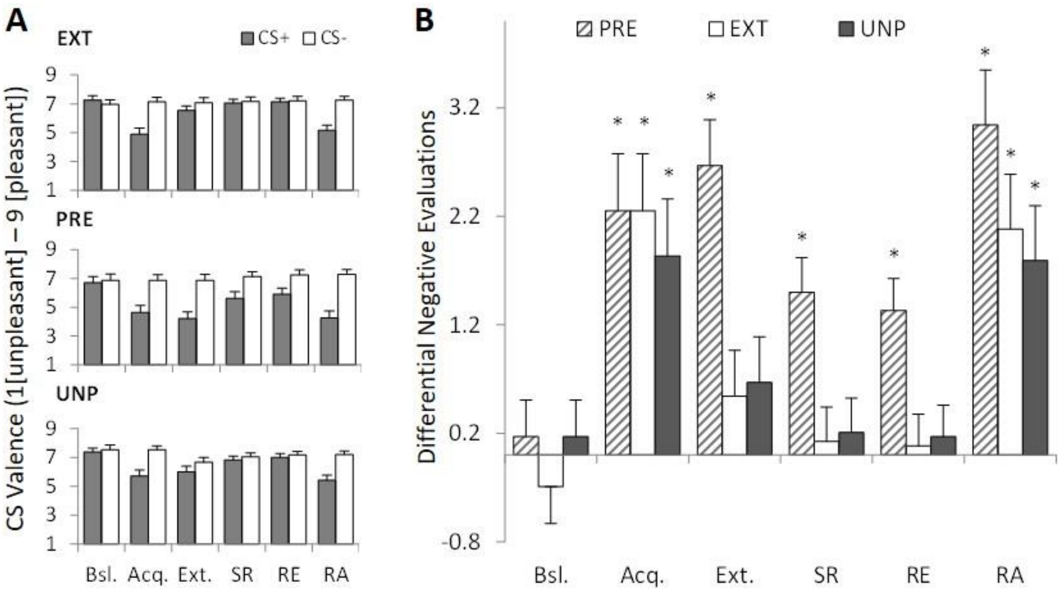

Figure 3. Mean conditioned stimulus (CS) valence ratings (**A**) and differential negative evaluations of the CS+, relative to the CS- (**B**) at baseline (Bsl.), after acquisition (Acq.), extinction (Ext.), spontaneous recovery (SR), reinstatement (RE), and reacquisition (RA). CS valence was rated from 1 (unpleasant) to 9 (pleasant). Error bars represent standard errors. The asterisk in panel B indicates significant negative evaluation of the CS+, relative to the CS-, at $p \leq .001$.

Extinction. Assessment of extinction of differential evaluations revealed a main effect of CS type, $F(1, 69) = 50.03$, $p < .001$, $\eta p^2 = .42$, as well as interactions of CS type x time, $F(1, 69) = 9.26$, $p = .003$, $\eta p^2 = .12$, and time x group, $F(2, 69) = 4.58$, $p = .014$, $\eta p^2 = .12$, which were qualified by a CS type x time x group interaction, $F(2, 69) = 5.61$, $p = .006$, $\eta p^2 = .14$. The three-way interaction reflects a significant decrease in negative evaluations from post-acquisition to post-extinction as well as elimination of differential evaluations in the EXT and UNP group, but not in the PRE group. Differential negative evaluations in the PRE group remained significant after extinction training, $F(1, 69) = 39.70$, $p < .001$, $\eta p^2 = .37$.

Tests of fear recovery. Assessment of spontaneous recovery and reinstatement data showed that this pattern was also reflected in ratings obtained after spontaneous recovery, CS type x group, $F(2, 69) = 8.39$, $p = .001$, $\eta p^2 = .20$, and after reinstatement, CS type x group, $F(2, 69) = 8.39$, $p = .001$, $\eta p^2 = .20$ (see also Figure 3B). No group differences were found in valence ratings obtained after reacquisition; negative evaluations of the CS+, relative to the CS-, were observed in all groups, as reflected in the significant main effect of CS type, $F(1, 69) = 62.76$, $p < .001$, $\eta p^2 = .48$, and non-significant CS type x group interaction, $F(2, 69) = 1.68$, $p = .193$, $\eta p^2 = .05$. In summary, analysis of valence ratings showed that, compared to electrodermal responding, post-test ratings of CS valence exhibited a different pattern of conditioned responding, whereby occasional CS-US pairings, but not unpaired US presentations, maintained differential evaluations after extinction, spontaneous recovery and reinstatement. Additionally, all groups rated the CS+ as significantly more unpleasant than the CS- after reacquisition, indicating that differential evaluations were reacquired.

Discussion

Our primary aim was to investigate if occasional presentations of the US during extinction training would provide enhanced protection from fear recovery than non-reinforced extinction. The overall pattern of results supported this hypothesis, although there were important differences between the effects of partially reinforced and unpaired extinction training. The results showed no differences in the acquisition and extinction of differential SCRs across groups. Group differences emerged during tests of fear recovery, whereby spontaneous recovery of extinguished differential SCRs was observed after non-reinforced extinction, but not after extinction conducted with occasional presentations of paired or unpaired USs. Analysis of SCRs during reinstatement failed to show significant group differences, indicating that the presentation of three unsignaled shocks did not reinstate differential SCRs in any group. The reacquisition test yielded somewhat unexpected results, indicating reacquisition of differential SCRs subsequent to partially reinforced and non-reinforced extinction, but not after unpaired extinction training.

In contrast to electrodermal data, post-test ratings of CS+/- valence did not reflect any benefits of occasional presentations of the US during extinction over non-reinforced extinction. While acquisition of differential negative evaluations was in line with the electrodermal data, ratings obtained after the subsequent phases showed a different pattern of results. Negative evaluations of the CS+, relative to the CS-, were observed in the group that received partially reinforced extinction, but not in the groups that received non-reinforced or unpaired extinction, after extinction, spontaneous recovery, and reinstatement. Differential evaluations did not differ across groups after reacquisition, as the CS+ was rated as less pleasant than the CS- in all groups. The combined results demonstrate that partially reinforced and unpaired extinction training offer enhanced protection from fear recovery assessed by SCRs, relative to non-reinforced extinction. However, they also indicate that physiological and verbal indices of conditioned fear may be differentially sensitive to occasional presentations of the US during extinction, as reflected in the different pattern of results obtained from the analysis of electrodermal data and the analysis of valence ratings. Our results also indicate that occasional presentations of unpaired USs during extinction may be more effective in the long-lasting reduction of fear, as indexed by SCRs, than partially reinforced extinction, as unpaired extinction may prevent reacquisition of extinguished responding.

The overall pattern of results can be interpreted as being broadly consistent with past research to the extent that a) occasional presentations of the US during extinction resulted in superior attenuation of recovery from extinction effects than non-reinforced extinction (Bouton et al., 2004; Culver et al., 2018) and b) we observed a dissociation between physiological and verbal measures of fear (e.g. Culver et al., 2018; Luck & Lipp, 2015; Schultz, Balderston, Geiger, & Helmstetter, 2013; Thompson & Lipp, 2017). However, there were several important differences between our findings and those reported in past research. The results of the present study did not converge with previous reports (Bouton et al., 2004; Culver et al., 2018) that interference with reacquisition was greater following partially reinforced than non-reinforced extinction – reacquisition in our study did not differ between the control and partially reinforced group. While this finding was unexpected, in particular since partially reinforced extinction prevented spontaneous recovery of extinguished responding, the divergent results between this and Bouton et al.'s study may reflect on the use of an aversive conditioning preparation, as opposed to appetitive conditioning, as well as on the larger number of extinction trials employed by Bouton et al.

Repeated extinction sessions, conducted across multiple days, may have strengthened the animal subjects' memory of occasional CS-US trials occurring in the presence of many CS-no US trials, which, according to Bouton et al.'s (2004) adaptation of sequential theory (Capaldi, 1966, 1994), would slow the rate of reacquisition. It is possible that the 24 extinction trials employed in our study were not sufficient to create a robust extinction memory. On the other hand, Culver and colleagues (2018) employed the same number of extinction trials, albeit with six, instead of five, paired US presentations, and reported decreased reacquisition after reinforced than non-reinforced extinction. However, a closer examination of Culver et al.'s data suggests that differential SCRs were initially larger in the partially reinforced extinction group than in the control group (after the 1st and 2nd CS-US trial) and only decreased during the second half of reacquisition. The authors suggested that physiological responding decreased because participants acquired "physiological toughness," meaning an improved ability to cope with repeated exposure to aversive stimuli, as a result of partially reinforced extinction training. There was no evidence of decreased SCRs during reacquisition in the partially reinforced extinction group in our data, although this may reflect on the use of a partial reinforcement schedule during reacquisition. In contrast to continuous reinforcement (Culver et al., 2018), alternating CS-US with CS-no US trials may create more uncertainty about future threat. From an evolutionary perspective, even occasional threats may pose a risk to our survival. Hence, it may be of advantage to adopt a "better safe than sorry" approach in response to occasional threat encounters. In this sense, partially reinforced extinction training may protect against spontaneous recovery of fear, as a CS-only trial does not signal imminent danger, but the fear response may return when the likelihood of future threat increases, as would be the case after exposure to a CS-US trial. This proposition could be explored in future research through a direct comparison of reacquisition on a partial and continuous reinforcement schedule.

Analysis of electrodermal data during reacquisition further indicated that the mere reduction of the reinforcement schedule during extinction training, relative to that used during acquisition, may not be sufficient to prevent reacquisition of differential SCRs. This proposition was supported by the differential pattern of reacquisition results observed between the partially reinforced and the unpaired extinction group. Our results indicated that, in contrast to partially reinforced extinction, an unpaired extinction procedure that involves the occasional presentations of the US in the inter-trial interval may not only reduce, but also prevent, reacquisition of differential SCRs. These results are in line with previous animal research conducted in the appetitive (Bouton et al., 2004) and aversive setting (Frey & Butler, 1977; Mickley et al., 2009; Rauhut et al., 2001; Thomas et al., 2005). Similarly, Vervliet et al. (2010) reported that extinction training consisting of eight non-reinforced CS trials and six unpaired electrotactile USs, presented during the inter-trial interval, reduced renewal of extinguished fear in humans. While the experimental methods differed across these studies, including the number of CS and US trials and the ratio of CS to US trials during extinction training, the outcomes of the present study and past research suggest that extinction training during which the US is retained, but no longer paired with the CS, is superior in the reduction of fear recovery to non-reinforced or partially reinforced extinction.

As a possible underlying mechanism, it has been proposed that unpaired presentations of the US weaken, or even eliminate, the CS-US association that was formed during acquisition (Frey & Butler, 1977; Rescorla & Skucy, 1969; Rescorla & Wagner, 1972; Vervliet et al., 2010). While the associative strength of the CS would also be reduced during non-reinforced extinction (Rescorla & Wagner, 1972), the presentation of unexpected unpaired USs would enhance extinction learning through increased prediction errors (e.g. Fernández, Boccia, & Pedreira, 2016; Todd, Vurbic, & Bouton, 2014). As the CS-US association is weakened, or eliminated, subsequent responding to the CS would be reduced.

A second mechanism proposed by Rauhut et al. (2001) involves US habituation. However, this proposition has not been supported through previous research (Thomas et al., 2005) or the results of the present study. Comparisons of unconditioned responses to the USs presented during extinction training showed no significant differences between the partially reinforced and unpaired extinction groups. Similarly, analysis of US valence ratings showed that the US was rated as equally unpleasant in both groups, making a US habituation explanation unlikely.

Finally, Bouton et al.’s (2004) model would not readily account for the results observed in our unpaired extinction group. While the model proposes that unpaired extinction weakens the US’s exclusive association with the acquisition context, it would make the same prediction about partially reinforced extinction. Therefore, it cannot account for the differences in responding during reacquisition between the paired and unpaired US group. Overall, the present results appear to be consistent with a weakened CS-US association explanation (Rescorla & Wagner, 1972; Vervliet et al., 2010). The reduced associative strength of the CS would explain the reduced recovery from extinction effects, but this hypothesis requires further examination, as extinction is generally recognized as a form of new learning, not unlearning (Bouton, 2002). Future fear conditioning research conducted with human participants may provide further insight into the mechanisms underlying extinction training with occasionally paired or unpaired USs through an examination of US expectancy ratings (Lovibond, 2006). Assessment of US expectancies on each trial of training may provide further information about what is learned during reinforced extinction training and how this learning may affect subsequent recovery of conditioned responding. It should be also noted that while the concurrent recording of multiple indices of conditioned responding may give us a better understanding of mechanisms underlying human fear learning (Lipp, 2006a), careful consideration must be given to the experimental parameters, to prevent response interference, such as movement-induced artefacts in SCRs due to the manual handling of a US expectancy scale.

As a limitation, we should note that the interpretation of reacquisition results must be treated with caution, as the respective group comparison only yielded a trend towards significance, indicating reacquisition of extinguished SCRs in the non-reinforced and partially reinforced extinction group, but not in the unpaired extinction group. Nevertheless, the results of spontaneous recovery showed that partially reinforced and unpaired extinction provided enhanced protection from fear recovery, as indexed by SCRs, compared to non-reinforced extinction training. Further, isolated comparisons of reacquisition data from the partially reinforced and unpaired extinction group supported previous observations (Bouton et al., 2004) that an unpaired extinction procedure results in stronger reduction of reacquisition of extinguished conditioned responding than partially reinforced extinction. Hence, the present extension of Bouton et al.'s findings provides further support for the utility of an extinction procedure involving occasional presentations of paired or unpaired USs in the reduction of fear recovery in humans and serves to inform future research.

It remains to be investigated why manipulations of extinction training showed no differential effects on the reinstatement of differential SCRs. It is possible that the large number of preceding CS-only trials, meaning CS trials during extinction training and those presented during tests of spontaneous recovery, reduced the effect of the reinstatement manipulation and, thus, prevented the observation of between group differences (for a review of reinstatement research, see Haaker et al., 2014). This proposition could be tested in future research, explicitly designed to assess the effect of occasional US presentations during extinction training on reinstatement. In this regard, tests of fear recovery, including reinstatement testing, could be conducted 24 hours after the conclusion of extinction training (e.g. Schiller et al., 2010; Thompson & Lipp, 2017). Additional areas of examination involve the applicability of the present extinction procedures to fears conditioned to fear-relevant CSs, such as snakes (Lipp, 2006b; Öhman & Mineka, 2001), as well as to pre-existing fears and phobias in clinical populations. Recruitment of clinical populations will enable future research to test the translational utility of reinforced extinction, including the applicability of reinforced extinction procedures to well-established (i.e. consolidated) fear associations (Dudai, 2012; Dudai, Karni, & Born, 2015).

Another aspect requiring further examination pertains to the differential effects of extinction training with occasional US presentations on physiological and verbal indices of conditioned fear. This dissociation is in accordance with previous fear conditioning research, suggesting that response systems which are governed largely by conscious, cognitive processes, such as ratings of CS valence or US expectancy, and physiological indices of conditioned responding are differentially sensitive to (manipulations of) extinction (Culver et al., 2018; Lipp & Edwards, 2002; Lipp, Oughton, & LeLievre, 2003; Schultz et al., 2013; Thompson & Lipp, 2017). However, the differences between ratings in the partially reinforced and unpaired extinction group at the end of extinction training may also point to different underlying mechanisms, as discussed previously.

In conclusion, the counterintuitive suggestion that presentations of aversive events during extinction training may enhance extinction learning and thus reduce subsequent recovery of the extinguished fear response (Craske et al., 2014) has been supported through the results of the present study. In this regard, our findings further indicate that unpaired extinction training may be more effective in the reduction of recovery from extinction effects than partially reinforced extinction training, as unpaired extinction may guard against reacquisition of fear. When applied to the clinical setting, our results indicate that treatments focusing on the mere reduction of distress (in line with non-reinforced extinction) may be less effective in preventing fear recovery than treatments that provide clients with an opportunity to learn about the likelihood of future threat encounters (see also Craske et al., 2014; Weisman & Rodebaugh, 2018). Returning to our previous example of social anxiety disorder, during occasional exposure to a feared social situation and the feared outcome (CS-US pairing) clients may learn that the feared outcome occurs less often than expected. This learning may reduce the return of fear when the feared social situation (CS) is encountered between treatment sessions. For clients who are exposed to occasional CS-US pairings in daily life, such as negative feedback at work, it may be reassuring to know that such encounters may be beneficial for their overall treatment outcomes. In this sense, psychoeducation could be utilized to educate clients that occasional return of fear to cues in the environment is not an indicator of ineffective treatment, but an aspect that contributes to the reduction of maladaptive fears and, thereby, to the success of treatment.

We should add, however, that the translational utility of occasionally reinforced extinction requires further examination. It was not the authors' intention to suggest that a reinforced extinction procedure is readily applicable to the clinical setting, but to enhance our understanding of the mechanisms that affect extinction learning and to provide suggestions for future pre-clinical and clinical research (for a more detailed discussion of potential clinical applications, see: Craske, 2015; Craske et al., 2014; Krompinger, Van Kirk, Garner, Potluri, & Elias, 2018; Weisman & Rodebaugh, 2018). In this regard the procedure may lend itself more readily to the treatment of some, but not all, fears, phobias, or anxiety disorders. Conditions such as social anxiety disorder or specific phobias may benefit most from occasionally reinforced extinction, although the literature indicates the procedure may also be applicable to obsessive compulsive disorder (Krompinger et al., 2018). Clinical applications should also be guided by ethical considerations, in line with current exposure practices. As such, real life exposure to feared outcomes may be feasible from a pragmatic and ethical point of view in some, but not all, situations. Imaginal exposure may be considered in cases where real life exposure is not feasible; however, the application of occasionally reinforced extinction during imaginal exposure requires further investigation.

Translating the unpaired extinction procedure to the clinical setting may be slightly more challenging, but not impossible if we consider that unpaired extinction does not involve the elimination of the feared outcome, but the presentation of the feared outcome (US) at unexpected times, separate from the situation (CS) that typically predicts the arrival of the US. It is conceivable that such training could be incorporated into imaginal exposure (e.g. Abramowitz & Arch, 2014), although more research is necessary to test this hypothesis.

Acknowledgements

This work was supported by grants number DP120100750 and SR120300015 from the Australian Research Council, as well as through the contribution of an Australian Government Research Training Program Scholarship.

Competing Financial Interests

The authors declare no competing financial interests.

References

Abramowitz, J. S., & Arch, J. J. (2014). Strategies for improving long-term outcomes in cognitive behavioral therapy for obsessive-compulsive disorder: Insights from learning theory. *Cognitive and Behavioral Practice*, *21*, 20-31. doi:10.1016/j.cbpra.2013.06.004

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine, 43*, 897-910. doi:10.1017/S003329171200147X

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*, 80-99. doi:10.1037/0033-2909.114.1.80

Bouton, M. E. (2000). A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychology, 19(1S),* 57-63. doi:10.1037/0278-6133.19.Suppl1.57

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*, 976-986. doi:10.1016/S0006-3223(02)01546-9

Bouton, M. E., Woods, A. M., & Pineño, O. (2004). Occasional reinforced trials during extinction can slow the rate of rapid reacquisition. *Learning and Motivation, 35*, 371-390. doi:10.1016/j.lmot.2004.05.001

Capaldi, E. J. (1966). Partial reinforcement: A hypothesis of sequential effects. *Psychological Review, 73*, 459-477. doi:10.1037/h0023684

Capaldi, E. J. (1994). The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. *Psychonomic Bulletin & Review, 1,* 156-181. doi:10.3758/BF03200771

Carleton, R. N., Norton, M. A. P. J., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders, 21*, 105-117. doi:10.1016/j.janxdis.2006.03.014

Craske, M. G. (2015). Optimizing exposure therapy for anxiety disorders: An inhibitory learning and inhibitory regulation approach. *Verhaltenstherapie, 25*, 134-143. doi:10.1159/000381574

Craske, M. G., & Mystkowski, J. L. (2006). Exposure therapy and extinction: Clinical studies. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 37-52). Washington: American Psychological Association.

Craske, M. G., & Stein, M. B. (2016). Anxiety. *The Lancet, 388*, 3048-3059. doi:10.1016/S0140-6736(16)30381-6

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10-23. doi:10.1016/j.brat.2014.04.006

Culver, N. C., Stevens, S., Fanselow, M. S., & Craske, M. G. (2018). Building physiological toughness: Some aversive events during extinction may attenuate return of fear. *Journal of Behavior Therapy and Experimental Psychiatry, 58*, 18-28. doi:10.1016/j.jbtep.2017.07.003

Davey, G. C. L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research and Therapy, 14*, 29-66. doi:10.1016/0146-6402(92)90010-L

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Bernston (Eds.), *Handbook of psychophysiology* (pp. 159-181). Cambridge: Cambridge University Press.

Dudai, Y. (2012). The restless engram: Consolidations never end. *Annual Review of Neuroscience, 35*, 227-247. doi:10.1146/annurev-neuro-062111-150500

Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron, 88*, 20-32. doi:10.1016/j.neuron.2015.09.004

Dunsmoor, J. E., Campese, V. D., Ceceli, A. O., LeDoux, J. E., & Phelps, E. A. (2015). Novelty-facilitated extinction: Providing a novel outcome in place of an expected threat diminishes recovery of defensive responses. *Biological Psychiatry, 78*, 203-209. doi:10.1016/j.biopsych.2014.12.008

Fernández, R. S., Boccia, M. M., & Pedreira, M. E. (2016). The fate of memory: Reconsolidation and the case of prediction error. *Neuroscience & Biobehavioral Reviews, 68*, 423-441. doi:10.1016/j.neubiorev.2016.06.004

Foa, E. B., & McLean, C. P. (2016). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: The case of OCD and PTSD. *Annual Review of Clinical Psychology, 12*, 1-28. doi:10.1146/annurev-clinpsy-021815-093533

Forster, K., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124. doi:10.3758/BF03195503

Frey, P. W., & Butler, C. S. (1977). Extinction after aversive conditioning: An associative or nonassociative process? *Learning and Motivation, 8*, 1-17. doi:10.1016/0023-9690(77)90063-7

Haaker, J., Golkar, A., Hermans, D., & Lonsdorf, T. B. (2014). A review on human reinstatement studies: An overview and methodological challenges. *Learning & Memory, 21*, 424-440. doi:10.1101/lm.036053.114

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227-239. doi:10.1348/014466505X29657

Krompinger, J. W., Van Kirk, N. P., Garner, L. E., Potluri, S. I., & Elias, J. A. (2018). Hope for the worst: Occasional reinforced extinction and expectancy violation in the treatment of OCD. *Cognitive and Behavioral Practice*. doi:10.1016/j.cbpra.2017.12.002

Lipp, O. V. (2006a). Human fear learning: Contemporary procedures and measurement In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 37-52). Washington: American Psychological Association.

Lipp, O. V. (2006b). Of snakes and flowers: Does preferential detection of pictures of fear-relevant animals in visual search reflect on fear-relevance? *Emotion, 6*, 296-308. doi:10.1037/1528-3542.6.2.296

Lipp, O. V., & Edwards, M. S. (2002). Effect of instructed extinction on verbal and autonomic indices of Pavlovian learning with fear-relevant and fear-irrelevant conditional stimuli. *Journal of Psychophysiology, 16*, 176-186. doi:10.1027//0269-8803.16.3.176

Lipp, O. V., Oughton, N., & LeLievre, J. (2003). Evaluative learning in human Pavlovian conditioning: Extinct, but still there? *Learning and Motivation, 34*, 219-239. doi:10.1016/S0023-9690(03)00011-0

Lissek, S., Pine, D. S., & Grillon, C. (2006). The strong situation: A potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology, 72*, 265-270. doi:10.1016/j.biopsycho.2005.11.004

Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans - Biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience & Biobehavioral Reviews, 80*, 703-728. doi:10.1016/j.neubiorev.2017.07.007

Lovibond, P. F. (2006). Fear and avoidance: An integrated expectancy model. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 117-132). Washington: American Psychological Association.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*, 335-343. doi:10.1016/0005-7967(94)00075-U

Luck, C. C., & Lipp, O. V. (2015). A potential pathway to the relapse of fear? Conditioned negative stimulus evaluation (but not physiological responding) resists instructed extinction. *Behaviour Research and Therapy, 66*, 18-31. doi:10.1016/j.brat.2015.01.001

Lykken, D. T. (1972). Range correction applied to heart rate and to GSR data. *Psychophysiology, 9*, 373-379. doi:10.1111/j.1469-8986.1972.tb03222.x

Mickley, G. A., DiSorbo, A., Wilson, G. N., Huffman, J., Bacik, S., Hoxha, Z., . . . Kim, Y.-H. (2009). Explicit disassociation of a conditioned stimulus and unconditioned stimulus during extinction training reduces both time to asymptotic extinction and spontaneous recovery of a conditioned taste aversion. *Learning and Motivation, 40*, 209-220. doi:10.1016/j.lmot.2009.01.001

Napier, R. M., Macrae, M., & Kehoe, E. J. (1992). Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 182-192. doi:10.1037/0097-7403.18.2.182

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

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532-552. doi:10.1037/0033-295X.87.6.532

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology, 46*, 984-995. doi:10.1111/j.1469-8986.2009.00852.x

Prokasy, W. F., & Ebel, H. C. (1967). Three components of the classically conditioned GSR in human subjects. *Journal of Experimental Psychology, 73*, 247-256. doi:10.1037/h0024108

Rauhut, A. S., Thomas, B. L., & Ayres, J. J. B. (2001). Treatments that weaken Pavlovian conditioned fear and thwart its renewal in rats: Implications for treating human phobias. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 99-114. doi:10.1037/0097-7403.27.2.99

Rescorla, R. A. (1973). Effects of US habituation following conditioning. *Journal of Comparative and Physiological Psychology, 82*, 137-143. doi:10.1037/h0033815

Rescorla, R. A., & Skucy, J. C. (1969). Effect of response-independent reinforcers during extinction. *Journal of Comparative & Physiological Psychology, 67*, 381-389. doi:10.1037/h0026793

Rescorla, R. A., & Wagner, A. W. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Ricker, S. T., & Bouton, M. E. (1996). Reacquisition following extinction in appetitive conditioning. *Animal Learning & Behavior, 24*, 423-436. doi:10.3758/bf03199014

Schiller, D., Monfils, M.-H., Raio, C. M., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature, 463*, 49-53. doi:10.1038/nature08637

Schultz, D. H., Balderston, N. L., Geiger, J. A., & Helmstetter, F. J. (2013). Dissociation between implicit and explicit responses in postconditioning UCS revaluation after fear conditioning in humans. *Behavioral Neuroscience, 127*, 357-368. doi:10.1037/a0032742

Thomas, B. L., Longo, C. L., & Ayres, J. J. B. (2005). Thwarting the renewal (relapse) of conditioned fear with the explicitly unpaired procedure: Possible interpretations and implications for treating human fears and phobias. *Learning and Motivation, 36*, 374-407. doi:10.1016/j.lmot.2004.11.005

Thompson, A., & Lipp, O. V. (2017). Extinction during reconsolidation eliminates recovery of fear conditioned to fear-irrelevant and fear-relevant stimuli. *Behaviour Research and Therapy, 92*, 1-10. doi:10.1016/j.brat.2017.01.017

Todd, T. P., Vurbic, D., & Bouton, M. E. (2014). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiology of Learning and Memory, 108*, 52-64. doi:10.1016/j.nlm.2013.08.012

van den Akker, K., Havermans, R. C., & Jansen, A. (2015). Effects of occasional reinforced trials during extinction on the reacquisition of conditioned responses to food cues. *Journal of Behavior Therapy and Experimental Psychiatry, 48*, 50-58. doi:10.1016/j.jbtep.2015.02.001

Vervliet, B., Vansteenwegen, D., & Hermans, D. (2010). Unpaired shocks during extinction weaken the contextual renewal of a conditioned discrimination. *Learning and Motivation, 41*, 22-31. doi:10.1016/j.lmot.2009.08.001

Vurbic, D., & Bouton, M. E. (2014). A contemporary behavioral perspective on extinction. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley Blackwell Handbook of Operant and Classical Conditioning* (pp. 53-76). Oxford, UK: John Wiley & Sons, Ltd.

Weisman, J. S., & Rodebaugh, T. L. (2018). Exposure therapy augmentation: A review and extension of techniques informed by an inhibitory learning approach. *Clinical Psychology Review, 59*, 41-51. doi:10.1016/j.cpr.2017.10.010

Woods, A. M., & Bouton, M. E. (2007). Occasional reinforced responses during extinction can slow the rate of reacquisition of an operant response. *Learning and Motivation*, *38*, 56-74. doi:10.1016/j.lmot.2006.07.003