Behaviour Research and Therapy 120 (2019) 103416

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Reducing negative stimulus valence does not attenuate the return of fear: Two counterconditioning experiments

Eva A.M. van Dis*, Muriel A. Hagenaars, Claudi L.H. Bockting, Iris M. Engelhard

Utrecht University, Department of Clinical Psychology, the Netherlands

ARTICLE INFO

Keywords:
Counterconditioning
Evaluative conditioning
Return of fear
Fear extinction
Positive valence training

ABSTRACT

Exposure-based treatment for anxiety disorders is effective for many patients, but relapse is not uncommon. One predictor of the return of fear is the negative valence of fear-relevant stimuli. The aim of the current experiments was to examine whether counterconditioning with positive film clips reduces this negative stimulus valence as well as the return of fear, compared to standard extinction training and to an extinction training with non-contingent exposure to the positive film clips. Participants were 87 students in Experiment 1 (three-day paradigm), and 90 students in Experiment 2 (one-day paradigm). They first underwent a differential acquisition phase, in which one of three pictures was paired with an electric shock. They were then randomly allocated to one of the three intervention groups. Afterwards, they underwent a test phase in which pictures were presented without shock (to measure spontaneous recovery of fear), which was followed by unsignaled shocks to induce reinstatement of extinguished fear. Outcome variables were self-reported stimulus valence, shock expectancy, skin conductance, and fear-potentiated startle. In both experiments, counterconditioning decreased negative stimulus valence, relative to the other interventions, but it did not reduce spontaneous fear recovery or fear reinstatement. Overall, our findings do not support the notion that counterconditioning reduces return of fear.

Anxiety disorders are among the leading causes of the global burden of disease attributable to mental disorders (Whiteford et al., 2013). The gold-standard treatment is exposure-based therapy (Olatunji, Cisler, & Deacon, 2010), in which patients are exposed to fear-relevant, innocuous stimuli or situations to disconfirm their threat expectancy (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). Although the treatment is initially effective for many patients (e.g., Cuijpers, Cristea, Karyotaki, Reijnders, & Huibers, 2016; Hofmann & Smits, 2008), relapse rates of about 19–62% have been reported (see Vervliet, Craske, & Hermans, 2013).

Lab studies using fear conditioning paradigms have examined factors involved in the return of fear. These paradigms typically begin with an acquisition phase in which one neutral stimulus, such as a picture, is repeatedly followed by a negative unconditioned stimulus (US_{neg}), such as a mild electric shock, and another neutral image is not. This usually results in shock expectancy and fear reactions to the first picture, which now serves as conditioned stimulus (CS). In a subsequent extinction phase (the laboratory analog of exposure therapy), the CS is repeatedly presented without the US_{neg}. This generally extinguishes fear. Extinguished fear can return after the passage of time (i.e., spontaneous recovery), a context switch, or non-signaled US_{neg} presentations (i.e., reinstatement; see Bouton, 2002, and Vervliet et al., 2013).

The return of fear is associated with the person's negative attitude towards fear-relevant stimuli after extinction. This was found in de novo fear conditioning studies (e.g., Dirikx, Hermans, Vansteenwegen, Baeyens, & Eelen, 2004, 2007; Zbozinek, Hermans, Prenoveau, Liao, & Craske, 2015) and in subclinical studies (e.g., Huijding & De Jong, 2009; Vasey, Harbaugh, Buffington, Jones, & Fazio, 2012). This suggests that interventions that decrease negative stimulus valence may reduce the return of fear. Several experiments have examined the effects of positive valence training on the return of fear. First, Zbozinek, Holmes, and Craske (2015) found that positive imagery before extinction training, compared to positive verbal training prior to extinction, reduced negative stimulus valence and reinstatement of fear. However, as suggested by Zbozinek and Craske (2017), increased positive affect during extinction may also have enhanced extinction learning, thereby reducing the return of fear. Therefore, it is unclear whether reduced negative stimulus valence resulted in reduced reinstatement. Second, Dour, Brown, and Craske (2016) showed that positive valence training during exposure for spider fear resulted in less negative stimulus valence, less spontaneous recovery of spider fear, and less behavioural avoidance after a reinstatement manipulation. However, the positive valence training group received more exposure to spiders than the control group. Third, a recent fear conditioning study demonstrated

* Corresponding author. Department of Clinical Psychology, Utrecht University, PO Box 80140, 3508 TC, Utrecht, the Netherlands.
E-mail address: e.a.m.vandis@uu.nl (E.A.M. van Dis).

https://doi.org/10.1016/j.brat.2019.103416
Received 21 November 2018; Received in revised form 21 May 2019; Accepted 3 June 2019
Available online 04 June 2019
0005-7967/

that positive, relative to neutral, verbal information reduced negative stimulus valence but did not attenuate reinstatement (Luck & Lipp, 2018; Exp 2). Fourth, another study used counterconditioning as positive valence training (Kang, Vervliet, Engelhard, van Dis, & Hagenaars, 2018), in which the CS is paired with a positive US (see De Houwer, Thomas, & Baeyens, 2001). Lab experiments have shown that counterconditioning reduces negative stimulus valence (e.g., Engelhard, Leer, Lange, & Olatunji, 2014; Kerkhof, Vansteenwegen, Baeyens, & Hermans, 2011; Raes & De Raedt, 2012). Kang et al. (2018) found that, compared to extinction training, counterconditioning attenuated the return of fear, which was measured with threat expectancy. Unexpectedly, the groups did not differ in negative stimulus valence, perhaps because the positive US (comic pictures) was not potent enough to affect evaluative learning. Given the mixed results of these studies, more controlled research is needed before positive valence training is implemented in clinical practice to reduce return of fear.

The aim of the current research was to examine whether positive valence training through counterconditioning attenuates the return of fear (while controlling for positive affect and amount of exposure). Participants underwent fear acquisition and were then randomly allocated to one of three groups: counterconditioning (in which the CS was paired with positive film clips, which can be more potent than static images; Rottenberg, Ray, & Gross, 2007), extinction training, or extinction training with unpaired presentations of positive film clips (to control for positive affect induced by the film clips). We expected that counterconditioning, compared to the other interventions, would reduce post-intervention negative stimulus valence (hypothesis 1) and the return of fear (i.e., spontaneous fear recovery and reinstatement; hypothesis 2).

1. Experiment 1

We used a three-day fear conditioning paradigm, over a period of nine days, with the following phases: Acquisition (Day 1), Intervention (Day 2), and Spontaneous Recovery and Reinstatement (Day 9).

1.1. Method

Participants. Ninety-eight native Dutch-speaking students aged between 18 and 30 were recruited via Utrecht University, Facebook, and Proefbunny.nl. The exclusion criteria were: self-reported current psychiatric diagnosis, a history of heart or epileptic problems, oversensitivity to loud noises, pregnancy, psychoactive medication use, and fear of dogs (see Stimuli below). Six participants were excluded because of unsuccessful fear acquisition (n = 3), negatively rating the positive film clips (n = 2), and misunderstanding how to use the US_{neg} expectancy scale (n = 1). Five additional participants dropped out: two because of the experimenter falling ill, two because they found the shock/startle probe too unpleasant, and one was a no-show on Day 3. The final sample size comprised 87 participants (72 females and 25 males; mean age = 21.39, SD = 2.48) that were randomly assigned to condition (stratified for gender). This study was approved by the ethics committee of the Faculty of Social and Behavioural Sciences at Utrecht University (FETC16-054) and was pre-registered on the Open Science Framework (https://osf.io/bvfx8/).

Stimuli. One CS+ and two CSs- depicted three neutral faces derived from the Chicago Faces Database (Ma, Correll, & Wittenbrink, 2015) that were presented for 8 s (see below). To enhance stimulus differentiation, the CS+ was a picture of a male face, and the CSs- were pictures of female faces (or vice versa). The stimuli were fully counterbalanced across participants.

The negative US (US_{neg}) was a 2-ms electric shock that was delivered by a Digitimer DS7A through an electrode band that was attached to the wrist of the dominant arm. On Day 1, participants determined a "really annoying, but not painful" shock intensity during a standard work-up procedure (as described by Orr et al., 2000).

The positive US (US_{pos}) consisted of eight different 6-s funny film clips in which a baby laughs at a dog that is trying to catch soap bubbles from the air. The fragments were derived from a 59-s YouTube video (see Jess0rT, 2011).

1.2. Measures

Eysenck Personality Questionnaire (EPQ). Neuroticism was measured with the Dutch translation of the Neuroticism scale of the Eysenck Personality Questionnaire (EPQ-N; Sanderman, Arrindell, Ranchor, Eysenck, & Eysenck, 2012; Sanderman, Arrindell, & Ranchor, 1991). It consists of 22 self-report items (e.g., "Are you often troubled about feelings of guilt?") that are rated on a dichotomous scale (0 = *no*, 1 = *yes*). Cronbach's α was 0.82 in the present study.

US_{neg} unpleasantness and expectancy. US_{neg} unpleasantness was rated on an 11-point scale, ranging from 0 (*not unpleasant at all*) to 10 (*very unpleasant*). US_{neg} expectancy was rated with a visual analog scale (VAS) with three anchors: 0 (*certainly no shock*), 50 (*uncertain*), and 100 (*certainly a shock*).

US_{pos} valence. US_{pos} valence was rated on an 11-point scale, ranging from 0 (*negative*) to 10 (*positive*).

CS valence, arousal, and fear. For each CS rating, the CS was presented on a white background to enhance differentiation between the rating contexts and experimental trials. CS valence ("How negative or positive do you find this picture?") and arousal ("How arousing do you find this picture?") were rated using the Self-Assessment Manikin scale (Bradley & Lang, 1994), ranging from 1 (*negative/not arousing*) to 9 (*positive/arousing*). Fear (i.e., "How fearful are you when seeing this picture?") was rated on a 10-point Likert-scale, ranging from 1 (*not fearful at all*) to 9 (*very fearful*).[1]

Affect rating. Affect ("How do you feel at this moment?") was rated on a VAS ranging from 0 (*unpleasant*) to 100 (*pleasant*).

Physiological measures. BioSemi hardware unit and ActiView 7.06 were used to acquire physiological data at a 2048 Hz sampling rate. Two 4-mm Ag–AgCl CMS/DRL electrodes were positioned on the forehead and served as a reference for all physiological measurements.

Skin conductance response (SCR). Skin conductance was recorded using two 8-mm passive Nihon Kohden electrodes that were attached to the index and middle fingers of the left hand (all participants were right-handed). SCR was calculated by subtracting the baseline (mean activity during 2 s immediately prior to CS onset) from the peak skin conductance activity between 1 and 7 s after CS onset (see Pineles, Orr, & Orr, 2009). Negative values and values smaller than 0.01 μS were recoded to 0. We applied a z-transformation of each raw SCR across all phases to account for inter-individual variance (as recommended by Lonsdorf et al., 2017).

Fear-potentiated startle (FPS). Orbicularis oculi activity was recorded using two 4-mm Ag/AgCl electrodes placed under the left eye (see Blumenthal et al., 2005). FPS was calculated as the difference score between the peak (21–150 ms after startle probe onset) from the baseline (mean activity during 50 ms, starting from 30 ms before startle probe onset). We applied an intra-individual z-transformation across all phases of the raw FPS data (see Blumenthal et al., 2005).

Post-auricular reflex. The post-auricular reflex is an implicit, psychophysiological measure of valence (e.g., Benning, Patrick, & Lang, 2004). It was measured only to explore its value for future studies but there was no differential responding to CSs during the acquisition phase. Therefore, it is not mentioned further.

Trial procedure. The trial procedure is largely based on Zbozinek, Hermans, et al. (2015) and Zbozinek, Holmes, et al. (2015). Each trial began with an 8-s CS presentation (in the middle of the computer screen on a black background). During the first 6 s, participants rated US_{neg}

[1] Results on self-reported arousal and fear are reported in the Supplementary Materials.

Table 1
Overview of the experimental design of Experiment 1.

Group	Day 1		Day 2	Day 9		
	Habituation	Acquisition	Intervention	Spontaneous recovery	Reinstatement	Test
CC	CS+ (2)	CS+/US_{neg} (8)	CS+/US_{pos} (8)	CS+ (2)	US_{neg} (3)	CS+ (2)
	CS1- (2)	CS1- (8)	CS1- (8)	CS1- (2)		CS1- (2)
	CS2- (2)	CS2- (8)	CS2- (8)	CS2- (2)		CS2- (2)
EXT	CS+ (2)	CS+/US_{neg} (8)	CS+ (8)	CS+ (2)	US_{neg} (3)	CS+ (2)
	CS1- (2)	CS1- (8)	CS1- (8)	CS1- (2)		CS1- (2)
	CS2- (2)	CS2- (8)	CS2- (8)	CS2- (2)		CS2- (2))
EXT+	CS+ (2)	CS+/US_{neg} (8)	CS+ (8)	CS+ (2)	US_{neg} (3)	CS+ (2)
	CS1- (2)	CS1- (8)	CS1- (8)	CS1- (2)		CS1- (2)
	CS2- (2)	CS2- (8)	CS2-/US_{pos} (8)	CS2- (2)		CS2- (2)

Note. CC = counterconditioning; CS = conditioned stimulus; EXT = extinction; EXT+ = extinction and positive material; US_{neg} = negative unconditioned stimulus.

expectancy. Startle probes (50 ms bursts of white noise at 95 dB) were presented 7 s after CS onset.

During the acquisition phase, the US_{neg} was presented 7.5 s after CS onset during all CS + presentations. In all phases, CSs were presented in a pseudorandom order (with a maximum of two consecutive CS presentations per phase). Trials ended with a 6-s presentation of a black screen (or a 6-s presentation of US_{pos}, see below). The inter-trial interval (ITI) was a black screen with a white fixation cross that appeared for 15, 20 or 25 s (counterbalanced). Noise alone (NA) trials followed one out of three trials (counterbalanced) to measure baseline startle responding.

General procedure and Intervention. Table 1 shows the general procedure. On Day 1, participants provided written informed consent. They then washed their hands without soap, were connected to physiological and shock electrodes, and began the shock work-up procedure. Next, they rated unpleasantness of the US_{neg} and completed the EPQ-N. Then, they started the Habituation phase, consisting of 6 CS presentations trials (in which they practiced rating the US_{neg} expectancy scale) and 10 startle probes. Participants then completed CS and affect ratings. Thereafter, the Acquisition phase started, in which each CS was presented 8 times. The CS+ was always paired with US_{neg}, whereas CS1- and CS2- were not. Finally, participants rated CS valence and affect again.

On Day 2, participants were reconnected to physiological and shock electrodes and received 10 startle probes to habituate. They filled out the CS and affect ratings and continued with the intervention phase, in which each CS was presented 8 times and was never followed by the US_{neg}. Reinforcement differed for the groups. In the counterconditioning group (CC), the CS+ was always followed by $US_{pos.}$, in the extinction group (EXT), there was no reinforcement, and in the extinction with positive material group (EXT+), CS2- was always followed by US_{pos}. Participants then rated the CS and affect scales.

On Day 9, participants were again reconnected to physiological and shock electrodes and received 10 startle probes to habituate. After completing the CS and affect ratings, they continued with a spontaneous recovery phase in which each CS was presented twice without reinforcement (see Zbozinek, Holmes, et al., 2015). Participants then completed CS and affect ratings again. Next, they received three non-signaled shocks (same intensity as Day 1) with 15 and 20 s inter-stimulus intervals (fixed order). They continued with a reinstatement test phase in which each CS was presented twice without reinforcement (see Kang et al., 2018; Zbozinek, Holmes, et al., 2015). Participants then completed the CS, affect, and US_{pos} ratings. Finally, they were debriefed and received 20 euros or course credits.

Data preparation. Due to technical problems, data were missing completely at random (MCAR; van Buuren, 2012) for SCR, FPS (n = 9), and neuroticism scores (EPQ-N; n = 27). Multiple imputation techniques were not applied due to an unconnected file matching missing data pattern (e.g., missing data for SCR and FPS during Day 1 and 2, but not for Day 9).

Data analysis. First, to test whether randomization was successful, one-way ANOVAs were performed on age, neuroticism scores, shock level, US_{neg} unpleasantness, and affect. Second, to examine whether fear acquisition and extinction took place for US_{neg} expectancy, FPS, and SCR, we used three 3 (Stimulus: CS+, CS1-, CS2-) × 8 (Time: all acquisition or intervention trials) × 3 (Group: CC, EXT, EXT+) mixed ANOVAs. Third, to test whether CC and EXT+ groups had higher affect ratings compared to EXT following the Intervention phase, a 2 (Time: pre-intervention, post-intervention) × 3 (Group: CC, EXT, EXT+) mixed ANOVA was used.

To test the first Hypothesis on post-intervention group differences in CS valence, a 3 (Stimulus: CS+, CS1-, CS2-) × 2 (Time: pre-intervention, post-intervention) × 3 (Group: CC, EXT, EXT+) mixed ANOVA was conducted. To test the second hypothesis on group differences in spontaneous recovery and reinstatement, separate 3 (Stimulus: CS+, CS1-, CS2-) × 2 (Time: last Intervention trial, first Spontaneous recovery trial; *or* last Spontaneous recovery trial, first Reinstatement trial) × 3 (Group: CC, EXT, EXT+) ANOVAs were used for US_{neg} expectancy, SCR, and FPS.

In case sphericity assumptions were not met, we applied Huynh-Feldt ($\varepsilon > 0.75$) or Greenhouse-Geisser ($\varepsilon < 0.75$) corrections.

1.3. Results

Randomization checks. Groups did not significantly differ in age, neuroticism scores, shock level, US_{neg} unpleasantness, or baseline affect (all $Fs < 2.14$, all $ps > .123$), which suggests that randomization was successful. Evaluations of positive film clips also did not significantly differ between the CC (M = 8.28, SD = 1.14) and EXT+ (M = 8.44, SD = 1.42) groups, $t(50) < 1$, p = .649.

Acquisition phase. Acquisition (Stimulus × Time) was reflected by increases in US_{neg} expectancy, SCR, and FPS, with stronger responses to the CS+ over time compared to both s- (Stimulus × Time: all $Fs > 4.06$, $ps < .001$, simple effects: all $ts > 2.64$, $ps < .011$), see Fig. 1. From pre to post acquisition, CS negative valence increased for CS+, but not for CSs-, Stimulus × Time: $F(1.69, 141.72) = 53.58$, $p < .001$, simple effects: $t(86) = 8.67$, $p < .001$, see Table 2. There were no three-way interactions with Group (all $Fs < 1.21$, all $ps > .245$).

Intervention phase. US_{neg} expectancy and SCR decreased during the intervention phase, with larger decreases for the CS+ compared to both CSs- (Stimulus × Time: both $Fs > 3.45$, $ps < .001$; simple effects: both $ts > 2.22$, $ps < .030$). There was no significant Stimulus × Time interaction for FPS, $F(12.48, 1047.88) < 1$, p = .487, but there was a main effect for Time, $F(4.36, 366.27) = 80.45$, $p < .001$, which reflects decreased responding during the intervention phase, and there was a main effect for Stimulus, $F(2, 168) = 15.78$, $p < .001$. There were no significant three-way interactions with Group

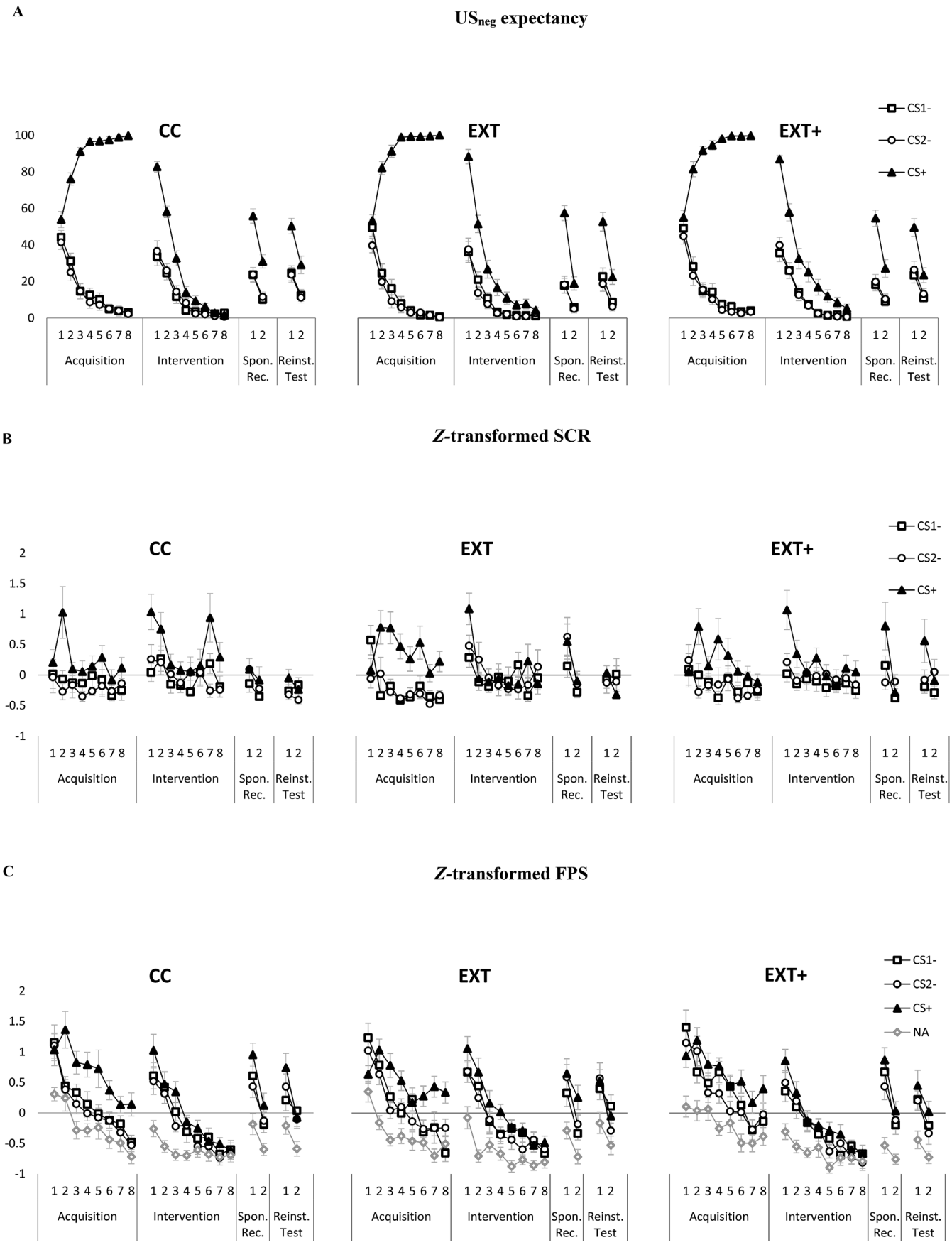

Fig. 1. US_{neg} expectancy, skin conductance response (SCR), and fear potentiated startle response (FPS) in Experiment 1. CC = counterconditioning; CS = conditioned stimulus; EXT = extinction; EXT+ = extinction and positive material; NA = noise alone; Spon. Rec. = Spontaneous Recovery; Reinst. = Reinstatement; US_{neg} = negative unconditioned stimulus.

Table 2
Means (standard deviations) of CS valence and affect ratings in Experiment 1.

		Acquisition		Intervention		Spontaneous Recovery		Reinstatement
		Pre	Post	Pre	Post	Pre	Post	Post
CC	CS+	4.97 (1.43)	3.21 (1.82)	3.79 (1.37)	5.93 (1.33)	5.76 (1.15)	5.76 (1.09)	5.07 (1.16)
	CS1-	4.66 (1.61)	5.41 (1.62)	5.62 (1.21)	5.52 (1.30)	5.38 (1.05)	5.31 (1.23)	4.55 (1.48)
	CS2-	4.83 (1.65)	5.52 (1.60)	5.31 (1.51)	5.72 (1.46)	5.64 (1.37)	5.86 (1.21)	5.43 (1.20)
	Affect	68.79 (16.17)	58.52 (15.77)	68.83 (17.92)	74.45 (13.98)	72.79 (17.95)	74.79 (16.85)	68.38 (17.34)
EXT	CS+	5.24 (1.81)	2.79 (1.76)	3.69 (1.39)	4.97 (1.12)	5.62 (1.40)	6.03 (1.18)	4.72 (1.18)
	CS1-	5.10 (1.61)	5.79 (1.78)	5.66 (1.32)	6.03 (1.30)	5.45 (1.39)	5.66 (1.29)	4.28 (1.31)
	CS2-	5.38 (1.45)	6.14 (1.53)	6.00 (1.49)	6.07 (1.49)	5.79 (1.29)	5.97 (1.21)	4.83 (1.34)
	Affect	60.52 (18.21)	49.93 (20.35)	66.00 (19.81)	64.28 (20.99)	73.38 (18.88)	73.93 (19.95)	65.15 (21.93)
EXT+	CS+	5.00 (1.67)	3.10 (1.52)	3.76 (1.53)	5.00 (1.22)	5.83 (1.20)	6.03 (1.38)	5.10 (1.37)
	CS1-	4.69 (1.61)	5.69 (1.77)	5.79 (1.54)	5.83 (1.73)	5.66 (1.47)	6.14 (1.27)	4.52 (1.30)
	CS2-	5.21 (1.26)	6.03 (1.32)	5.72 (1.58)	6.14 (1.92)	5.89 (1.63)	5.85 (1.35)	4.93 (1.69)
	Affect	60.48 (18.36)	52.24 (18.23)	67.45 (16.26)	71.52 (18.98)	73.83 (17.98)	78.07 (17.67)	65.96 (21.48)

Note. CC = counterconditioning; CS = conditioned stimulus; EXT = extinction; EXT+ = extinction and positive material.

(all $Fs < 1$, all $ps > .690$).

For the affect ratings, there was a trend for a Time (pre, post intervention) × Group interaction, $F(2, 84) = 2.45$, $p = .092$, $\eta_p^2 = 0.06$. Post-hoc analyses showed that the intervention phase increased positive affect for the CC group ($M = 74.45$, $SD = 13.98$), $t(28) = 2.62$, $p = .014$, and the EXT+ group ($M = 71.52$, $SD = 18.98$), $t(28) = 2.22$, $p = .035$, but not for the EXT group ($M = 64.28$, $SD = 20.99$), $t(28) < 1$, $p = .596$ (see Table 2). The increase in positive affect did not differ between the CC and EXT+ groups, $t(56) < 1$, $p = .585$. Before the Intervention phase, affect ratings did not differ across groups, $F(2, 84) < 1$, $p = .837$, $\eta_p^2 = 0.00$, but there was a non-significant trend for a Group effect afterwards, $F(2, 84) = 2.39$, $p = .097$, $\eta_p^2 = 0.05$.

Hypothesis 1. More positive CS + valence after CC, compared to EXT and EXT+. There was a Stimulus × Time × Group interaction for CS valence, $F(3.55, 149.08) < 1$, $p = .035$, $\eta_p^2 = 0.06$. The Stimulus × Group interaction was not significant before the intervention, $F(4, 168) < 1$, $p = .523$, $\eta_p^2 = 0.02$, but it was significant afterwards, $F(4, 168) = 4.69$, $p = .002$, $\eta_p^2 = 0.10$. Post-hoc analyses with Bonferroni corrections revealed that after the intervention, CS+ valence was more positive in the CC group compared to both control groups ($ps < .015$), and did not differ between the EXT and EXT+ groups ($p = .999$). These results support the hypothesis that negative CS+ valence was more strongly reduced in the CC group relative to the EXT and EXT+ groups, see Table 2. However, on Day 9, the Stimulus × Group effect was no longer significant, $F(4, 162) < 1$, $p = .850$, $\eta_p^2 = 0.01$.

Hypothesis 2a. Less spontaneous recovery in CC relative to EXT and EXT+.

US_{neg} expectancy. The Stimulus × Time interaction was significant, $F(1.63, 137.12) = 99.04$, $p < .001$, $\eta_p^2 = 0.54$. The increase in US_{neg} expectancy was stronger for CS+ than for both CSs- (both $ts > 10.36$, $ps < .001$), which demonstrates differential spontaneous recovery (see Fig. 1A). However, interaction effects with Group were not significant (all $Fs < 1$, $ps > .407$).

SCR. The Stimulus × Time interaction was not significant, $F(1.85, 155.09) < 1$, $p = .620$, $\eta_p^2 = 0.01$, but SCR increased over time, $F(1, 84) = 4.02$, $p = .048$, $\eta_p^2 = 0.05$, which indicates the occurrence of non-differential spontaneous recovery (see Fig. 1B). There was also a main effect for Stimulus, $F(1.92, 161.29) = 5.78$, $p = .004$, $\eta_p^2 = 0.06$. However, again, interaction effects of Group were not significant (all $Fs < 2.17$, $ps > .078$).

FPS. The Stimulus × Time interaction was not significant, $F(2, 168) < 1$, $p = .375$, $\eta_p^2 = 0.01$, but there was a main effect for Time, $F(1, 84) = 89.83$, $p < .001$, which indicates a non-differential spontaneous recovery effect (see Fig. 1C), and for Stimulus, $F(2, 168) = 3.96$, $p = .021$, $\eta_p^2 = 0.05$. Again, interaction effects with Group were not significant (all $Fs < 1.01$, $ps > .406$).

Hypothesis 2b. Less reinstatement in CC relative to EXT and EXT+.

US_{neg} expectancy. The Stimulus × Time interaction was significant, $F(1.33, 111.63) = 9.20$, $p < .001$, $\eta_p^2 = 0.10$ (see Fig. 1A), due to a stronger increase in US_{neg} expectancy following the CS+, compared to both CSs- (both $ts > 2.96$, $ps < .004$). This means that differential reinstatement was successful. However, there were no interaction effects of Group (all $Fs < 1.05$, all $ps > .354$).

SCR. There was no Stimulus × Time interaction, $F(2, 168) = 2.11$, $p = .124$, $\eta_p^2 = 0.03$, but there was a main effect for Time, $F(1, 84) = 4.45$, $p = .038$, $\eta_p^2 = 0.05$, which demonstrates non-differential reinstatement (see Fig. 1B), and Stimulus, $F(1.94, 162.74) = 3.48$, $p = .034$, $\eta_p^2 = 0.04$. There were no significant interactions with Group (all $Fs < 1.69$, all $ps > .192$).

FPS. Again, the Stimulus × Time interaction was not significant, $F(2, 168) < 1$, $p = .875$, $\eta_p^2 = 0.00$, but there was a main effect for Time, $F(1, 84) = 35.43$, $p < .001$, $\eta_p^2 = 0.30$, which indicates non-differential reinstatement (see Fig. 1C), and Stimulus, $F(1.91, 160.47) = 7.27$, $p = .001$, $\eta_p^2 = 0.08$. There were no interaction effects of Group (all $Fs < 1$, $ps > .612$).

In summary, spontaneous recovery and reinstatement were differential for US_{neg} expectancy ratings, and non-differential for SCR and FPS. However, in contrast to our hypotheses, CC did not attenuate spontaneous recovery or reinstatement of fear relative to the control groups.

Exploratory analyses. Regression analyses were performed to explore whether post-intervention CS+ positive valence predicted less spontaneous recovery (i.e., difference between CS+ trial 1 in spontaneous recovery phase and CS+ trial 8 in intervention phase) and reinstatement (i.e., difference between CS+ trial 1 in reinstatement phase and CS+ trial 2 in spontaneous recovery phase), measured with US_{neg} expectancy, FPS, or SCR. This was only the case for spontaneous recovery measured with SCR (Beta = −0.22, $p = .039$); (all other Betas < 0.17, $ps > .136$).

1.4. Discussion Experiment 1

As predicted, counterconditioning outperformed both extinction procedures in reducing negative stimulus valence at the end of Day 2. However, it did not reduce the return of fear on any of the outcome measures one week later. To our knowledge, only one earlier multiple-day fear conditioning study has been published that showed a correlation between post-extinction CS+ negative valence (Day 2) and

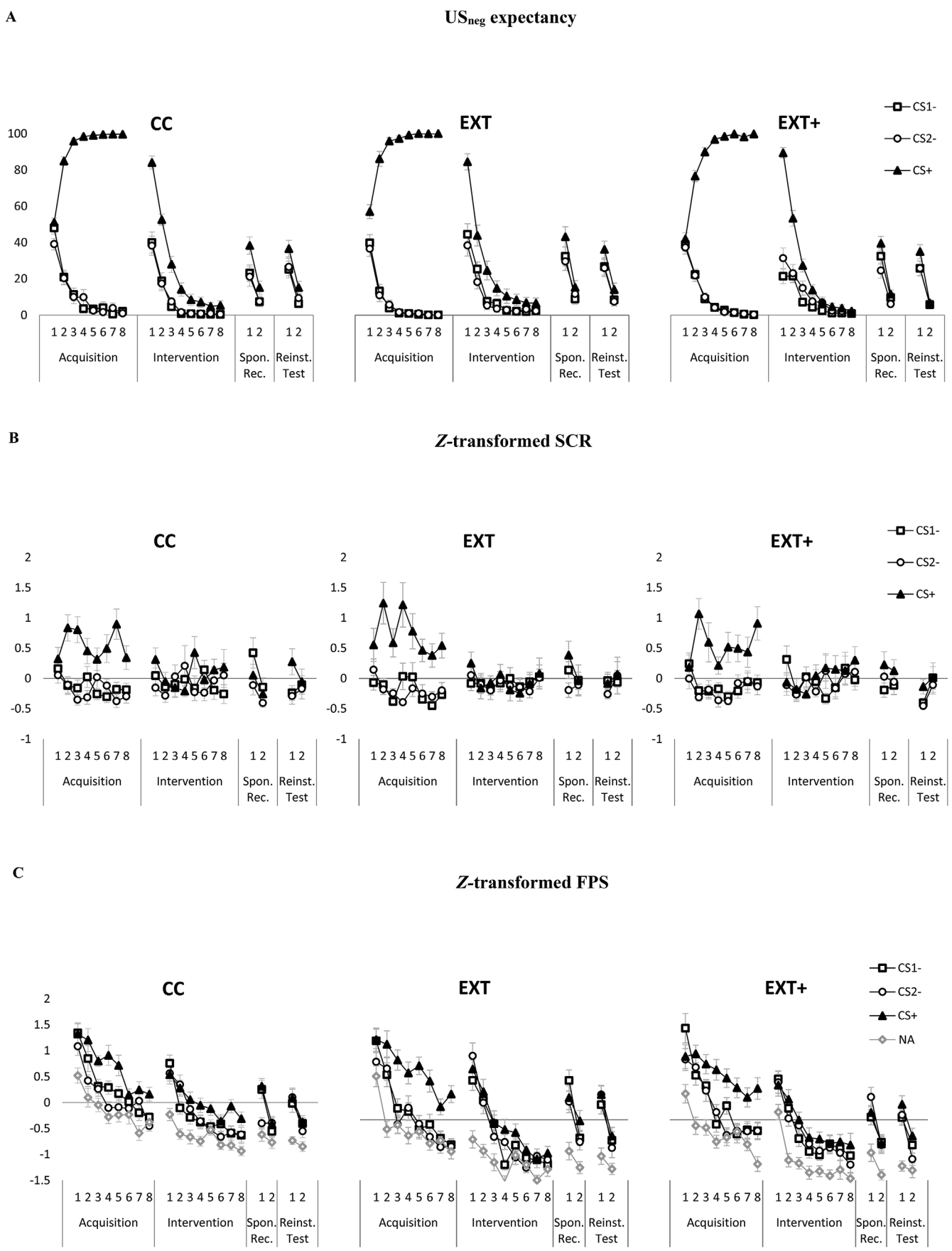

Fig. 2. US_{neg} expectancy, skin conductance response (SCR), and fear potentiated startle response (FPS) in Experiment 2. CC = counterconditioning; CS = conditioned stimulus; EXT = extinction; EXT+ = extinction and positive material; NA = noise alone; Spon. Rec. = Spontaneous Recovery; Reinst. = Reinstatement; US_{neg} = negative unconditioned stimulus.

Table 3
Means (standard deviations) of CS valence and affect ratings in Experiment 2.

		Acquisition		Intervention	Spontaneous Recovery	Reinstatement
		Pre	Post	Post	Post	Post
CC	CS+	5.13 (1.46)	3.00 (1.46)	5.53 (1.66)	5.53 (1.68)	5.60 1.30)
	CS1-	4.93 (1.72)	6.13 (1.61)	5.97 (1.77)	6.13 (1.76)	5.93 (1.80)
	CS2-	4.93 (1.68)	6.00 (1.68)	5.93 (2.00)	6.00 (1.53)	6.17 (1.60)
	Affect	64.10 (19.78)	55.00 (18.27)	68.23 (15.80)	68.97 (17.66)	64.67 (19.96)
EXT	CS+	4.87 (1.76)	2.70 (1.62)	4.43 (1.43)	4.83 (1.39)	5.30 (1.70)
	CS1-	4.97 (1.52)	6.03 (1.63)	5.60 (1.45)	5.80 (1.30)	5.90 (1.67)
	CS2-	5.50 (1.17)	6.30 (1.49)	5.90 (1.60)	6.07 (1.51)	6.17 (1.62)
	Affect	66.10 (18.76)	55.93 (17.46)	61.67 (15.42)	63.50 (15.57)	63.23 (16.79)
EXT+	CS+	5.17 (1.76)	3.20 (1.52)	5.27 (1.31)	5.33 (1.37)	5.47 (1.53)
	CS1-	4.93 (1.62)	6.10 (1.60)	5.53 (1.61)	5.70 (1.29)	5.70 (1.32)
	CS2-	5.03 (1.59)	5.73 (1.80)	6.64 (1.73)	6.03 (1.50)	6.00 (1.49)
	Affect	61.40 (18.40)	55.23 (19.47)	67.80 (16.39)	65.87 (18.59)	61.43 (16.79)

Note. CC = counterconditioning; CS = conditioned stimulus; EXT = extinction; EXT+ = extinction and positive material.

reinstatement of fear (Day 3) (Zbozinek, Hermans, et al., 2015). However, their post-extinction CS+ valence ratings did not differ from the pre-spontaneous recovery and pre-reinstatement ratings, whereas, in our experiment, the group effects on CS+ valence did not persist for one week. Other fear conditioning experiments that detected a positive correlation between post-extinction negative stimulus valence and the return of fear used a one-day fear conditioning paradigm (e.g., Dirikx, Hermans, Vansteenwegen, Baeyens, & Eelen, 2007, 2004; Hermans et al., 2005), in which a reinstatement phase took place immediately after post-extinction CS valence ratings. Hence, in those one-day studies, return of fear may be related to negative stimulus valence right before reinstatement rather than to the valence right after extinction (one week earlier in our experiment). Theoretically, there is no reason to assume that the return of fear would be reduced if the stimulus is negative before reinstatement, regardless of whether it was positive before. To test whether return of fear is associated with reduced negative stimulus valence right before a return of fear test, we decided to conduct a replication experiment with all phases on one day.

2. Experiment 2

2.1. Method

Experiment 2 was a replication of Experiment 1, except that the Acquisition, Extinction, Spontaneous Recovery and Reinstatement phases took place on one day, and the post-auricular reflex was not measured. Results on self-reported arousal and fear are reported in the Supplementary Materials.

Participants. Ninety-eight participants were recruited at Utrecht University and via the Internet. Two participants dropped out because they found the shock/startle probe too unpleasant. Data from three participants were excluded because they rated the positive film clips negatively ($n = 1$), did not follow instructions ($n = 1$), or misunderstood how to use the US_{neg} expectancy scale ($n = 1$). Due to a counterbalancing error in the Extinction group, and before data analysis, we excluded 33 EXT participants and tested an additional 30 participants in this group. The final sample included 90 participants (75 females and 15 males; mean age = 21.18, $SD = 2.01$) who were equally distributed across groups. Participants received 12 euros or course credits as compensation for their participation. This study received approval from the ethics committee of the Faculty of Social and Behavioural Sciences at Utrecht University (FETC16-054) and was pre-registered on the Open Science Framework (https://osf.io/mxa5z/).

General procedure. The procedure of Experiment 1 was executed on one day. After participants provided informed consent, they were connected to physiological and shock electrodes and continued with the shock work-up procedure. Then, they completed the EPQ-N, followed by the Habituation, Acquisition, Intervention, Spontaneous Recovery (starting with a CS1- or CS2- presentation) and Reinstatement phases. Between these phases, they rated CS valence, arousal, and fear and affect. Afterwards, participants were thanked and debriefed.

2.2. Results

Randomization checks. There were no significant group differences in age, neuroticism scores, shock level, and baseline affect (all $Fs < 2.71$, all $ps > .072$). However, groups differed in US_{neg} unpleasantness, $F(2, 87) = 3.77$, $p = .027$, $\eta_p^2 = 0.08$, with lower shock evaluation in the EXT compared to the EXT+ group (mean difference = 0.50), $t(58) = 2.55$, $p = .013$. There were no other group differences ($ts < 1.54$, $ps > .128$). The evaluation of film clips were not different between the CC ($M = 8.42$, $SD = 0.96$) and EXT+ ($M = 8.55$, $SD = 1.02$) groups, $F(1, 60) < 1$, $p = .609$, $\eta_p^2 = 0.00$.

Acquisition phase. Acquisition was reflected by US_{neg} expectancy, SCR, and FPS variables, with stronger responding over time to the CS+ relative to both CSs- (Stimulus × Time: all $Fs > 3.12$, $ps < .001$; simple effects: all $ts > 2.97$, $ps < .005$), see Fig. 2. After the acquisition phase, negative valence had increased for CS+ (but not for CSs-), Stimulus × Time: $F(1.70, 169.97) = 97.30$, $p < .001$, simple effects: $t(89) = 11.63$, $p < .001$, see Table 3. There were no three-way interaction effects (all $Fs < 1.73$, all $ps > .072$), indicating that groups did not differ in acquisition.

Intervention phase. US_{neg} expectancy decreased during the Intervention phase; Stimulus × Time: $F(3.59, 311.97) = 49.82$, $p < .001$, $\eta_p^2 = 0.36$, simple effects: $t(89) = 10.24$, $p < .001$. For SCR, we did not observe a significant Stimulus × Time interaction, $F(13.66, 1188.39) < 1$, $p = .653$, $\eta_p^2 = 0.01$, Time, $F(6.61, 575.09) = 1.50$, $p = .169$, $\eta_p^2 = 0.02$, or main effect for Stimulus, $F(2, 174) = 1.91$, $p = .152$, $\eta_p^2 = 0.02$. With respect to FPS, there was no Stimulus × Time effect, $F(13.56, 1179.71) = 1.21$, $p = .263$, $\eta_p^2 = 0.01$, but there were main effects for Time, $F(6.11, 531.13) = 86.00$, $p < .001$, $\eta_p^2 = 0.50$, and Stimulus, $F(1.93, 168.16) = 14.38$, $p < .001$, $\eta_p^2 = 0.14$ (see Fig. 2). There were no three-way interaction effects (all $Fs < 1.98$, all $ps > .056$).

The Time (pre, post intervention) × Group interaction was not significant for affect ratings, $F(2, 87) = 2.31$, $p = .106$, $\eta_p^2 = 0.05$. Positive affect increased in all groups from pre to post intervention, $F(1, 87) = 44.38$, $p < .001$, $\eta_p^2 = 0.34$ (see Table 3).

In summary, extinction occurred for US_{neg} expectancy and FPS. Because SCR did not decrease during the Intervention phase, we did not examine spontaneous recovery and reinstatement effects for SCR (see Fig. 2B for visual inspection).

Hypothesis 1. More positive CS+ valence after CC, compared to EXT and EXT+. There was no significant Stimulus × Time (pre, post intervention) × Group interaction for CS valence, $F(3.46, 150.58) = 1.82$, $p = .138$, $\eta_p^2 = 0.04$, but there was a marginally significant Stimulus × Group effect after the intervention, $F(3.74, 162.60) = 2.47$, $p = .051$, $\eta_p^2 = 0.05$, see Table 3. Post-hoc analyses adjusted with Bonferroni correction showed that after the intervention phase, the CC group reported more positive CS+ valence compared to the EXT group ($p = .014$), but not compared to the EXT+ group ($p = .999$). CS+ evaluation did not differ significantly between the EXT and EXT+ groups ($p = .093$). These results partially support our hypothesis that negative CS+ valence would be lower after the intervention phase in the CC group relative to the other two groups.

Hypothesis 2a. Less spontaneous recovery in CC relative to EXT and EXT+.

US_{neg} expectancy. There was a significant Stimulus × Time effect, $F(1.88, 163.61) = 10.34$, $p < .001$, $\eta_p^2 = 0.11$, with a stronger increase in US_{neg} expectancy to the CS+ compared to both CSs- (both $ts > 2.64$, both $ps < .010$). This indicates successful differential spontaneous recovery. There were no interaction effects of Group (all $Fs < 1$, all $ps > .561$), see Fig. 2A.

FPS. There was no significant Stimulus × Time effect, $F(2, 174) = 2.32$, $p = .102$, $\eta_p^2 = 0.02$. There were main effects for Time, $F(1, 87) = 104.86$, $p < .001$, $\eta_p^2 = 0.55$, which demonstrates non-differential spontaneous recovery, and Stimulus, $F(2, 174) = 3.27$, $p = .040$, $\eta_p^2 = 0.04$. The Stimulus × Time × Group effect was also significant, $F(4, 174) = 3.08$, $p = .018$, $\eta_p^2 = 0.07$. FPS increased over time for all stimuli in all groups ($ts > 1.99$, $ps < .056$), but not for CS2- in the CC group, $t(29) = 1.42$, $p = .166$, see Fig. 2C.

Hypothesis 2b. Less reinstatement in CC relative to EXT and EXT+.

US_{neg} expectancy. A Stimulus × Time interaction, $F(1.73, 150.27) = 3.58$, $p = .037$, $\eta_p^2 = 0.04$, revealed a stronger increase in US_{neg} expectancy following the CS+ compared to both CSs- (both $ts > 2.10$, $ps < .039$). This demonstrates successful differential reinstatement. There were no interaction effects including Group (all $Fs < 1$, all $ps > .559$), see Fig. 2A.

FPS. There was no significant Stimulus × Time effect, $F(2, 174) < 1$, $p = .956$, $\eta_p^2 = 0.00$. A main effect of Time, $F(1, 87) = 59.36$, $p < .001$, $\eta_p^2 = 0.41$, indicated non-differential reinstatement. There were no main effect for Stimulus, $F(2, 174) = 2.17$, $p = .118$, $\eta_p^2 = 0.02$, or interaction effects with Group (all $Fs < 1$, all $ps > .667$), see Fig. 2C.

These results suggest that spontaneous recovery and reinstatement occurred for US_{neg} expectancy (i.e., differential return of fear) and FPS (i.e., non-differential return of fear). Contrary to our hypothesis, fear responses were not attenuated in the CC group relative to the control groups.

Exploratory analyses. Following Experiment 1, regression analyses were used to explore whether post-intervention CS+ positive valence predicted less spontaneous recovery and less reinstatement measured with US_{neg}, FPS, or SCR, but it did not (largest Beta = −0.10, $p = .343$).

2.3. Discussion Experiment 2

Findings from Experiment 2 were generally in line with Experiment 1. Counterconditioning again reduced negative stimulus valence, but only compared to standard extinction training and not compared to extinction training with exposure to positive material. Again, counterconditioning did not reduce the return of fear. In contrast to Experiment 1, SCR did not decrease in the Intervention phase. It is unclear why SCR was low at the start of the intervention.

3. General discussion

We conducted two experiments to examine whether positive valence training through counterconditioning reduces negative stimulus valence and the return of fear. Overall, our findings do not support the notion that it reduces the return of fear. The first main finding is that counterconditioning reduced negative stimulus valence, compared to standard extinction training, which is in line with previous research (e.g., Engelhard et al., 2014; Kerkhof et al., 2011; Raes & De Raedt, 2012). However, some studies did not find this effect. This may be related to methodological differences, such as the use of different reinforcements (e.g., financial reward: Meulders, Karsdorp, Claes, & Vlaeyen, 2015; comic pictures: Kang et al., 2018), or testing an additional effect of counterconditioning on exposure in vivo, which already includes techniques that may reduce negative stimulus valence, such as therapeutic modelling (De Jong, Vorage, & Van Den Hout, 2000). We also found that negative stimulus valence was reduced directly after counterconditioning but spontaneously recovered after one week (Experiment 1). Thus, counterconditioning did not reduce negative stimulus valence more than extinction did in some studies, and its long-term effects were not found in Experiment 1. Together, these findings may indicate the relevance of testing boundary conditions of counterconditioning as positive valence training. In addition, in Experiment 1, counterconditioning reduced negative stimulus valence more compared to extinction training in which the same positive material was presented (but unrelated to the CS+). This suggests that the effects should not be attributed to general positive mood induction. Nevertheless, this effect was not replicated in Experiment 2, perhaps because negative affect before the intervention was higher in the one-day paradigm than in the three-day paradigm (see Tables 2 and 3).

The second main finding of the current experiments is that counterconditioning did not attenuate the return of fear relative to extinction training. To our knowledge, only one human fear conditioning study so far tested whether counterconditioning reduced the return of fear (Kang et al., 2018). Our findings are at odds with this study, in which counterconditioning did not reduce negative valence but did attenuate spontaneous recovery and reinstatement of threat expectancy (Kang et al., 2018). Two methodological differences may account for these divergent findings. First, we presented the US_{neg} *during* CS presentation throughout the acquisition phase (following Zbozinek, Holmes, et al., 2015) and the US_{pos} right *after* CS offset during the intervention phase, whereas Kang et al. (2018) presented the US_{neg} at CS *offset* in both phases. The same timing of the US presentation in each phase in Kang et al. (2018) may have enhanced learning during counterconditioning. Second, our US_{pos} comprised 6-s film clips which apparently enhanced the reduction of negative stimulus valence, whereas Kang et al. used 3-s comic pictures. In our experiments, counterconditioning may not have facilitated associative learning due the longer US duration or complexity.

Our findings are in line with rodent studies, in which no beneficial effect of counterconditioning (relative to extinction) on the return of fear has been detected (Bouton & Peck, 1992; Brooks, Hale, Nelson, & Bouton, 1995; Holmes, Leung, & Westbrook, 2016; Kerkhof et al., 2012). One explanation could be that, as with extinction, counterconditioning creates new, secondary learning about the CS-US relationship. That is, after counterconditioning, the CS possesses new meanings: its original meaning (CS-US_{neg}), as well as two additional meanings (CS-no US_{neg} and CS-US_{pos}). This may leave the original meaning intact and thereby vulnerable to the return of fear or relapse (e.g., Bouton, 2002; Craske, 2015). The current experiments suggest that advantages of counterconditioning with respect to reducing negative stimulus valence do not outweigh the disadvantage of creating ambiguity (i.e., two new associations) about fear-relevant stimuli. In this sense, counterconditioning may not be used as a method to prevent the return of fear.

However, this does not necessarily imply that counterconditioning

or related procedures are never effective in reducing a return of conditioned responding. First, because individuals tend to learn faster from negative than positive outcomes (Rozin & Royzman, 2001), counterconditioning with a negative (instead of positive) US may yield stronger learning effects and therefore be more effective in reducing a return of appetitive compared to fear responding. Two appetitive conditioning studies indeed found that, relative to extinction, counterconditioning with a negative US (a highly disliked liquid) reduced the return of appetitive responding (i.e., expectancy to eat chocolate and chocolate consumption; Van Gucht, Baeyens, Vansteenwegen, Hermans, & Beckers, 2010, 2013). Second, the "surprise" aspect of counterconditioning may be effective in itself, as it enhances prediction error. Two fear conditioning experiments showed that novelty-facilitated extinction (i.e., extinction in which a CS is followed by a surprising and novel nonthreat outcome) reduced the return of fear more than standard extinction training (Dunsmoor, Campese, Ceceli, LeDoux, & Phelps, 2015; Lucas, Luck, & Lipp, 2018, but see; Krypotos & Engelhard, 2018). Novelty-facilitated extinction is procedurally akin to counterconditioning (pairing a fear relevant stimulus with a neutral instead of positive US), but appears to be more effective in reducing return of fear.

Our third main finding is that reduced negative stimulus valence did not lower return of fear. Thus far, attempts to test this hypothesis have yielded inconclusive evidence. Two studies found that positive valence training reduced the spontaneous recovery of spider fear (Dour et al., 2016), and fear reinstatement in a conditioning paradigm (Zbozinek, Holmes, et al., 2015), but another study showed that positive valence training did not reduce reinstatement (Luck & Lipp, 2018). These findings, together with current findings, provide no compelling evidence for a direct causal relationship between post-extinction negative stimulus valence and the return of fear. How may these inconsistencies in evidence be reconciled? One explanation is that the positive correlation between negative stimulus valence and return of fear (e.g., Dirikx et al., 2007, 2004; Hermans et al., 2005; Huijding & De Jong, 2009; Vasey et al., 2012) is spurious. That is, there may be a third variable that explains the relation between post-extinction negative stimulus valence and return of fear, such as individual differences in positive affect. Indeed, positive valence training increased positive affect and reduced negative stimulus valence and reinstatement in a previous study (Zbozinek, Holmes, et al., 2015; but see; van Veen, Zbozinek, Engelhard, van Dis, & Craske, 2018).

Our findings may be relevant for the treatment of anxiety disorders in which negative stimulus valence may impede exposure therapy (e.g., due to disgust in spider phobia, Smits, Telch, & Randall, 2002; or blood-injection-injury phobia, Olatunji, Smits, Connolly, Willems, & Lohr, 2007). About 20% of anxiety patients drop out during treatment, and about 11% even refrain from starting with therapy (Fernandez, Salem, Swift, & Ramtahal, 2015). Counterconditioning may be a useful additional strategy to reduce negative stimulus valence to lower the threshold for individuals to expose themselves to fear-relevant situations. In that sense, it may increase the therapy's acceptability and reduce avoidance (e.g., Chen, & Bargh, 1999) and dropout. These are empirical questions that await future clinical research.

Strengths of the current research include the multimodal assessment of fear and a replication of results in an independent second experiment, which increases confidence in the robustness of our findings. Several limitations should also be noted. First, we only used an explicit measure of stimulus valence, which may be susceptible to demand characteristics and may reflect judgment-related processes instead of genuine changes in stimulus valence (Gawronski, Gast, & De Houwer, 2015). Future research could include implicit measures, such as affective priming (e.g., Engelhard et al., 2014; Raes & De Raedt, 2012). Second, participants were undergraduates. Future studies may preselect individuals suffering from anxiety or characterized by high neuroticism scores, as there might be more room for improvement in attenuating return of fear in these individuals (see Haaker, Golkar, Hermans, & Lonsdorf, 2014). Third, the effect of counterconditioning on negative stimulus valence was no longer present after one week. Further research with multiple day paradigms may use a more powerful manipulation of stimulus valence, for example by repeatedly exposing participants to the CS/US_{pos} relation throughout the week.

In conclusion, counterconditioning seems to be promising as positive valence training. Our findings do not support the notion that it reduces the return of fear and there was no direct relationship between post-intervention negative stimulus valence and the return of fear. More research is needed to test boundary conditions of counterconditioning effects on stimulus valence and how (sub)clinical groups may profit from counterconditioning.

Funding

CB is now at Amsterdam AMC, location AMC, University of Amsterdam and at the Institute for Advanced Study. This work was supported by a Vici grant (453-15-005) from the Netherlands Organisation for Scientific Research awarded to IME. EvD, MH, and IME designed the studies. EvD performed statistical analyses and drafted the manuscript. MH, CB, and IME critically revised the manuscript.

Acknowledgements

We thank Ashley Ferguson, Mijntje Boon, Julia van Buuren, Sylvia Veerman and Herma Walrave for testing participants.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2019.103416.

References

Benning, S. D., Patrick, C. J., & Lang, A. R. (2004). Emotional modulation of the post-auricular reflex. *Psychophysiology, 41*(3), 426–432. https://doi.org/10.1111/j.1469-8986.00160.x.

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology, 42*(1), 1–15. https://doi.org/10.1111/j.1469-8986.2005.00271.x.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*(10), 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9.

Bouton, M. E., & Peck, C. A. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning & Behavior, 20*(4), 313–321. https://doi.org/10.3758/BF03197954.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry, 25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Brooks, D. C., Hale, B., Nelson, J. B., & Bouton, M. E. (1995). Reinstatement after counterconditioning. *Animal Learning & Behavior, 23*(4), 383–390. https://doi.org/10.3758/BF03198938.

van Buuren, S. (2012). *Flexible imputation of missing data.* London: Chapman and Hall.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Pers. Soc. Psychol. Bull. 25*(2), 215–224. https://doi.org/10.1177/0146167299025002007.

Craske, M. (2015). Optimizing exposure therapy for anxiety disorders: An inhibitory learning and inhibitory regulation approach. *Verhaltenstherapie, 25*(2), 134–143. https://doi.org/10.1159/000381574.

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. https://doi.org/10.1016/j.brat.2014.04.006.

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry, 15*(3), 245–258. https://doi.org/10.1002/wps.20346.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853–869. https://doi.org/10.1037//D033-29O9.127.6.853.

De Jong, P. J., Vorage, I., & Van Den Hout, M. A. (2000). Counterconditioning in the treatment of spider phobia: Effects on disgust, fear and valence. *Behaviour Research and Therapy, 38*(11), 1055–1069. https://doi.org/10.1016/S0005-7967(99)00135-7.

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2004).

Reinstatement of extinguished conditioned responses and negative stimulus valence as a pathway to return of fear in humans. *Learning & Memory, 11*(5), 549–554. https://doi.org/10.1101/lm.78004.

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2007). Reinstatement of conditioned responses in human differential fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(3), 237–251. https://doi.org/10.1016/j.jbtep.2006.04.001.

Dour, H. J., Brown, L. A., & Craske, M. G. (2016). Positive valence reduces susceptibility to return of fear and enhances approach behavior. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 277–282. https://doi.org/10.1016/j.jbtep.2015.09.010.

Dunsmoor, J. E., Campese, V. D., Ceceli, A. O., LeDoux, J. E., & Phelps, E. A. (2015). Outcome in place of an expected threat diminishes recovery of defensive responses. *Biological Psychiatry, 78*(3), 203–209. https://doi.org/10.1016/j.biopsych.2014.12.008.

Engelhard, I. M., Leer, A., Lange, E., & Olatunji, B. O. (2014). Shaking that icky feeling: Effects of extinction and counterconditioning on disgust-related evaluative learning. *Behavior Therapy, 45*(5), 708–719. https://doi.org/10.1016/j.beth.2014.04.003.

Fernandez, E., Salem, D., Swift, J. K., & Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology, 83*(6), 1108–1122. https://doi.org/10.1037/ccp0000044.

Gawronski, B., Gast, A., & De Houwer, J. (2015). Is evaluative conditioning really resistant to extinction? Evidence for changes in evaluative judgements without changes in evaluative representations. *Cognition & Emotion, 29*(5), 816–830. https://doi.org/10.1080/02699931.2014.947919.

Haaker, J., Golkar, A., Hermans, D., & Lonsdorf, T. B. (2014). A review on human reinstatement studies: An overview and methodological challenges. *Learning & Memory, 21*(9), 424–440. https://doi.org/10.1101/lm.036053.114.

Hermans, D., Dirikx, T., Vansteenwegenin, D., Baeyens, F., Van Den Bergh, O., & Eelen, P. (2005). Reinstatement of fear responses in human aversive conditioning. *Behaviour Research and Therapy, 43*(4), 533–551. https://doi.org/10.1016/j.brat.2004.03.013.

Hofmann, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *Journal of Clinical Psychiatry, 69*(4), 621–632. https://doi.org/10.4088/JCP.v69n0415.

Holmes, N. M., Leung, H. T., & Westbrook, R. F. (2016). Counter conditioned fear responses exhibit greater renewal than extinguished fear responses. *Learning & Memory, 23*(4), 141–150. https://doi.org/10.1101/lm.040659.115.

Huijding, J., & De Jong, P. J. (2009). Implicit and explicit attitudes toward spiders: Sensitivity to treatment and predictive value for generalization of treatment effects. *Cognitive Therapy and Research, 33*(2), 211–220. https://doi.org/10.1007/s10608-007-9167-5.

Jess0rT (2011). *Hysterical bubbles! (original) - laughing baby*. Retrieved from https://www.youtube.com/user/Jess0rT.

Kang, S., Vervliet, B., Engelhard, I. M., van Dis, E. A. M., & Hagenaars, M. A. (2018). Reduced return of threat expectancy after counterconditioning verus extinction. *Behaviour Research and Therapy, 108*, 78–84. https://doi.org/10.1016/j.brat.2018.06.009.

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011). Counterconditioning: An effective technique for changing conditioned preferences. *Experimental Psychology, 58*(1), 31–38. https://doi.org/10.1027/1618-3169/a000063.

Kerkhof, I., Vansteenwegen, D., Beckers, T., Dirikx, T., Bayens, F., D'Hooge, R., et al. (2012). The role of negative affective valence in return of fear. In A. D. Gervaise (Ed.). *Psychology of fear: New research* (pp. 153–170). New York: Nova Science Publishers.

Krypotos, A. M., & Engelhard, I. M. (2018). Testing a novelty-based extinction procedure for the reduction of conditioned avoidance. *Journal of Behavior Therapy and Experimental Psychiatry, 60*, 22–28. https://doi.org/10.1016/j.jbtep.2018.02.006.

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., et al. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews, 77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026.

Lucas, K., Luck, C. C., & Lipp, O. V. (2018). Novelty-facilitated extinction and the reinstatement of conditional human fear. *Behaviour Research and Therapy, 109*, 68–74. https://doi.org/10.1016/j.brat.2018.08.002.

Luck, C. C., & Lipp, O. V. (2018). Verbal instructions targeting valence alter negative conditional stimulus evaluations (but do not affect reinstatement rates). *Cognition & Emotion, 32*(1), 61–80. https://doi.org/10.1080/02699931.2017.1280449.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5.

Meulders, A., Karsdorp, P. A., Claes, N., & Vlaeyen, J. W. S. (2015). Comparing counterconditioning and extinction as methods to reduce fear of movement-related pain. *The Journal of Pain, 16*(12), 1353–1365. https://doi.org/10.1016/j.jpain.2015.09.007.

Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: A review of meta-analytic findings. *Psychiatric Clinics of North America, 33*(3), 557–577. https://doi.org/10.1016/j.psc.2010.04.002.

Olatunji, B. O., Smits, J. A. J., Connolly, K., Willems, J., & Lohr, J. M. (2007). Examination of the decline in fear and disgust during exposure to threat-relevant stimuli in blood-injection-injury phobia. *Journal of Anxiety Disorders, 21*(3), 445–455. https://doi.org/10.1016/j.janxdis.2006.05.001.

Orr, S. P., Metzger, L. J., Lasko, N. B., Macklin, M. L., Peri, T., & Pitman, R. K. (2000). De novo conditioning in trauma-exposed individuals with and without posttraumatic stress disorder. *Journal of Abnormal Psychology, 109*(2), 290–298. https://doi.org/10.1037/0021-843X.109.2.290.

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*(5), 984–995. https://doi.org/10.1111/j.1469-8986.2009.00852.x.

Raes, A. K., & De Raedt, R. (2012). The effect of counterconditioning on evaluative responses and harm expectancy in a fear conditioning paradigm. *Behavior Therapy, 43*(4), 757–767. https://doi.org/10.1016/j.beth.2012.03.012.

Rottenberg, J., Ray, R. D., & Gross, J. J. (2007). Emotion elicitation using films. In J. A. Coan, & J. J. B. Allen (Eds.). *Handbook of emotion elicitation and assessment* (pp. 9–28). New York: Oxford University Press.

Rozin, P., & Royzman, E. B. (2001). *Negativity Bias , Negativity Dominance , and Contagion, 5*(4), 296–320. https://doi.org/10.1207/s15327957pspr0504_2.

Sanderman, R., Arrindell, W., & Ranchor, A. (1991). Crosscultural comparison of personality traits: The Netherlands and england. *Psychological Reports, 69*, 1091–1096. https://doi.org/10.2466/pr0.1991.69.3f.1091.

Sanderman, R., Arrindell, W. A., Ranchor, A. V., Eysenck, H. J., & Eysenck, S. B. G. (2012). *Het meten van persoonlijkheidskenmerken met de Eysenck Personality Questionnaire (EPQ), een handleiding (Tweede her)*. Groningen: Research Insitute SHARE.

Smits, J. A. J., Telch, M. J., & Randall, P. K. (2002). An examination of the decline in fear and disgust during exposure-based treatment. *Behaviour Research and Therapy, 40*(11), 1243–1253. https://doi.org/10.1016/S0005-7967(01)00094-8.

Van Gucht, D., Baeyens, F., Hermans, D., & Beckers, T. (2013). The inertia of conditioned craving. Does context modulate the effect of counterconditioning? *Appetite, 65*, 51–57. https://doi.org/10.1016/j.appet.2013.01.019.

Van Gucht, D., Baeyens, F., Vansteenwegen, D., Hermans, D., & Beckers, T. (2010). Counterconditioning reduces cue-induced craving and actual cue-elicited consumption. *Emotion, 10*(5), 688–695. https://doi.org/10.1037/a0019463.

Vasey, M. W., Harbaugh, C. N., Buffington, A. G., Jones, C. R., & Fazio, R. H. (2012). Predicting return of fear following exposure therapy with an implicit measure of attitudes. *Behaviour Research and Therapy, 50*(12), 767–774. https://doi.org/10.1016/j.brat.2012.08.007.

van Veen, S. C., Zbozinek, T. D., Engelhard, I. M., van Dis, E. A. M., & Craske, M. G. (2018). *Positive mood induction does not reduce return of fear: A virtual reality exposure study for public speaking anxiety.* (in preparation).

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology, 9*(1), 215–248. https://doi.org/10.1146/annurev-clinpsy-050212-185542.

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., et al. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the Global Burden of Disease Study 2010. *The Lancet, 382*(9904), 1575–1586. https://doi.org/10.1016/S0140-6736(13)61611-6.

Zbozinek, T. D., & Craske, M. G. (2017). The role of positive affect in enhancing extinction learning and exposure therapy for anxiety disorders. *Journal of Experimental Psychopathology, 8*(1), 13–39. https://doi.org/10.5127/jep.052615.

Zbozinek, T. D., Hermans, D., Prenoveau, J. M., Liao, B., & Craske, M. G. (2015). Post-extinction conditional stimulus valence predicts reinstatement fear: Relevance for long-term outcomes of exposure therapy. *Cognition & Emotion, 29*(4), 654–667. https://doi.org/10.1080/02699931.2014.930421.

Zbozinek, T. D., Holmes, E. A., & Craske, M. G. (2015). The effect of positive mood induction on reducing reinstatement fear: Relevance for long term outcomes of exposure therapy. *Behaviour Research and Therapy, 71*, 65–75. https://doi.org/10.1016/j.brat.2015.05.016.