Verbal instructions targeting valence alter negative conditional stimulus evaluations (but do not affect reinstatement rates)

Camilla C. Luck[1, 2] & Ottmar V. Lipp[1, 2]

[1]School of Psychology and Speech Pathology, Curtin University, Australia

[2]ARC-SRI: Science of Learning Research Centre

Author Notes

Address for correspondence:

Camilla C. Luck, School of Psychology and Speech Pathology, Curtin University, GPO Box U1987 Perth WA 6845, Australia. Email: c.luck@curtin.edu.au

Acknowledgements

This work was supported by grants number DP120100750 and SR120300015 from the Australian Research Council.

Word count: 11897

Abstract

Negative conditional stimulus (CS) valence acquired during fear conditioning may enhance fear relapse and is difficult to remove as it extinguishes slowly and does not respond to the instruction that unconditional stimulus (US) presentations will cease. We examined whether instructions targeting CS valence would be more effective. In Experiment 1, an image of one person (CS+) was paired with an aversive US, while another (CS-) was presented alone. After acquisition, participants were given positive information about the CS+ poser and negative information about the CS- poser. Instructions reversed the pattern of differential CS valence present during acquisition and eliminated differential electrodermal responding. In Experiment 2, we compared positive and negative CS revaluation by providing positive/negative information about the CS+ and neutral information about CS-. After positive revaluation, differential valence was removed and differential electrodermal responding remained intact. After negative revaluation, differential valence was strengthened and differential electrodermal responding was eliminated. Unexpectedly, the instructions did not affect the reinstatement of differential electrodermal responding.

Key words: Fear conditioning; reinstatement; cognitive intervention; conditional stimulus valence; electrodermal responding.

Efficacious treatments for anxiety disorders are available but relapse is a widespread and pervasive problem – approximately one to two thirds of recovered patients will relapse within eight years (Craske, 1999). Consequently, the research focus has shifted from the assessment of short term treatment effects, to the long term success of efficacious treatments. In laboratory based research, the acquisition, reduction, and relapse of human fear can be modelled using classical fear conditioning – an approach which gives researchers the ability to manipulate variables of potential clinical importance (Craske, Hermans, & Vansteenwegen, 2006).

During differential fear conditioning, two neutral conditional stimuli (CS) and an aversive unconditional stimulus (US) are presented throughout the experiment. Acquisition, a laboratory analogue for the development of a fear, involves pairing one CS, the CS+, with the US, while the other, the CS-, is presented alone (Lipp, 2006). The CS+ becomes a predictor of the US and acquires negative valence (becomes unpleasant) leading to the development of differential physiological responding and differential valence evaluations between CS+ and CS- (De Houwer, Thomas, & Baeyens, 2001; Lipp, 2006). After acquisition, this differential responding can be reduced in an extinction or a counterconditioning procedure. Extinction, a laboratory analogue for fear reduction via exposure therapy, involves presenting both CS+ and CS- without the US. Throughout extinction, differential physiological responding and valence evaluations between CS+ and CS- reduce (Lipp, 2006). With sufficient exposure both will return to pre-acquisition levels, however, negative valence extinguishes slower than physiological responding (Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010) and often the CS+ retains some negative valence at the end of extinction. During counterconditioning, instead of the aversive US, the CS+ is followed by an appetitive stimulus. Participants learn that the aversive US will not be presented anymore and due to pairings with the appetitive stimulus the

CS+ gradually becomes more pleasant. Counterconditioning can be more effective at removing the negative valence acquired during fear acquisition – at least in the short term.

The fear associations acquired during acquisition are not erased by extinction or counterconditioning training and differential responding can return without re-training or re-exposure (for a review see Vervliet, Craske, & Hermans, 2013; Van Gucht, Baeyens, Hermans, & Beckers, 2013). This phenomenon, known as the return of fear, can occur due to the mere passage of time (spontaneous recovery); a change in context (renewal); or the unsignaled presentation of the US (reinstatement; Bouton, 2002). Higher levels of negative CS+ valence have been shown to correlate with higher reinstatement rates (Dirkx, Hermans, Vansteenwegen & Baeyens, 2004; Hermans et al., 2005; Zbozinek, Hermans, Prenoveau, Liao & Craske, 2015) and negative CS+ valence seems to be slow to extinguish and resists instructed extinction (Luck & Lipp, 2015a).

Instructed extinction, an experimental manipulation which targets the CS-US relationship, involves informing participants before extinction that US presentations will cease (for a review see Luck & Lipp, 2016a). Luck and Lipp (2015a) report that instructed extinction eliminates conditional electrodermal responding and fear potentiated startle but leaves conditional valence evaluations intact; a pattern of results that is observed regardless of whether the US electrode is removed or remains attached (Luck & Lipp, 2015b). These findings may indicate that the CS-US relationship is not the appropriate target for an instructional intervention intended to reduce negative CS+ valence, but an instructional intervention targeting CS valence may be better suited to reduce the negative CS+ valence acquired during acquisition. A CS targeted intervention may also be more likely to change the valence of the CS intrinsically (the CS+ is pleasant) rather than creating an additional CS-US association (the CS+ predicts pleasant stimuli), which would later compete with the CS-US association learned during acquisition and potentially result in a relapse of negative valence. If the value of the CS+ is changed intrinsically, the valence change may also be less susceptible to relapse.

In the current study, we examined whether an instructional intervention targeting CS valence would affect conditional valence evaluations acquired during classical fear conditioning. During acquisition, an image of one person (CS+) was followed by an aversive electrotactile stimulus, while another (CS-) was presented alone. At the end of acquisition, participants received information about the individuals, but not the CS-US contingency. The revaluation group received positive character information about the CS+ poser and negative character information about the CS- poser, while the control group received neutral information about both posers. Relative to controls, we hypothesized that targeted instructions would increase CS+ valence, such that the CS+ would become more pleasant, and reduce CS- valence, such that the CS- would become less pleasant. In the control group, we expected no effect of instructions, but that differential valence would extinguish throughout extinction and recover after a reinstatement procedure (three unsignalled US presentations). In the revaluation group, we expected the valence changes produced by the revaluation instructions to persist throughout extinction and not to be affected by the reinstatement procedure.

Electrodermal responding was measured to examine whether CS revaluation instructions would also influence physiological responding. Electrodermal responding is a very useful and common measure of fear learning, but is not selectively sensitive to fear, showing the same response patterns in non-aversive conditioning as well (Lipp, Siddle, & Dall, 2003). Electrodermal responding can reflect a number of psychological processes, including the anticipation of a salient event, emotional arousal, and orienting (Lipp, 2006). During fear conditioning, these processes tend to co-occur – the CS+ predicts the US, acquires negative emotional arousal, and elicits enhanced orienting responses, relative to CS-. Some experimental manipulations, however, may differentially affect these processes, which renders the interpretation of electrodermal responding difficult. Instructions targeting CS valence should not affect expectations of the CS-US contingency, but may reverse emotional arousal and increase orienting to both the CS+ and the CS- (although not necessarily equally). In such circumstances, electrodermal responding is likely to reflect the most salient process. Extrapolating from instructed extinction studies, in which electrodermal responding to CS+ reduces after participants are informed that the US will no longer be presented, we predict that the unchanged CS-US contingency will be the most salient process. Therefore, we hypothesized that differential conditional electrodermal responding would not be affected by the instructions, but would reduce throughout the extinction phase. In line with the research suggesting that negative CS+ valence is correlated with higher reinstatement rates (Dirkx et al., 2004; Hermans et al., 2005; Zbozinek et al., 2015), we predicted that after the reinstatement procedure, the extent of differential responding would be smaller in the revaluation group in comparison with the control group. An affective priming task was included to provide an implicit measure of CS valence and offline CS valence evaluations were measured before habituation, after reinstatement, after extinction, and post-experimentally, to complement the online CS evaluations and to test for possible renewal effects (Bouton, 2002; see Lipp & Purkis, 2006 for an examination of assessment timing on CS valence).

Experiment 1

Method

Participants

Thirty two undergraduate students (21 female), aged between 18 – 44 years (*M* = 23.91) volunteered participation in exchange for course credit. Sample size was determined based on previous experience examining instruction effects in fear conditioning. Participants provided informed consent and were randomly assigned to the control or revaluation group. One participant's electrodermal responses were lost due to problems with the recording device.

Apparatus/Stimuli

The conditional stimuli (CS) were 506 × 650 pixel color pictures of Caucasian, male adults [NimStim database: images M_NE_C: models 20, 21, 32, 31, Tottenham et al. (2009)] displaying neutral facial expressions. The pictures were presented on a 24 inch color LCD screen for 6 s and the trial sequence was arranged in a pseudo-random order, such that a CS+/CS- was not presented more than twice consecutively. Counter-balancing was performed across participants, varying the nature of the first trial (CS+/CS-), the face used as CS+/CS-, and the two faces used in the experiment (out of the possible four). The unconditional stimulus (US) was a 200 ms electrotactile stimulus, pulsed at 50 Hz and delivered by a Grass SD9 Stimulator to the participant's preferred forearm.

CS evaluations, electrodermal responding, and respiration were recorded with a Biopac MP150 system, using acqKnowledge Version 4.1 at a sampling frequency of 1000 Hz. DMDX 5.1.3.4 software (Forster & Forster, 2003) was used to control the stimulus presentation and timing, deliver the instructional manipulations, record the before and after pleasantness ratings, and record the reaction times and errors from the affective priming task. Online CS valence was measured with a Biopac Variable Assessment Transducer with the anchors 0 (*very negative*) to 9 (*very positive*). Electrodermal responding was recorded with two 8 mm Ag/AgCl electrodes filled with an isotonic gel and DC amplified at a gain of 5 μSiemens per Volt. Respiration was measured with a respiratory effort transducer fitted with an adjustable Velcro strap to control for changes in electrodermal responding that occurred due to excessive movement or deep breaths. No such cases were identified and therefore no electrodermal responses were excluded.

Procedure

Participants washed their hands, provided informed consent, and were seated in front of a monitor in a separate room adjacent to the control room. The respiratory effort transducer was fitted around their waist, the electrodermal electrodes were attached to the thenar and hypothenar prominences of their non-preferred hand, and the shock electrode was placed on their preferred forearm. Participants underwent a shock work-up procedure to set the intensity of the electrotactile stimulus to a level that was subjectively experienced as 'unpleasant but not painful' and then asked to relax and watch the blank computer screen while a 3 min baseline of their electrodermal activity was recorded. After the baseline recording, participants rated the CS faces on a 1 (*unpleasant*) to 9 (*pleasant*) Likert scale with the keyboard (Ratings A) and completed an affective priming task, using the four possible CS faces as primes. The face primes were presented for 200 ms followed by a blank screen for 100 ms and then a pleasant, or an unpleasant target word for 1 s, or until a response was made. CS–word pairs were presented in a random order and each pair was presented twice, forming 96 trials in total. After the affective priming task, participants were informed that faces would be displayed on the screen and asked to pay attention to which face was followed by the electrotactile stimulus. They were asked to use the Biopac Assessment Transducer with their preferred hand to indicate how pleasant/unpleasant they found each face while it was on the screen, ensuring that the movement did not interfere with the electrodermal recordings and the presence/absence of the US did not influence the evaluations. The participants were asked to move the dial back to neutral after each trial.

After these instructions, the conditioning task, consisting of habituation, acquisition, extinction, and the reinstatement test phase, was started. During habituation, both CS+ and CS- were presented alone four times. During acquisition, CS+ was presented 8 times, with the offset of the CS+ coinciding with the onset of the US in a 100% reinforcement schedule, while CS- was presented eight times alone. After acquisition, the experimenter informed the participants that there would be a break in the experiment for them to receive information about the individuals shown in the pictures. In the revaluation group, two positive and negative instructions were counter-balanced across participants. The positive instructions were: 'This is Ben (CS+ face), Ben is passionate about helping people in need and has been raising money for the local homeless shelter. Ben is currently preparing to travel to Africa to help teach the children in the local schools'; and 'This is Chris (CS+ face), Chris is a very community oriented citizen who regularly volunteers at the local children's hospital. Chris is currently preparing to travel to Africa to help build a new school for the local children.' The negative instructions were: This is Chris (CS- face), Chris is currently in jail for setting fire to his elderly neighbor's house after they had an argument. Chris has been caught fighting with the other inmates and has attempted to escape a number of times'; and 'This is Ben (CS- face), Ben has a history of domestic violence and is currently in jail for beating his pregnant girlfriend. Ben was moved to a higher security prison, after he managed to send threatening letters to his girlfriend'. Participants in the control group were given neutral information about both posers. Two control instructions were used and which instruction was given for CS+ and CS- was counter-balanced across participants. The control instructions were: 'This is Chris, Chris works as a builder and is about to start work on a new housing estate downtown. When Chris finishes work it takes him about 30 minutes to drive home'; and 'This is Ben, Ben works at the local supermarket and spends most of his time working on the tills and serving customers. When Ben finishes work, it takes him about 20 minutes to drive home'.

After the manipulation, participants were informed that the experiment would continue and extinction, consisting of eight unreinforced presentations of CS+ and CS-, was started. Participants were not informed about the CS–US contingency and the shock electrode was left attached. Following extinction, participants completed a second ratings task (Ratings B), and a second affective priming task, both identical to the first tasks. After the last affective priming trial, three unsignaled US presentations were administered, followed by four unreinforced presentations of CS+ and CS- (test phase). Participants then completed a third rating task (Ratings C) and the electrodes were removed. In the control room a post-experimental questionnaire was administered – the questionnaire required participants to identify which faces were presented during acquisition training, which face was paired with the US, and which information they had received for each face. It also included another rating task using a different scale (-3 [*very unpleasant*] to +3 [*very pleasant*]) and an evaluation of how pleasant/unpleasant they found the electrotactile stimulus (-3 [*very unpleasant*] to +3 [*very pleasant*]).

Scoring and Response Definition

The online valence ratings were recorded as voltage deviations and scored as the largest positive or negative deviation recorded during the 6 s CS presentation from a 1 s pre-CS baseline voltage ('neutral' position). Electrodermal responding was scored in multiple latency windows as recommended by Prokasy and Kumpfer (1973) and Luck and Lipp (2016b). First interval responding (FIR) was defined as responses starting within 1-4 s of CS onset, second interval responding (SIR) was defined as responses starting within 4-7 s of CS onset. Responses to the US were scored during acquisition as responses starting within 7-10 s of CS+ onset (1-4 s of US onset). Both first and second interval responding are sensitive to fear learning and will show both the acquisition and extinction of conditional fear, FIR however, is more sensitive to the orienting elicited by the CS onset, while, SIR is more sensitive to the anticipation of the US (Öhman, 1973). The largest response starting within the latency window was scored and response magnitude was calculated as the difference between response onset and response peak (Prokasy & Kumpfer, 1973). First and second interval responses were square root transformed to reduce the positive skew of the distribution (Dawson et al., 2007), and then range corrected to reduce the effect of individual differences in response size (Boucsein et al., 2012; Dawson, Schell, & Filion, 2007). The reference for the range correction was the largest response displayed by the participant, which was typically the response to the first or second presentation of the US. During habituation, only FIRs were scored as they are more sensitive to orienting processes and anticipatory processes are not expected during habituation (Öhman, 1973). As a measure of spontaneous electrodermal responding, any discernible electrodermal response displayed during the baseline period was counted (Dawson et al., 2007).

Scoring and Response Definition

To reduce the influence of trial by trial variability, FIR, SIR, and CS valence evaluations were averaged into blocks of two consecutive trials. For comparison with Ratings A, B, and C, the post-experimental rating scores (Ratings D) were transformed to a 1-9 pleasantness scale. All analyses were conducted with IBM SPSS Statistics 22 with an alpha cut-off of .05, interactions have been followed-up with simple effect contrasts, and Pillai’s trace statistics of the multivariate solution are reported. Bonferroni corrections were applied to the follow-up analyses by dividing the critical alpha value (.05) by the number of follow-up comparisons. In cases where the Bonferroni correction changed the overall conclusion (significant or non-significant), the corrected *p* value (Wright, 1992; uncorrected p × number of comparisons) has also been reported.

Results

Preliminary Checks

The descriptive statistics for the preliminary checks are presented in Table 1. A Pearson's chi-square test revealed that the female to male ratio was marginally higher in the control than in the revaluation group, $\chi^2(1) = 3.46$, $p = .063$. A series of independent samples t-tests revealed that the groups did not differ in age, spontaneous electrodermal responding, US intensity, or US valence, all *t*s < .985, *p*s > .332. A 2 (Group: instruction, control) × 4 (Block: 1, 2, 3, 4) factorial ANOVA confirmed that unconditional electrodermal responding during acquisition did not differ between groups, all $F < 0.83$, *p*s > .373, $\eta p^2 < .081$. One participant from the revaluation group was not able to identify which face was used as the CS-, and one participant from the revaluation group identified the incorrect instruction story (but of the same valence). When these participants were removed from the analyses the conclusions remained the same and therefore results from the full sample are reported.

Habituation

The CS valence evaluations and first interval responding (see the first panels of Figures 1 and 2, respectively) recorded during habituation were subjected to separate 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 2 (Block: 1, 2) mixed-model factorial ANOVAs.

Online CS Evaluations. No significant effects were detected, all $F < 2.76$, $p > .107$, $\eta p^2 < .085$.

First Interval Responding. Responding decreased from block one to block two, $F(1, 29) = 16.26$, $p < .001$, $\eta p^2 = .359$.

Acquisition

The CS valence evaluations, and electrodermal first and second interval responding recorded during acquisition were subjected to separate 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 4 (Block: 1, 2, 3, 4) mixed model factorial ANOVAs and are presented in Figures 1 (second panel), 2 (second panel), and 3 (first panel), respectively.

Online CS Evaluations. A main effect of CS, $F(1, 30) = 5.63$, $p = .024$, $\eta p^2 = .158$, was moderated by a CS × Block interaction, $F(3, 28) = 4.30$, $p = .013$, $\eta p^2 = .316$. Evaluations of CS+ and CS- did not differ during blocks one or two, $F(1, 30) < 2.77$, $p > .106$, $\eta p^2 < .085$, but CS+ was evaluated as less pleasant than CS- during blocks three and four, $F(1, 30) > 7.68$, $p < .010$, $\eta p^2 > .203$.

First Interval Responding. Responses to CS+ were larger than to CS-, $F(1, 29) = 39.82$, $p < .001$, $\eta p^2 = .579$, and responding declined across blocks, $F(3, 27) = 9.53$, $p < .001$, $\eta p^2 = .514$.

Second Interval Responding. A main effect of CS, $F(1, 29) = 43.33$, $p < .001$, $\eta p^2 = .599$, and a main effect of block, $F(3, 27) = 5.88$, $p = .003$, $\eta p^2 = .395$, were qualified by a CS × Block interaction, $F(3, 27) = 7.31$, $p = .001$, $\eta p^2 = .448$. Responding during block one did not differ between CS+ and CS-, $F(1, 29) = 2.41$, $p = .131$, $\eta p^2 = .077$, but during subsequent blocks responding to CS+ was larger than responding to CS-, all $F(1, 29) > 25.75$, $p < .001$, $\eta p^2 > .469$.

Instructional Manipulation

To assess the immediate influence of the instructions, separate 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 2 (Phase: last block of acquisition, first block of extinction) mixed model factorial ANOVAs were conducted on the CS valence evaluations, and first and second interval responding.

Online CS Evaluations. A CS × Phase interaction, $F(1, 30) = 40.46$, $p < .001$, $\eta p^2 = .574$, and a CS × Group interaction, $F(1, 30) = 7.11$, $p = .012$, $\eta p^2 = .191$, were moderated by a CS × Phase × Group interaction, $F(1, 30) = 16.83$, $p < .001$, $\eta p^2 = .359$. During the last block of acquisition, participants in the control group evaluated CS+ as less pleasant than CS-, $F(1, 30) = 11.24$, $p = .002$, $\eta p^2 = .273$. In the revaluation group, the pattern was similar, but not significant after Bonferroni correction, $F(1, 30) = 3.76$, $p = .062$, $\eta p^2 = .111$ ($p_{adjusted} = .248$). During the first block of extinction, the control group evaluated CS- as marginally more pleasant than CS+, $F(1, 30) = 3.69$, $p = .064$, $\eta p^2 = .109$ (not significant after Bonferroni correction, $p_{adjusted} = .256$), but the revaluation group evaluated CS+ as more pleasant than CS-, $F(1, 30) = 12.16$, $p = .002$, $\eta p^2 = .288$. Valence evaluations of CS+ or CS- did not differ between groups during the last block of acquisition, $F(1, 30) < 1.41$, $p > .245$, $\eta p^2 < .046$, but during the first block of extinction, evaluations of CS- were more unpleasant in the revaluation group, $F(1, 30) = 27.87$, $p < .001$, $\eta p^2 = .482$, while valence evaluations to CS+ did not differ between the groups, $F(1, 30) = 2.36$, $p = .135$, $\eta p^2 = .073$.

First Interval Responding. A main effect of CS, $F(1, 29) = 20.26$, $p < .001$, $\eta p^2 = .411$, and a CS × Phase interaction, $F(1, 29) = 5.57$, $p = .025$, $\eta p^2 = .161$, were moderated by a CS × Phase × Group interaction, $F(1, 29) = 5.51$, $p = .026$, $\eta p^2 = .160$. During the last block of acquisition responding to CS+ was greater than responding to CS- in both the control group, $F(1, 29) = 7.69$, $p = .010$, $\eta p^2 = .210$, and the revaluation group, $F(1, 29) = 17.57$, $p < .001$, $\eta p^2 = .377$. This pattern persisted in the control group during the first block of extinction, $F(1, 29) = 8.05$, $p = .008$, $\eta p^2 = .217$, however, differential responding to CS+ and CS- had been eliminated by the first block of extinction in the revaluation group, $F(1, 29) = 0.07$, $p = .793$, $\eta p^2 = .002$. A comparison between the groups revealed no difference in responding to the CS+ or CS- during the last block of acquisition or the first block of extinction, $F(1, 29) < 1.79$, $p > .192$, $\eta p^2 < .059$.

Second Interval Responding. A main effect of CS, $F(1, 29) = 23.61$, $p < .001$, $\eta p^2 = .449$, and a main effect of phase, $F(1, 29) = 16.97$, $p < .001$, $\eta p^2 = .369$, were moderated by a CS × Phase interaction, $F(1, 29) = 15.03$, $p = .001$, $\eta p^2 = .341$. During the last block of acquisition, responses to CS+ were larger than to CS-, $F(1, 29) = 29.71$, $p < .001$, $\eta p^2 = .506$, but during the first block of extinction, responding to CS+ and CS- did not differ, $F(1, 29) = 2.88$, $p = .100$, $\eta p^2 = .090$.

Extinction

The CS valence evaluations, and first and second interval responding recorded during extinction were subjected to separate 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 4 (Block: 1, 2, 3, 4) mixed model factorial ANOVAs and are presented in the right panels of Figures 1, 2, and 3, respectively.

Online CS Evaluations. Evaluations in block two were less pleasant than evaluations in blocks three, $p = .006$, and four, $p = .016$ (marginal after Bonferroni correction, $p_{adjusted} = .096$), $F(3, 28) = 3.12$, $p = .042$, $\eta p^2 = .250$. A main effect of group, $F(1, 30) = 4.89$, $p = .035$, $\eta p^2 = .140$, was moderated by a CS × Group interaction, $F(1, 30) = 17.54$, $p < .001$, $\eta p^2 = .369$. Differential evaluations were not present in the control group, $F(1, 30) = 2.53$, $p = .122$, $\eta p^2 = .078$, but the revaluation group evaluated CS- as less pleasant than CS+, $F(1, 30) = 18.76$, $p < .001$, $\eta p^2 = .385$. Evaluations of CS- were more unpleasant in the revaluation group than in the control group, $F(1, 30) = 26.52$, $p < .001$, $\eta p^2 = .469$, while evaluations of CS+ were marginally more pleasant in the revaluation group than in the control group, $F(1, 30) = 4.14$, $p = .051$, $\eta p^2 = .121$, (but not significant after Bonferroni correction, $p_{adjusted} = .102$).

First Interval Responding. Responding in block one was larger than responding in blocks two, p = .012, three, p = .001, and four, p < .001, F(3, 27) = 7.44, p = .001, $\eta p^2 = .453$, Responses to CS+ were marginally larger than to CS-, F(1, 29) = 3.52, p = .071, $\eta p^2 = .108$.

Second Interval Responding. A CS × Group interaction, F(1, 29) = 4.78, p = .037, $\eta p^2 = .142$, revealed that differential responding was not present in either group, F(1, 29) < 2.79, p > .105, $\eta p^2 < .089$. Responding to CS- in the control group was larger than in the revaluation group, F(1, 29) = 4.31, p = .047, $\eta p^2 = .129$ (not significant after Bonferroni correction, $p_{adjusted}$ = .094), but responding to CS+ did not differ between the groups, F(1, 29) = 0.16, p = .695, $\eta p^2 = .005$.

Reinstatement Test

To examine the influence of the reinstatement procedure on CS valence evaluations, and electrodermal first, and second interval responding, separate 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 2 (Phase: last block of extinction, first block of test phase) mixed model factorial ANOVAs were conducted.

Online CS Evaluations. Evaluations were more negative in the first block of the test phase than in the last block of extinction, F(1, 30) = 8.60, p = .006, $\eta p^2 = .223$. A CS × Group interaction, F(1, 30) = 15.54, p < .001, $\eta p^2 = .341$, revealed that the control group did not differentially evaluate CS+ and CS-, F(1, 30) = 2.21, p = .148, $\eta p^2 = .069$, but the revaluation group evaluated CS- as less pleasant than CS+, F(1, 30) = 16.72, p < .001, $\eta p^2 = .358$. A comparison between groups revealed that evaluations of CS- were more negative in the revaluation group than in the control group, F(1, 30) = 19.40, p < .001, $\eta p^2 = .393$, while evaluations of CS+ were more pleasant in the revaluation group than in the control group, F(1, 30) = 4.56, p = .041, $\eta p^2 = .132$ (but marginal after Bonferroni correction, $p_{adjusted}$ = .082).

First Interval Responding. Responding increased from the last block of extinction to the first block of the test phase, $F(1, 29) = 13.92$, $p = .001$, $\eta p^2 = .324$.

Second Interval Responding. Responding increased from the last block of extinction to the first block of the test phase, $F(1, 29) = 10.11$, $p = .003$, $\eta p^2 = .259$, and responding in the control group was larger than in the revaluation group, $F(1, 29) = 13.99$, $p = .001$, $\eta p^2 = .325$.

Test Phase

The CS valence evaluations, and electrodermal first and second interval responding (see the fourth panels of Figures 1, 2, and 3, respectively) recorded during the test phase were subjected to separate 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 2 (Block: 1, 2) mixed-model factorial ANOVAs.

Online CS Evaluations. A main effect of block, $F(1, 30) = 5.39$, $p = .027$, $\eta p^2 = .152$, a CS × Group interaction, $F(1, 30) = 13.25$, $p = .001$, $\eta p^2 = .306$, and a Block × Group interaction, $F(1, 30) = 6.13$, $p = .019$, $\eta p^2 = .170$, were detected. The CS × Group interaction revealed that in the control group, evaluations of CS+ and CS- did not differ, $F(1, 30) = 1.70$, $p = .203$, $\eta p^2 = .054$, however, in the revaluation group, CS- was evaluated as less pleasant than CS+, $F(1, 30) = 14.78$, $p = .001$, $\eta p^2 = .330$. A comparison between groups revealed that CS- was evaluated as more unpleasant in the revaluation group, $F(1, 30) = 15.06$, $p = .001$, $\eta p^2 = .334$, while CS+ was evaluated as marginally more pleasant in the revaluation group, $F(1, 30) = 3.89$, $p = .058$, $\eta p^2 = .115$ (but not after Bonferroni correction, $p_{adjusted} = .116$). The Block × Group interaction revealed that overall evaluations were more pleasant in block two than in block one in the control group, $F(1, 30) = 11.50$, $p = .002$, $\eta p^2 = .277$, but did not differ between blocks in the revaluation group, $F(1, 30) = 0.01$, $p = .914$, $\eta p^2 < .001$. Overall responding during block one, $F(1, 30) =$

0.79, p = .380, ηp^2 = .026, or two, $F(1, 30)$ = 3.32, p = .079, ηp^2 = .100, did not differ between groups.

First Interval Responding. Responding decreased from block one to block two, $F(1, 29)$ = 8.35, p = .007, ηp^2 = .224.

Second Interval Responding. A CS × Block interaction, $F(1, 29)$ = 4.97, p = .034, ηp^2 = .146 was detected. During block 1, responding to CS+ and CS- did not differ, $F(1, 29)$ = 1.42, p = .244, ηp^2 = .047, but during block two, responding to CS- was larger than responding to CS+, $F(1, 29)$ = 4.60, p = .040, ηp^2 = .137 (not significant after Bonferroni correction, $p_{adjusted}$ = .080).

Pleasantness Ratings

The conditional stimulus pleasantness ratings taken before habituation (ratings A), after extinction (ratings B), after reinstatement (ratings C), and post-experimentally (ratings D) were subjected to a 2 (Group: instruction, control) × 2 (CS: CS+, CS-) × 4 (Ratings: A, B, C, D) mixed-model factorial ANOVA, and are presented in Figure 4. A CS × Group interaction, $F(1, 30)$ = 15.70, p < .001, ηp^2 = .334, was moderated by a CS × Group × Ratings interaction, $F(3, 28)$ = 5.59, p = .004, ηp^2 = .374. In the control group, pleasantness ratings did not differ before conditioning, $F(1, 30)$ = 2.15, p = .153, ηp^2 = .067, but after extinction, CS+ was rated as less pleasant than CS-, $F(1, 30)$ = 7.53, p = .010, ηp^2 = .201 (not after Bonferroni correction, $p_{adjusted}$ = .80). After reinstatement, $F(1, 30)$ = 4.07, p = .053, ηp^2 = .119, and post-experimentally, $F(1, 30)$ = 3.79, p = .061, ηp^2 = .112, CS+ was evaluated as marginally less pleasant than CS-, but not after Bonferroni correction, $p_{adjusted =}$ .424, and .488, respectively. In the revaluation group, pleasantness ratings did not differ before conditioning, $F(1, 30)$ = 0.20, p = .655, ηp^2 = .007, but after extinction, after reinstatement, and post-experimentally, CS+ was evaluated as more pleasant than CS-, all $F(1, 30)$ > 10.31, p < .004, ηp^2 > .255. A comparison between groups revealed that evaluations of CS+ did not differ between groups before conditioning, $F(1, 30) = 0.06$, $p = .802$, $\eta p^2 = .002$, but after extinction, $F(1, 30) = 8.50$, $p = .007$, $\eta p^2 = .221$, after reinstatement, $F(1, 30) = 6.74$, $p = .014$, $\eta p^2 = .184$, and post-experimentally, $F(1, 30) = 4.63$, $p = .040$, $\eta p^2 = .134$, CS+ was evaluated as more pleasant in the revaluation group than in the control group (after Bonferroni correction, this difference was marginal after extinction, $p_{adjusted} = .056$, but not significant after reinstatement, $p_{adjusted} = .112$, and post-experimentally, $p_{adjusted} = .320$). Evaluations of the CS- did not differ between the groups before conditioning, $F(1, 30) = 1.21$, $p = .280$, $\eta p^2 = .039$, however after extinction, $F(1, 30) = 18.62$, $p < .001$, $\eta p^2 = .383$, reinstatement, $F(1, 30) = 12.26$, $p = .001$, $\eta p^2 = .290$, and post-experimentally, $F(1, 30) = 15.72$, $p < .001$, $\eta p^2 = .344$, the revaluation group evaluated the CS- as less pleasant than the control group.

Discussion

In the current study we examined whether conditional valence evaluations, acquired based on CS-US pairings, could be reduced, or reversed, using instructions targeting CS valence. After fear acquisition, the control group received neutral information about both CS posers, while the revaluation group received positive character information about the CS+ poser and negative character information about the CS- poser. We examined CS valence online, throughout conditioning training, and offline, in pleasantness ratings taken before habituation, after extinction, after reinstatement, and post-experimentally. An affective priming task was administered as an implicit measure of CS valence and electrodermal responding was measured throughout conditioning to examine whether the revaluation instructions would also affect a physiological index of fear learning.

Throughout acquisition, CS+ was evaluated as less pleasant than CS- in both groups, but unexpectedly, during the last block of acquisition this pattern of evaluations was only significant in the control group (marginal before Bonferroni correction in the revaluation group). After the revaluation instructions however, the pattern of differential CS valence reversed, such that participants evaluated CS+ as more pleasant than CS-. During the first block of extinction, negative CS+ valence was no longer present and throughout extinction the CS+ was evaluated as pleasant. The CS- was evaluated as unpleasant immediately and remained unpleasant during extinction. These valence changes were stable across extinction and were not affected by the reinstatement procedure. We also did not observe relapse of conditional valence in the ratings tasks taken after extinction, after reinstatement and post-experimentally. These results are encouraging as they suggest that not only can negative CS valence be removed using instructions targeting CS valence, but that such a change in valence might be less susceptible to relapse. Unexpectedly, we did not observe the same result in the affective priming task which, across groups, yielded no evidence of conditioning. This task was taken after the extinction phase and may not have been sensitive enough to detect residual valence, especially after a manipulation which aimed to reverse the pattern of differential valence acquired during acquisition.

Interestingly, CS+ valence also increased after the control instructions, but evaluations of CS+ were still marginally more negative than evaluations of CS- at the beginning of extinction. It is likely that giving the CS+ poser a name and occupation 'normalized' him, possibly having a weak positive revaluation effect. This could indicate that merely familiarizing participants with a social feared stimulus could weaken previously acquired negative valence. Differential evaluations in the control group extinguished very quickly and were not present in the post-experimental rating tasks. Prior research has indicated that CS evaluations are slow to extinguish and often their extinction requires extended training (Hofmann, et al., 2010). It is likely that the control instructions, which did slightly increase CS+ valence, also facilitated the speed of extinction.

Contrary to our hypothesis, electrodermal responding was also affected by the revaluation manipulation. Differential first and second interval responding was acquired during acquisition in both groups. This differential responding remained intact in the control group during the first block of extinction, but was eliminated in the revaluation group. Examination of Figure 2 suggests that differential responding was eliminated due to a decrease in first interval responding to CS+ and an increase to CS- . This pattern does not fit well with an emotional arousal account, as heightened positive and negative emotional arousal results in increased electrodermal responding (Cuthbert, Bradley & Lang 1996), but it is possible that these changes reflect an expectation that the contingencies would change after the manipulation. Although participants were not explicitly informed about the CS-US contingency, if they learned throughout acquisition that only one CS would be paired with the US, changing the CS valence in the revaluation group could have resulted in an increased expectancy of the US following CS-, and a decreased expectancy of the US following CS+. This would be consistent with evidence suggesting that negative stimuli are more readily associated with aversive events (Hamm, Vaitl, & Lang, 1989) and could explain why the revaluation group shows smaller first interval responding to the CS+ and larger responding to the CS-. Unexpectedly, differential second interval electrodermal responding was eliminated in both groups at the beginning of extinction. This result is surprising as second interval responding is typically a robust index of US anticipation and less sensitive to changes in orienting or emotional arousal which can be influenced by the instructions. As both groups show an elimination of differential second interval responding it seems unlikely that this effect is specific to the revaluation manipulation, however, it is not clear what caused this reduction in second interval responding.

Significant differential reinstatement did not occur in the revaluation group, however unfortunately, it is not possible to assess whether the revaluation instructions decreased reinstatement levels because we did not observe reinstatement in the control group either. It is possible that the affective priming task, which involved 48 unreinforced presentations of CS+ and CS- and was administered between extinction and reinstatement, could have functioned as extended extinction training. There is some evidence to suggest that massive extinction may attenuate the return of fear (see Denniston, Chang & Miller, 2003; Laborda & Miller, 2014) and therefore extending extinction training in this case could have reduced overall reinstatement rates.

Experiment 2

Experiment 1 revealed that the pattern of differential CS valence evaluations, acquired during acquisition training, can reversed by providing positive character information about the CS+ poser and negative character information about the CS- poser. This result is encouraging for clinical practice as it indicates that the negative valence towards the feared stimulus can be changed with instructions targeting the valence of the feared stimulus. In Experiment 1, the instructions revalued the feared stimulus (CS+) to be pleasant and the safety signal (CS-) to be unpleasant. Revaluating a safety signal is not desirable in a clinical setting and after providing evidence that CS valence can be changed via an instructional manipulation in Experiment 1, we conducted Experiment 2 to confirm that the CS+ can be positively revaluated without providing negative information about another stimulus. Participants in the positive revaluation group were given positive character information about the CS+ poser and neutral information about the CS-poser, and participants in the negative revaluation group were given negative character information about the CS+ poser and neutral information about the CS- poser. We hypothesized that the CS+ would be evaluated as more pleasant after positive revaluation and less pleasant after negative revaluation. We also expected these valence changes to be stable throughout extinction, reinstatement test, and the rating tasks completed after the experiment. We were interested in exploring whether the pattern of differential electrodermal responding following instructions in Experiment 1 would replicate and hypothesized that differential reinstatement of electrodermal responses would be larger in the negative revaluation group. The affective priming task and the rating task following extinction were removed to increase the chances of finding a significant reinstatement effect and to avoid an interruption between the extinction and the reinstatement procedure.

Method

Participants

Thirty four undergraduate students (25 female), aged between 17 – 46 years (*M* = 22.71) provided informed consent, volunteered participation in exchange for course credit, and were randomly assigned to the positive revaluation or the negative revaluation group.

Apparatus/Stimuli

The apparatus and stimuli used were the same as Experiment 1.

Procedure

The positive, negative, and neutral instructions were the same as in Experiment 1 and were counter-balanced across participants. Participants in the negative revaluation group received negative character information about the CS+ poser and neutral information about the CS- poser and participants in the positive revaluation group received positive character information about the CS+ poser and neutral information about the CS- poser. The affective priming measure and the rating task immediately following extinction were removed. An additional question was added to the post-experimental questionnaire to assess participants' expectations of the CS-US contingency immediately after the revaluation instructions. They were asked whether they expected the electrotactile stimulus to: stay the same, to stop, or to switch to the other face. The remainder of the procedure, the scoring, and the statistical analyses were conducted in the same manner as Experiment 1.

Results

Preliminary Checks

The descriptive statistics for the preliminary checks are presented in Table 1. A Pearson's chi-square test confirmed that the gender ratio did not differ between groups, $\chi^2(1) = 0.15$, $p = .697$. A series of independent samples t-tests revealed that the groups did not differ in age, spontaneous electrodermal responding, US intensity, or US valence, all $t < 1.61$, $p > .119$. A 2 (Group: positive, negative) × 4 (Block: 1, 2, 3, 4) factorial ANOVA confirmed that unconditional electrodermal responses did not differ between groups, all $F < 0.42$, $p > .748$, $\eta p^2 < .040$. All participants were able to report which instructions they had received, but one participant from the negative revaluation group was not able to correctly identify which face was used as the CS+. When this participant was removed from the analyses the conclusions did not change and therefore results from the full sample are reported. Seven participants (4 positive revaluation, 3 negative revaluation) indicated that they expected the US to swap to the CS-, and two participants (1 positive revaluation, 1 negative revaluation) indicated they expected the US to stop after the instructions. When these participants are removed from the analyses the conclusions did not change and therefore results from the entire sample are reported.

Habituation

The CS valence evaluations and first interval responding (see the first panels of Figures 5 and 6, respectively) recorded during habituation were subjected to separate 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 2 (Block: 1, 2) mixed-model factorial ANOVAs.

Online CS Evaluations. No significant effects were detected, all $F < 3.06$, $p > .089$, $\eta p^2 < .088$.

First Interval Responding. Responding decreased from block one to block two, $F(1, 32) = 18.12$, $p < .001$, $\eta p^2 = .361$.

Acquisition

The CS valence evaluations, electrodermal first and second interval responding recorded during acquisition were subjected to separate 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 4 (Block: 1, 2, 3, 4) mixed model factorial ANOVAs and are presented in Figures 5 (second panel), 6 (second panel), and 7 (first panel), respectively.

Online CS Evaluations. A main effect of CS, $F(1, 32) = 12.53$, $p = .001$, $\eta p^2 = .281$, was moderated by a CS × Block interaction, $F(3, 30) = 6.00$, $p = .002$, $\eta p^2 = .375$. During block one, $F(1, 32) = 2.05$, $p = .162$, $\eta p^2 = .060$, evaluations did not differ between CS+ and CS-, however during subsequent blocks CS+ was rated as less pleasant than CS-, all $F(1, 32) > 8.85$, $p < .007$, $\eta p^2 > .216$.

First Interval Responding. Main effects of CS, $F(1, 32) = 30.61$, $p < .001$, $\eta p^2 = .489$, and block, $F(3, 30) = 9.68$, $p < .001$, $\eta p^2 = .492$, were qualified by a CS × Block interaction, $F(3, 30) = 9.58$, $p < .001$, $\eta p^2 = .489$. During block one, responding to CS+ and CS- did not differ, $F(1, 32) = 0.09$, $p = .767$, $\eta p^2 = .003$, but during subsequent blocks responding to CS+ was larger than responding to CS-, all $F (1, 32) > 16.37$, ps $< .001$, ηp^2s $> .338$.

Second Interval Responding. A main effect of CS, $F(1, 32) = 48.22$, $p < .001$, $\eta p^2 = .601$, was moderated by a CS × Block interaction, $F(3, 30) = 9.01$, $p < .001$, $\eta p^2 = .474$. Responding to CS+ and CS- did not differ during block one, $F(1, 32) = 0.30$, $p = .589$, $\eta p^2 = .009$, however during subsequent blocks, responding to CS+ exceeded responding to CS-, all $F(1, 32) > 14.27$, $p < .002$, $\eta p^2 > .308$.

Instructional Manipulation

To assess the immediate influence of the instructions on CS valence evaluations, electrodermal first, and second interval responding, separate 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 2 (Phase: last block of acquisition, first block of extinction) mixed model factorial ANOVAs were conducted.

Online CS Evaluations. A main effect of CS, $F(1, 32) = 26.43$, $p < .001$, $\eta p^2 = .452$, a Phase × Group interaction, $F(1, 32) = 7.82$, $p = .009$, $\eta p^2 = .196$, and a CS × Group interaction, $F(1, 32) = 4.17$, $p = .049$, $\eta p^2 = .115$, were moderated by a CS × Phase × Group interaction, $F(1, 32) = 29.48$, $p < .001$, $\eta p^2 = .480$. During the last block of acquisition, CS+ was evaluated as less pleasant than CS-, in the negative revaluation group, $F(1, 32) = 5.98$, $p = .020$, $\eta p^2 = .157$ (marginal after Bonferroni correction, $p_{adjusted} = .080$), and in the positive revaluation group, $F(1, 32) = 11.77$, $p = .002$, $\eta p^2 = .269$. Whereas during the first block of extinction, CS+ was evaluated as less pleasant than CS- in the negative revaluation group, $F(1, 32) = 48.33$, $p < .001$, $\eta p^2 = .602$, but differential evaluations between CS+ and CS- were not present in the positive revaluation group, $F(1, 32) = 0.03$, $p = .875$, $\eta p^2 = .001$. During the last block of acquisition, the groups did not differ in evaluations of CS+, $(1, 32) = 3.41$, $p = .074$, $\eta p^2 = .096$ (after Bonferroni correction, $p_{adjusted} = .296$), or CS-, $F\ (1, 32) = 0.27$, $p = .604$, $\eta p^2 = .008$. After revaluation, the CS+ was evaluated as more pleasant by the positive revaluation group in comparison with the negative revaluation group, F (1, 32) = 17.34, $p < .001$, $\eta p^2 = .351$; while, CS- was evaluated as more unpleasant in the positive revaluation group in comparison with the negative revaluation group, $F(1, 32) = 6.10$, $p = .019$, $\eta p^2 = .160$ (marginal after Bonferroni correction, $p_{\text{adjusted}} = .076$).

First Interval Responding. A main effect of CS, $F(1, 32) = 16.74$, $p < .001$, $\eta p^2 = .343$, was moderated by a CS × Phase × Group interaction, $F(1, 32) = 5.34$, $p = .027$, $\eta p^2 = .143$. During the last block of acquisition responses to CS+ were larger than to CS- in the positive revaluation group, $F(1, 32) = 6.85$, $p = .013$, $\eta p^2 = .176$ (marginal after Bonferroni correction, $p_{\text{adjusted}} = .052$), and the negative revaluation group, $F(1, 32) = 9.95$, $p = .003$, $\eta p^2 = .237$. This pattern persisted in the positive revaluation group during the first block of extinction, $F(1, 32) = 11.96$, $p = .002$, $\eta p^2 = .272$, but differential responding was not present during the first block of extinction in the negative revaluation group, $F(1, 32) = 0.15$, $p = .699$, $\eta p^2 = .005$. A comparison between the groups revealed that responding to CS- was larger in the negative revaluation group in comparison with the positive revaluation group, $F(1, 32) = 4.76$, $p = .037$, $\eta p^2 = .129$, (not after Bonferroni correction, $p_{\text{adjusted}} = .148$); while responding to CS+ and CS- did not differ between the groups at any other time, $F(1, 32) < 0.77$, $p > .385$, $\eta p^2 < .025$.

Second Interval Responding. Responses to CS+ were larger than to CS-, $F(1, 32) = 15.14$, $p < .001$, $\eta p^2 = .321$.

Extinction

The CS valence evaluations, first interval responding, and second interval responding recorded during extinction were subjected to separate 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 4 (Block: 1, 2, 3, 4) mixed model factorial ANOVAs and are presented in Figures 5 (third panel), 6 (third panel), and 7 (second panel), respectively.

Online CS Evaluations. A main effect of CS, $F(1, 32) = 21.57$, $p < .001$, $\eta p^2 = .403$, a main effect of block, $F(3, 30) = 5.17$, $p = .005$, $\eta p^2 = .341$, and a CS × Group interaction, $F(1, 32) = 26.73$, $p < .001$, $\eta p^2 = .455$, were detected. Overall evaluations in block one were more unpleasant than in blocks two, $p < .021$, three, $p < .004$, and four, $p < .002$. The CS × Group interaction revealed that evaluations of CS+ were less pleasant than CS- in the negative revaluation group, $F(1, 32) = 48.16$, $p < .001$, $\eta p^2 = .601$, but that evaluations between CS+ and CS- did not differ in the positive revaluation group, $F(1, 32) = 0.14$, $p = .713$, $\eta p^2 = .004$. A comparison between groups revealed that the CS+ was evaluated as more pleasant, $F(1, 32) = 22.47$, $p < .001$, $\eta p^2 = .413$, and the CS- as more unpleasant, $F(1, 32) = 6.60$, $p = .015$, $\eta p^2 = .171$, in the positive revaluation group in comparison with the negative revaluation group.

First Interval Responding. Responding to CS+ was larger than responding to CS-, $F(1, 32) = 17.23$, $p < .001$, $\eta p^2 = .350$, and responding in block one was larger than responding in all subsequent blocks, all $p < .002$, $F(3, 30) = 7.77$, $p = .001$, $\eta p^2 = .437$.

Second Interval Responding. No significant effects were detected interaction during extinction, all $F < 2.67$, $p > .064$, $\eta p^2 < .212$.

Reinstatement Test

To assess the immediate influence of the reinstatement procedure on CS valence evaluations, first interval responding, and second interval responding separate 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 2 (Phase: last block of extinction, first block of the test phase) mixed model factorial ANOVAs were conducted.

Online CS Evaluations. A main effect of CS, $F(1, 32) = 25.52$, $p < .001$, $\eta p^2 = .444$, a CS × Group interaction, $F(1, 32) = 28.12$, $p < .001$, $\eta p^2 = .468$, and a Phase × Group interaction, $F(1, 32) = 5.63$, $p = .024$, $\eta p^2 = .150$, were detected. The CS × Group interaction revealed that the negative revaluation group evaluated CS+ as less pleasant than CS-, $F(1, 32) = 53.61$, $p < .001$, $\eta p^2 = .626$, but the positive revaluation group did not evaluate CS+ and CS- differently, $F(1, 32) = 0.03$, $p = .861$, $\eta p^2 = .001$. The Phase × Group interaction revealed that, while overall evaluations did not differ between the groups during the last block of extinction, $F(1, 32) = 1.65$, $p = .208$, $\eta p^2 = .049$, or the first block of the test phase, $F(1, 32) = 0.52$, $p = .476$, $\eta p^2 = .016$, overall evaluations in the positive revaluation group were less pleasant during the first block of the test phase, $F(1, 32) = 9.19$, $p = .005$, $\eta p^2 = .223$. Overall evaluations in the negative revaluation group did not differ between phases, $F(1, 32) = 0.11$, $p = .748$, $\eta p^2 = .003$.

First Interval Responding. A main effect of CS, $F(1, 32) = 4.67$, $p = .038$, $\eta p^2 = .127$, a main effect of phase, $F(1, 32) = 7.51$, $p = .010$, $\eta p^2 = .190$, and a marginal CS × Phase interaction, $F(1, 32) = 4.13$, $p = .050$, $\eta p^2 = .114$, were detected. During the last block of extinction differential responding to CS+ and CS- was not present, $F(1, 32) = 0.15$, $p = .702$, $\eta p^2 = .005$, but differential responding re-emerged during the first block of the test phase, with responding to CS+ exceeding responding to CS-, $F(1, 32) = 6.46$, $p = .016$, $\eta p^2 = .168$.

Second Interval Responding. A main effect of phase, $F(1, 32) = 4.33$, $p = .046$, $\eta p^2 = .119$, was moderated by a CS × Phase interaction, $F(1, 32) = 8.42$, $p = .007$, $\eta p^2 = .208$. During the last block of extinction, differential responding to CS+ and CS- was not present, $F(1, 32) = 0.60$, $p = .445$, $\eta p^2 = .018$, however, during the first block of the test phase responding to CS+ was marginally higher than responding to CS-, $F(1, 32) = 3.75$, $p = .062$, $\eta p^2 = .105$ (but not after Bonferroni correction, $p_{\text{adjusted}} = .124$).

Test Phase

The CS valence evaluations, electrodermal first and second interval responding (see the last panels of Figures 5, 6, and 7, respectively) recorded during the test phase were subjected to separate 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 2 (Block: 1, 2) mixed-model factorial ANOVAs.

Online CS Evaluations. A main effect of CS, $F(1, 32) = 22.89$, $p < .001$, $\eta p^2 = .417$, and a CS × Group interaction were detected, $F(1, 32) = 23.26$, $p < .001$, $\eta p^2 = .421$. The negative revaluation group evaluated CS+ as less pleasant than CS-, $F(1, 32) = 46.14$, $p < .001$, $\eta p^2 = .590$, while evaluations of CS+ and CS- did not differ in the positive revaluation group, $F(1, 32) < 0.01$, $p = .978$, $\eta p^2 < .001$. A comparison between groups revealed that the CS+ was evaluated as more pleasant, $F(1, 32) = 17.66$, $p < .001$, $\eta p^2 = .356$, and the CS- as less pleasant, $F(1, 32) = 7.37$, $p = .011$, $\eta p^2 = .187$, in the positive revaluation group in comparison with the negative revaluation group.

First Interval Responding. Responding to CS+ was larger than responding to CS-, $F(1, 32) = 12.71$, $p = .001$, $\eta p^2 = .284$, and responding decreased from block one to block two, $F(1, 32) = 6.81$, $p = .014$, $\eta p^2 = .176$.

Second Interval Responding. Responding decreased from block one to block two, $F(1, 32) = 4.63$, $p = .039$, $\eta p^2 = .126$.

Pleasantness Ratings

The conditional stimulus pleasantness ratings taken before habituation (ratings A), after extinction (ratings B), and post-experimentally (Ratings C) were subjected to a 2 (Group: positive, negative) × 2 (CS: CS+, CS-) × 3 (Ratings: A, B, C) mixed-model factorial ANOVA, and are presented in Figure 8. A main effect of CS, $F(1, 32) = 23.68$, $p < .001$, $\eta p^2 = .425$, a CS × Group, $F(1, 32) = 17.43$, $p < .001$, $\eta p^2 = .353$, and a CS × Ratings interaction, $F(2, 31) = 30.51$, $p < .001$, $\eta p^2 = .663$, were qualified by a CS × Group × Ratings interaction, $F(2, 31) = 9.92$, $p < .001$, $\eta p^2 = .390$. The negative revaluation group rated CS+ as more pleasant than CS- before conditioning, $F(1, 32) = 5.56$, $p = .025$, $\eta p^2 = .148$ (but not after Bonferroni correction, $p_{adjusted}$ = .150), however after reinstatement, $F(1, 32) = 40.80$, $p < .001$, $\eta p^2 = .560$, and post-experimentally, $F(1, 32) = 65.75$, $p < .001$, $\eta p^2 = .673$, CS+ was rated as less pleasant than CS-. In the positive revaluation group, evaluations of CS+ and CS- did not differ before conditioning, $F(1, 32) = 3.29$, $p = .079$, $\eta p^2 = .093$ (after Bonferroni correction, $p_{adjusted} = .474$), or after reinstatement, $F(1, 32) < 0.01$, $p > .999$, $\eta p^2 < .001$, but post-experimentally, CS+ was evaluated as less pleasant than CS-, $F(1, 32) = 4.77$, $p = .036$, $\eta p^2 = .130$ (not significant after Bonferroni correction, $p_{adjusted} = .216$). A comparison between the groups revealed that evaluations of CS+ did not differ before conditioning, $F(1, 32) = 1.24$, $p = .274$, $\eta p^2 = .037$, but after reinstatement and post-experimentally, the CS+ was evaluated as more pleasant in the positive revaluation group, $F(1, 32) > 14.23$, $p < .002$, $\eta p^2 > .307$. Evaluations of CS- did not differ before conditioning, $F(1, 32) = 0.36$, $p = .551$, $\eta p^2 = .011$, but after reinstatement, $F(1, 32) = 5.93$, p = .021, $\eta p^2 = .156$, and post-experimentally, $F(1, 32) = 4.69$, $p = .038$, $\eta p^2 = .128$, evaluations of CS- were less pleasant in the positive revaluation group (but these differences were not significant after Bonferroni correction, $p_{adjusted} = .126$, $p_{adjusted} = .228$, respectively).

Discussion

Experiment 2 was conducted to confirm that negative CS+ valence could be removed with a revaluation manipulation which did not rely on negatively revaluating a safety stimulus (CS-) and to assess whether reinstatement varies as a function of residual negative CS+ valence at the end of extinction. The positive revaluation group received positive character information about the CS+ poser and neutral information about the CS- poser, while the negative revaluation group received negative character information about the CS+ poser and neutral information about the CS- poser. As expected, following the manipulation, CS+ was evaluated as less pleasant in the negative revaluation group (strengthening differential valence evaluations) and as more pleasant in the positive revaluation group (abolishing differential valence evaluations). Positive revaluation removed the majority of the negative CS+ valence acquired during acquisition which is encouraging as it suggests that the positive revaluation instructions rendered the valence of the CS+ the same as that of a safety signal (CS-). As in Experiment 1, the revaluation effect was stable in both groups across extinction and after reinstatement. The positive revaluation group seemed to show some relapse in the post-experimental rating task, with CS+ evaluated as less pleasant than CS-. As this effect was not significant after Bonferroni correction and is not consistent with Experiment 1, it seems more work is required to confirm whether CS valence acquired during Pavlovian fear conditioning and altered using instructional revaluation is subject to relapse.

The valence manipulation did not affect electrodermal second interval responding in either group or electrodermal first interval responding in the positive revaluation group, but eliminated differential first interval responding in the negative revaluation group. Visual inspection of Figure 6 suggests that this elimination was due to an increase in responding to CS- rather than a decrease in responding to CS+. As CS- valence was not changed, this increase in responding is unlikely to reflect a change in emotional arousal. It is also unlikely to be driven by the expectation that the experimental contingencies would switch, as the removal of participants who expected a contingency change did not alter the results. It is not clear what drives the initial increase in responding to CS- in the negative revaluation group, but differential responding re-stabilizes when the entire extinction phase is considered, with CS+ eliciting larger responding than CS- throughout extinction, in both groups. Differential electrodermal responding had extinguished by the last block of extinction in both groups, and re-emerged following the reinstatement manipulation in both groups. Unexpectedly, there was no evidence that the size of this effect was moderated by the CS revaluation manipulation.

Overall Discussion

Across two experiments we investigated whether an instructional intervention targeting CS valence would affect CS valence evaluations acquired based on CS-US pairings throughout fear acquisition. In both experiments, we examined CS valence, online, throughout conditioning, and offline, in ratings tasks following extinction (Experiment 1 only), reinstatement, and post-experimentally in a different context. We measured electrodermal responding to examine whether the revaluation instructions would affect a physiological index of fear learning and were particularly interested in whether positively revaluating the CS+ would reduce the reinstatement of differential electrodermal responding. In Experiment 1, we compared positive CS+ revaluation and negative CS- revaluation with a control group who received neutral information about both CSs. In Experiment 2, we examined positive and negative CS+ revaluation between groups, by giving positive/negative information about the CS+ and neutral information about the CS-.

In both experiments, instructions immediately influenced CS valence evaluations in the predicted direction. In Experiment 1, after the revaluation instructions, negative CS+ valence was removed and CS+ was evaluated as more pleasant than CS-. In Experiment 2, the positive revaluation instructions removed the majority of negative CS+ valence, but the CS+ did remain slightly unpleasant (but not different from the CS-). The negative revaluation instructions increased negative CS+ valence and differential CS valence evaluations were strengthened by the manipulation. In both experiments, these valence changes were stable across the entire extinction phase and following the reinstatement procedure. In Experiment 1, the revaluation effect was also present in the ratings tasks taken following extinction, reinstatement, and post-experimentally. In Experiment 2, both positive and negative revaluation was stable across extinction and the reinstatement test phase, but some relapse was observed in the positive revaluation group during the post-experimental ratings task. This relapse was not significant after Bonferroni correction and requires further investigation.

Resistance to reinstatement could indicate that the revaluation manipulation was successful at intrinsically changing the value of the CS, rather than creating an additional CS-US association. There are debates about the type of association that underlies evaluative learning. Evaluative learning could occur because of a link between the CS and the unconditional response (signal-response learning; S-R; i.e. intrinsic change) or between the CS and the US (signal-signal learning; S-S; i.e. referential change). In US revaluation studies, evidence for both S-S learning (see Baeyens, Eelen, Van den Bergh, & Crombez, 1992) and S-R learning have been reported (see Gast & Rothermund, 2011). Gast and Rothermund (2011) argue that the type of link that is formed could depend on whether the US elicits evaluative responses during conditioning. This may depend on the stimuli that are used and whether there is a goal to evaluate (i.e. whether participants are asked to give evaluations during the conditioning task. As the CS revaluation manipulation targets the intrinsic value of the CS (i.e. Ben is a nice person) and does not give participants information about a new CS-US contingency (i.e. Ben will now be associated with pleasant stimuli), we believe it is plausible that the observed evaluative changes could reflect S-R learning, but more research will be required to determine the specific parameters underlying whether S-R or S-S learning will occur during evaluative conditioning and evaluative conditioning via instructions.

Despite the considerable difference in CS+ valence between the groups, we did not find evidence that positively revaluating the CS+ reduced reinstatement levels. Similar to Hermans et al. (2005) we conducted exploratory analyses to examine whether CS+ valence during the last block of extinction was correlated with reinstatement rates, but no significant correlations were detected in Experiment 1 or 2. It is possible that the moderating influence of CS+ valence on electrodermal responding was subsumed in an overall increase in arousal levels after the reinstatement procedure. The moderating influence of CS+ valence might be better detected using indices of fear learning which are less sensitive to changes in the participants' overall arousal levels.

Unexpectedly, CS revaluation also influenced electrodermal responding, but not consistently between the experiments. In Experiment 1, differential first interval responding remained intact in the control group, but was abolished in the revaluation group, a result which could suggest that positively revaluating the CS+ facilitates extinction learning. The pattern of electrodermal responding in Experiment 2, however, did not provide further evidence for this interpretation. In Experiment 2, while differential second interval responding was intact in both groups at the beginning of extinction, differential first interval responding was present in the positive revaluation group, but abolished in the negative revaluation group. This latter finding seems to be driven by increased responding to CS-, but it is not clear what might have caused this increase. CS- valence was not changed in Experiment 2 and therefore an emotional arousal account seems unlikely. It is possible that the participants suspected the contingencies to switch and therefore anticipation of the US may increase after the CS-. We did not find any evidence for this explanation when excluding the participants who reported expecting a contingency change, but it is possible that their number was underestimated in the post-experimental assessments of US expectancy.

The results of the current investigation provide strong evidence that it is possible to positively revaluate a feared stimulus with instructions specifically targeting valence and that such a revaluation is remarkably stable across the extinction and test phases. More research is required however to clarify some aspects of these results. It is not clear why CS revaluation influenced electrodermal responding in a different manner between experiments. An online assessment of US expectancy may provide a more reliable indication of participants' expectations of the CS-US contingency at the beginning of extinction. We also did not include a manipulation check at the end of the experiments to verify that the participants perceived the positive instructions as positive, the negative instructions as negative, and the neutral instructions as neutral. It is possible that the neutral instructions in particular were not viewed as neutral, but slightly positive as they suggested that the individuals were employed and functioning in society. Future research should also examine whether CS revaluation would reduce the reinstatement of differential fear learning using indices that are less sensitive to changes in general arousal levels, such as fear ratings and fear potentiated startle and whether the valence changes are stable after a longer time period. Although more work is required, the current results are encouraging as they indicate that negative valence towards a feared stimulus can be removed with instructions that target stimulus valence and that, once established, this valence change is remarkably stable.

Table 1.

Means and Standard Deviations for the Variables Assessed in the Preliminary Analyses.

	Experiment 1		Experiment 2	
	Control	Revaluation	Positive	Negative
Gender Ratio (M:F)	3:13	8:8	4:13	5:12
Age	23.50 (7.37)	24.31 (7.83)	23.06 (7.34)	22.35 (7.60)
Spontaneous EDA	19.44 (13.51)	22.13 (15.21)	12.76 (10.79)	17.88 (12.83)
US Level (volts)	38.81 (12.85)	42.50 (12.38)	50.88 (14.28)	60.29 (19.64)
US Valence	-1.63 (0.81)	-1.88 (0.62)	-1.59 (0.51)	-1.71 (0.69)

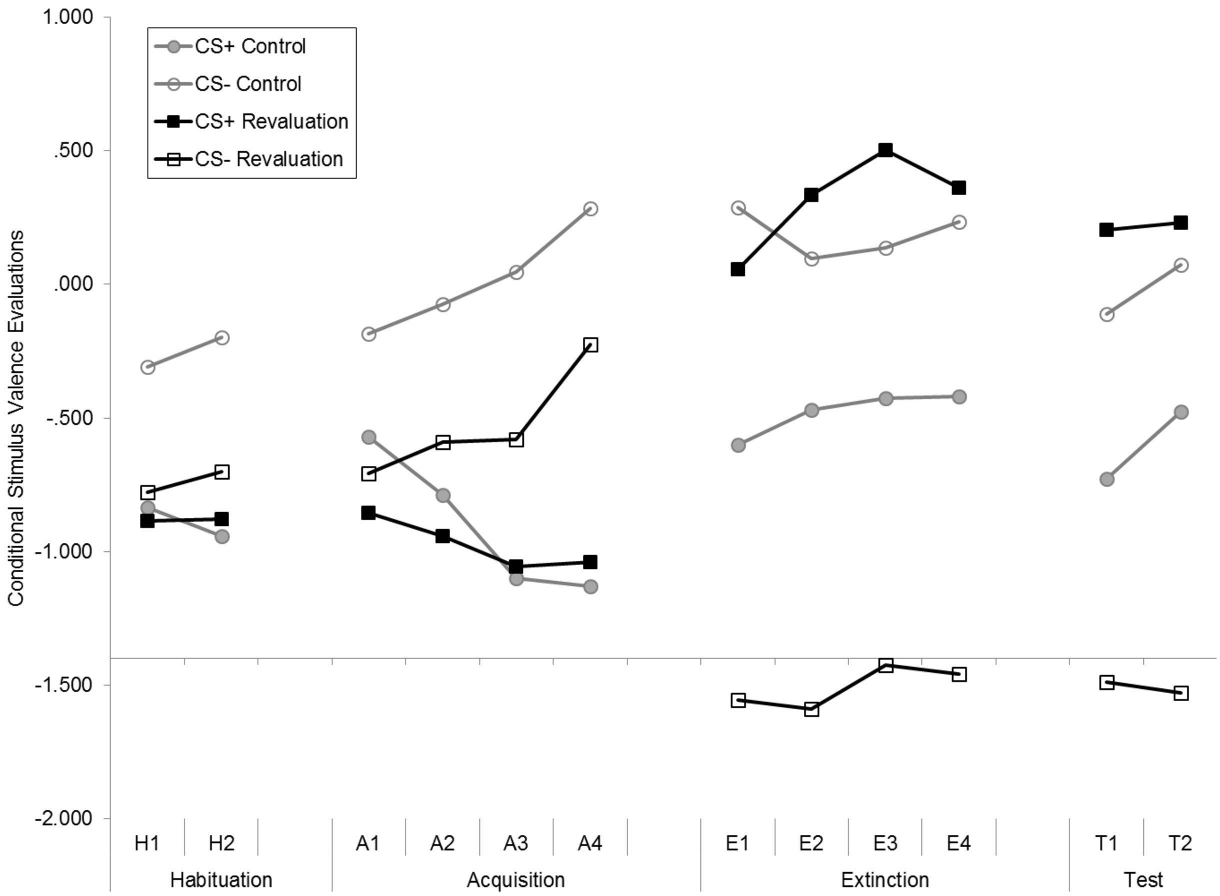

Figure 1. Conditional stimulus valence evaluations recorded throughout habituation, acquisition, extinction, and test phases of Experiment 1. Relative to neutral (zero value), negative values indicate negative evaluation and positive values indicate positive evaluation. (Range -2.5 to 2.5).

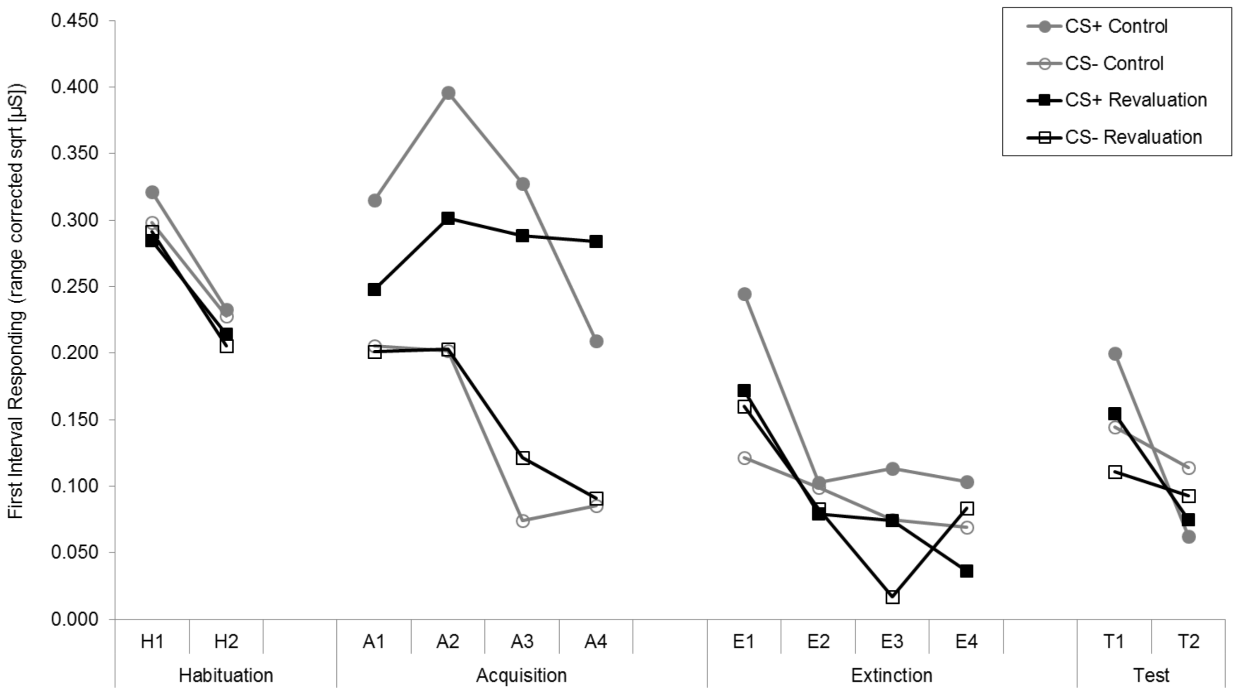

Figure 2. Electrodermal first interval responses recorded during habituation, acquisition, extinction, and test phases of Experiment 1.

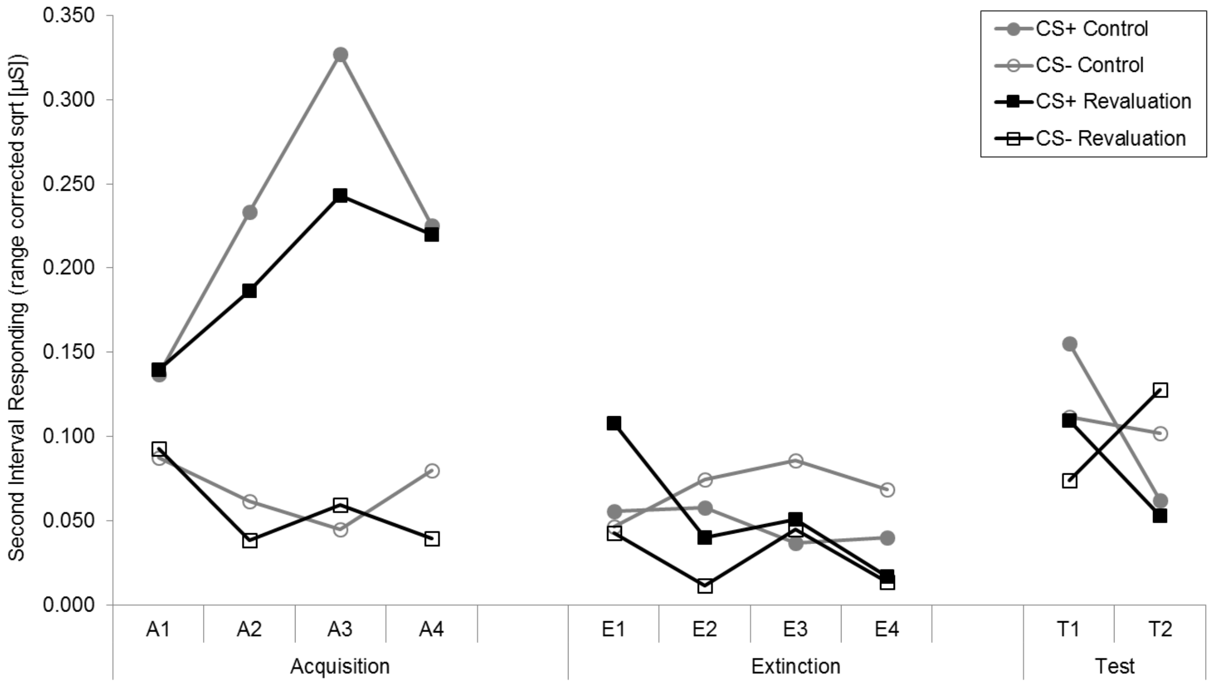

Figure 3. Electrodermal second interval responses recorded during acquisition, extinction, and test phases of Experiment 1.

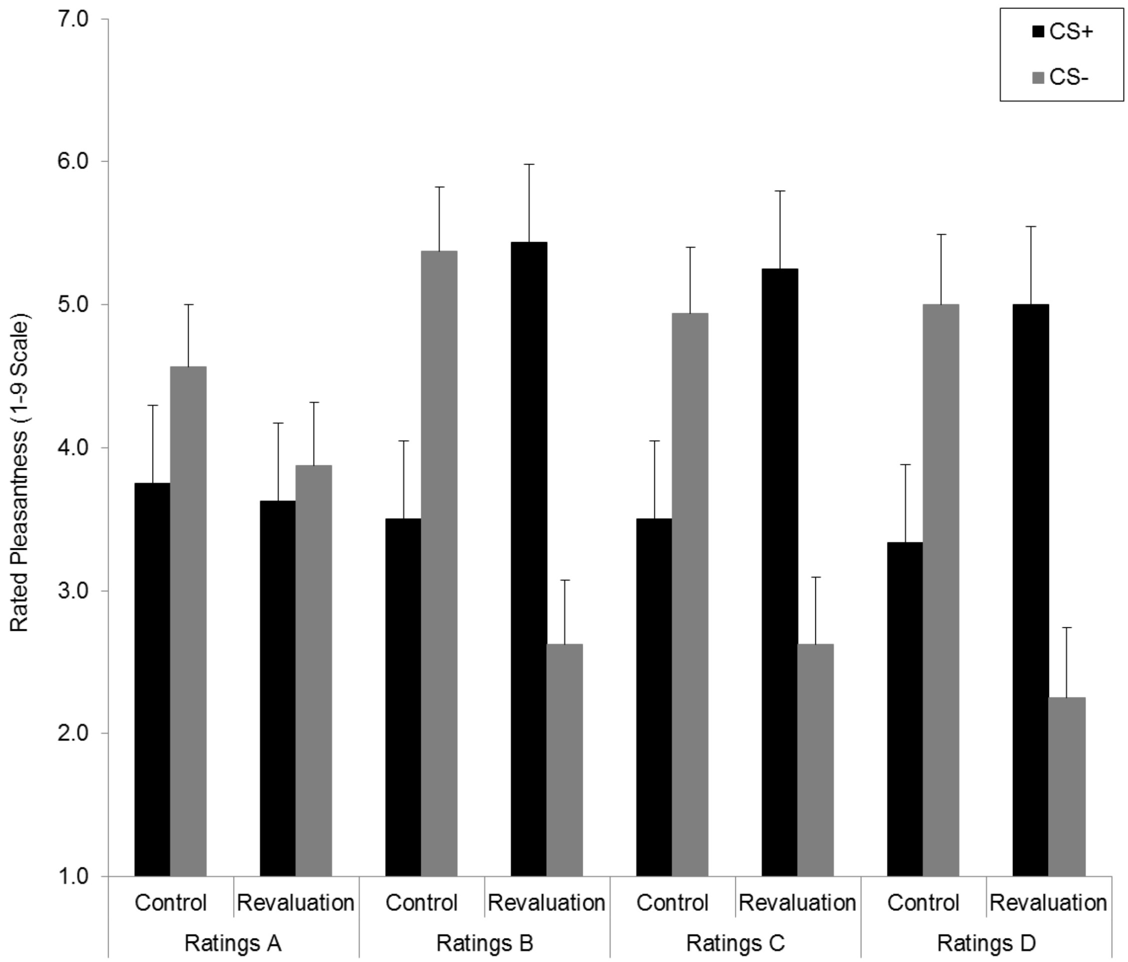

Figure 4. Conditional stimulus valence ratings taken before habituation (ratings A), after extinction (ratings B), after reinstatement (ratings C), and post-experimentally (ratings D) in Experiment 1. (Error bars indicate standard error of the mean; error bars can only be used to assess between group differences).

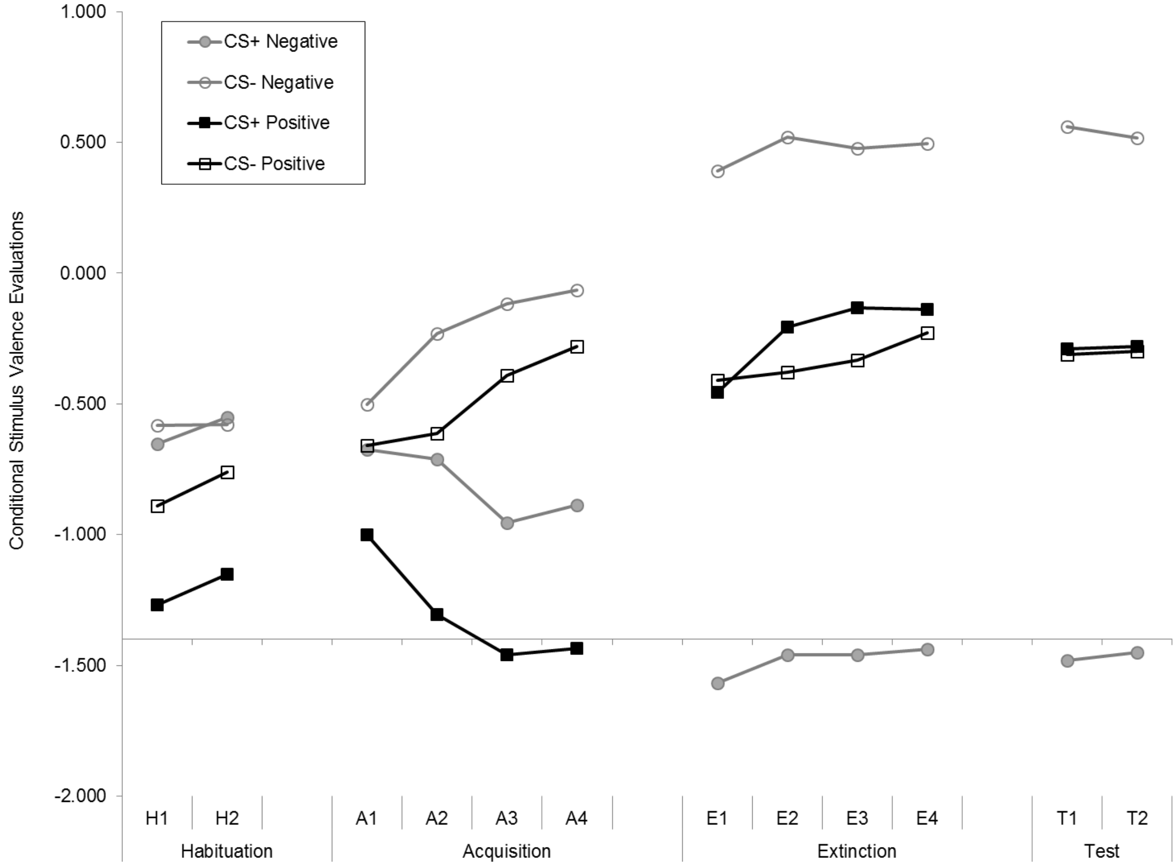

Figure 5. Conditional stimulus valence evaluations recorded throughout habituation, acquisition, extinction, and test phases of Experiment 2. Relative to neutral (zero value), negative values indicate negative evaluation and positive values indicate positive evaluation. (Range -2.5 to 2.5).

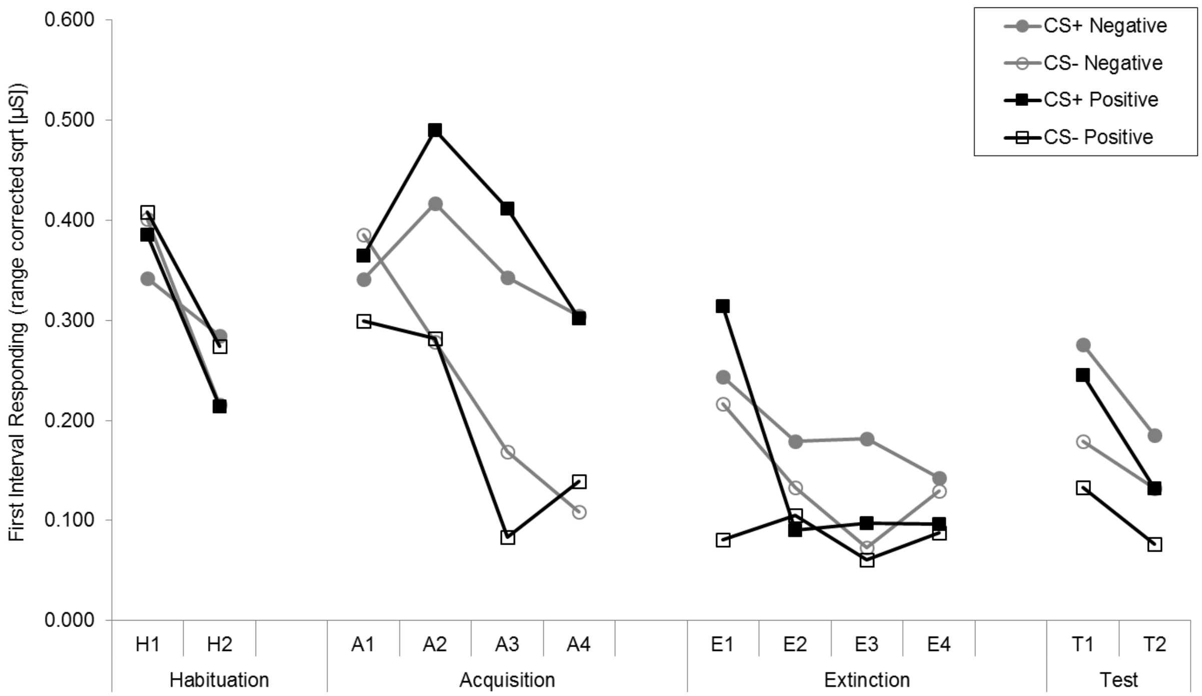

Figure 6. Electrodermal first interval responses recorded during habituation, acquisition, extinction, and test phases of Experiment 2.

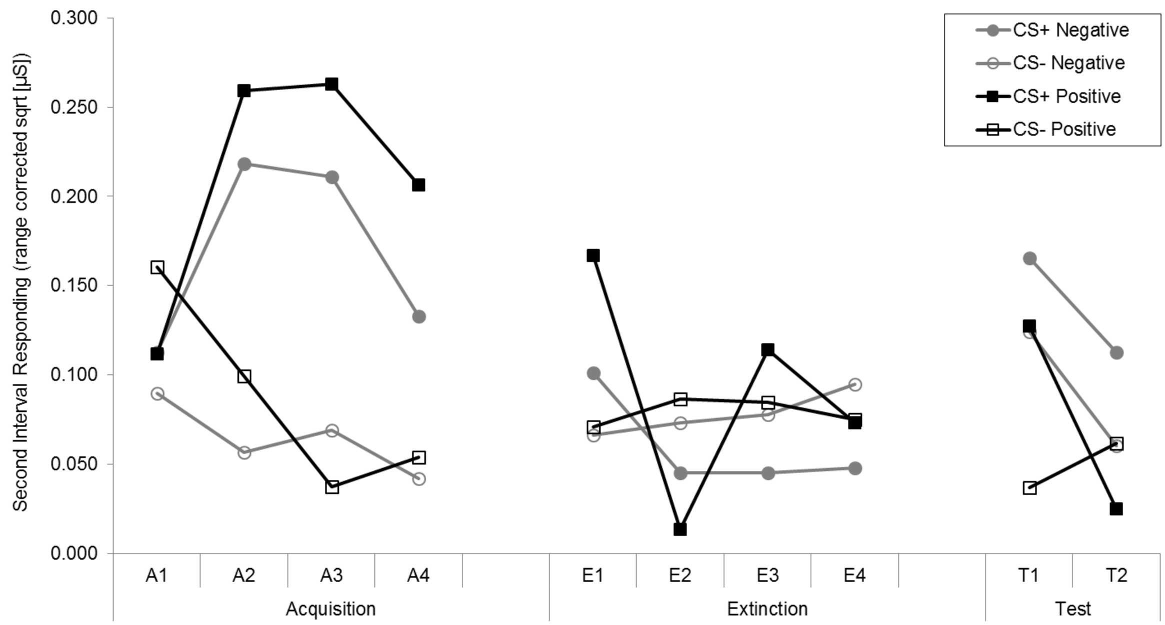

Figure 7. Electrodermal second interval responses recorded during acquisition, extinction, and test phases of Experiment 2.

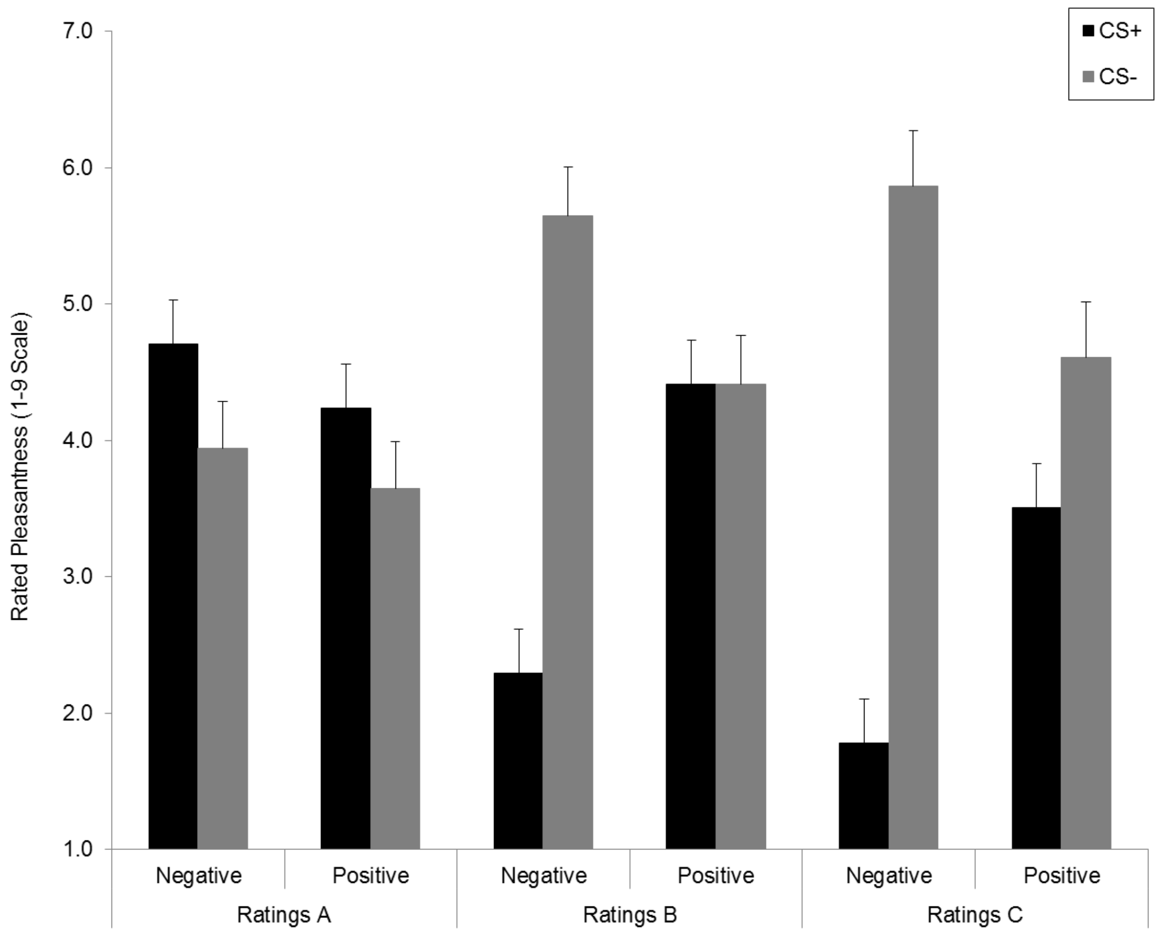

Figure 8. Conditional stimulus valence ratings taken before habituation (ratings A), after reinstatement (ratings B), and post-experimentally (ratings c) in Experiment 2. (Error bars indicate standard error of the mean; error bars can only be used to assess between group differences).

References

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1992). The content of learning in human evaluative conditioning: Acquired valence is sensitive to US-revaluation. *Learning and Motivation, 23,* 200-224. doi:10.1016/0023-9690(92)90018-H

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W.T., Dawson, M.E., & Filion, D. L. (2012). Publication recommendations for electrodermal measures. *Psychophysiology, 49*, 1017-1034. doi:10.1111/j.1469-8986.2012.01384.x

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

Craske, M. G. (1999). *Anxiety Disorders: Psychological Approaches to Theory and Treatment*: Westview Press.

M. G. Craske, D. Hermans & D. Vansteenwegen (Eds.). (2006). Fear and learning: From basic processes to clinical implications. Washington: APA Books.

Cuthbert, B. N., Bradley, M. M., & Lang, P. J. (1996). Probing picture perception: Activation and emotion. *Psychophysiology, 33*, 103-111. doi:10.1111/j.1469-8986.1996.tb02114.x

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.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*, 853-869. doi:10.1037/0033-2909.127.6.853

Denniston, J. C., Chang, R. C., & Miller, R. R. (2003). Massive extinction treatment attenuates the renewal effect. *Learning and Motivation, 34*, 68-86. doi:10.1016/S0023-9690(02)00508-8

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*, 549-554. doi:10.1101/lm.78004

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

Hamm, A. O., Vaitl, D., & Lang, P. J. (1989). Fear conditioning, meaning, and belongingness: A selective association analysis. *Journal of Abnormal Psychology, 98*, 395-406. doi:10.1037/0021-843X.98.4.395

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*, 533-551. doi:10.1016/j.brat.2004.03.013

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*, 390-421. doi:10.1037/a0018916

Laborda, M. A., & Miller, R. R. (2013). Preventing return of fear in an animal model of anxiety: additive effects of massive extinction and extinction in multiple contexts. *Behavior therapy, 44*, 249-261. doi:10.1016/j.beth.2012.11.001

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50*, 372-385. doi:10.1037/0003-066X.50.5.372

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

Lipp, O. V., & Purkis, H. M. (2006). The effects of assessment type on verbal ratings of conditional stimulus valence and contingency judgments: Implications for the extinction of evaluative learning. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 431-440. doi:10.1037/0097-7403.32.4.431

Lipp, O. V., Siddle, D.A.T., & Dall, P.J. (2003). The effects of unconditional stimulus valence and conditioning paradigm on verbal, skeleton-motor, and autonomic indices of human Pavlovian conditioning. *Learning and Motivation, 34*, 32-51. doi:10.1016/S0023-9690(02)00507-6

Luck, C. C., & Lipp, O. V. (2015a). 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

Luck, C. C., & Lipp, O. V. (2015b). To remove or not to remove? Removal of the unconditional stimulus electrode does not mediate instructed extinction effects. *Psychophysiology, 52*, 1248-1256. doi: 10.1111/psyp.12452

Luck, C. C., & Lipp, O. V. (2016a). Instructed extinction in human fear conditioning: History, recent developments, and future directions. *Australian Journal of Psychology, 68*, 209-227. doi:10.1111/ajpy.12135

Luck, C. C., & Lipp, O. V. (2016b). When orienting and anticipation dissociate — a case for scoring electrodermal responses in multiple latency windows in studies of human fear conditioning. *International Journal of Psychophysiology, 100*, 36-43. doi:10.1016/j.ijpsycho.2015.12.003

Luck, C. C., & Lipp, O. V. (2016a). Instructed extinction in human fear conditioning: history, recent developments and future directions. *Australian Journal of Psychology.*

Öhman, A. (1983). The orienting response during Pavlovian conditioning. In D. A. T. Siddle (Ed.), *Orienting and habituation: Perspectives in human research* (pp. 315-370). New York: Wiley.

Gast, A., & Rothermund, K. (2011). I like it because I said that I like it: Evaluative conditioning effects can be based on stimulus-response learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 466-476. doi:10.1037/a0023077

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242-249. doi:10.1016/j.psychres.2008.05.006

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. doi:10.1016/j.appet.2013.01.019

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear Extinction and Relapse: State of the Art. *Annual Review of Clinical Psychology, 9*, 215-248. doi:10.1146/annurev-clinpsy-050212-185542

Wright. S.P. (1992). Adjusted p-values for simultaneous inference. *Biometrics*, *48*, 1005-1013.: doi:10.2307/2532694

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 and Emotion, 29,* 654-667. .
doi:0.1080/02699931.2014.930421