

Startle Modulation and Explicit Valence Evaluations Dissociate During Backward Fear
Conditioning

Camilla C. Luck^{1,2} & Ottmar V. Lipp^{1,2}

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

²ARC-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: 8681

Abstract

Blink startle magnitude is linearly modulated by affect, such that relative to neutral stimuli, startle magnitude is inhibited during pleasant stimuli and potentiated during unpleasant stimuli. Andreatta, Mühlberger, Yarali, Gerber, and Pauli (2010) however, report a dissociation between startle modulation and explicit valence evaluations during backward conditioning, a procedure in which the unconditional stimulus (US) precedes the conditional stimulus (CS). Relative to controls, startles elicited during the CS were inhibited, suggesting that the CS had acquired positive valence, but participants still evaluated the CS as unpleasant after the experiment. In Experiment 1, we aimed to replicate this dissociation using a trial-by-trial measure of CS valence to measure startle modulation and CS valence simultaneously during forward and backward differential fear conditioning. In Experiment 2, we examined whether early and late portions of the CS could acquire differential valence by presenting startle probes at early and late probe positions during the CS. In both experiments, the dissociation between startle modulation and explicit valence evaluations in backward conditioning replicated, with CS+ evaluated as less pleasant than CS-, but startles elicited during CS+ inhibited relative to CS-. In Experiment 2, we provide preliminary evidence that this inhibition was present early, but not late, during the CS+. The results replicate the dissociation between implicit and explicit CS valence reported by Andreatta et al. (2010) using a trial-by-trial measure of valence. We also provide preliminary evidence that this dissociation may occur because the implicit and explicit measures are recorded at different times during the CS presentation.

Key words: blink startle potentiation, valence evaluations, backward conditioning, fear conditioning, emotion, fear potentiated startle.

During classical fear conditioning, a neutral conditional stimulus (CS) is repeatedly presented before an aversive unconditional stimulus (US; forward conditioning). The CS signals the aversive US, leading to the development of enhanced startle responses and negative valence (De Houwer, Thomas, & Baeyens, 2001; Lipp, 2006). Interestingly, simply reversing the temporal sequence of the CS and US, such that the US now precedes the CS (backward conditioning), leads to the development of positive implicit valence as indicated by diminished startle responding in humans (Andreatta, Muhlberger, Yarali, Gerber, & Pauli, 2010) and approach behavior in fruit-flies (Tanimoto, Heisenberg, & Gerber, 2004; for a review of this phenomenon see Gerber et al., 2014).

A CS becoming appetitive following an aversive US has been shown across a number of different species and paradigms and is believed to occur because the US offset elicits relief. Fruit-flies will avoid an odor associated with the onset of a shock but approach an odor associated with the offset of a shock (Tanimoto et al., 2004; Yarali et al., 2008). In rats and humans, if startle responding is assessed as an implicit measure of valence (see Lang, Bradley, & Cuthbert, 1990), startle responses elicited during a CS which signals the US are potentiated (indicating negative implicit valence), but startles elicited during a CS which follows the US are attenuated (indicating positive implicit valence) (see Andreatta et al., 2010; 2012; Mohammadi, Bergado-Acosta, & Fendt, 2014). This positive relief effect is in line with Solomon's (1980) opponent process theory, which suggests that an aversive US elicits two processes – a negative affective response upon US onset (*a*-process) and an opposing positive compensatory affective response shortly thereafter (*b*-process).

Interestingly however, a backwardly conditioned CS+ does not always acquire positive affective qualities across all response systems. Andreatta et al. (2010) reported that in human

participants, startles elicited during a backward CS+ were attenuated (positive implicit valence), but that explicit evaluations of the backward CS+ were negative. This dissociation, however, was not replicated by Andreatta et al. (2013) when the forward CS+ and the backward CS+ were associated with the same US, such that the US delivery was predictable (i.e. 8 s forward CS+ followed by a 200 ms US which after a 6 s delay was followed by the 8 s backward CS+). In this design, the backward CS+ developed positive explicit and implicit valence relative to the forward CS+, and did not differ from a CS- which had been presented alone. Participants were also not able to report the contingency between the backward CS+ and the US. Based on this, it seems likely that the backward CS+ was evaluated as pleasant, not because it elicited relief, but because it was perceived as a second CS-, or because conditioning to the forward CS+ overshadowed conditioning to the backward CS+.

The dissociation between explicit CS evaluations and startle modulation reported by Andreatta et al. (2010) is consistent with dual-process theories (see Bechara, 2005; Strack & Deutsch, 2004). These theories propose that there are two valence processing systems in the brain, an impulsive implicit system based on associative principles and a reflective explicit system which is more reliant on cognitive knowledge (Bechara, 2005; Strack & Deutsch, 2004). During backward conditioning, positive valence may be acquired via the impulsive processing system, while, negative valence is acquired via the reflective processing system. The notion that negative valence is acquired via the reflective system is consistent with evaluative conditioning studies using valenced pictures as USs. In this paradigm, the CS acquires the positive/negative valence of the US regardless of whether it is presented before, during, or after the CS (Mallan, Lipp, & Libera, 2008).

Dual process theories can explain the dissociation between implicit and explicit valence acquisition that was reported in backward conditioning by Andreatta et al. (2010), but the dissociation could also occur because of methodological differences in the measurement of implicit and explicit valence. Startle responses were measured throughout extinction training, but explicit CS valence evaluations were measured in a post-experimental assessment. Post-experimental measures do not always accurately capture CS valence during acquisition or extinction because they are sensitive to renewal effects (Bouton, 2002) and because participants tend to integrate valence across the different experimental contingencies when making post-experimental judgments (Lipp & Purkis, 2006). Measuring CS valence evaluations online, at the same time as the startle responses, would provide a more reliable index of CS valence and could help to identify whether the dissociation between implicit and explicit valence reported by Andreatta et al. (2010) reflects a true dissociation between different indices of stimulus valence.

We aimed to replicate the dissociation between startle modulation and explicit valence evaluation reported by Andreatta et al. (2010), using an online (trial-by-trial) assessment of explicit CS valence. In Experiment 1, we examined startle modulation and explicit CS valence during forward and backward conditioning, varying conditioning type within participants and order of conditioning between participants (reversal design). In Experiment 2, we examined startle modulation and explicit CS valence during forward and backward conditioning, varying conditioning type between participants. In Experiment 2, we presented startle probes in both early and late CS portions to examine whether early and late CS portions develop different valence. In both experiments, we included pre- and post- assessment of explicit CS valence to compare with the results of Andreatta et al. (2010). We also included an affective priming task (Fazio & Olsen, 2003) to examine whether the dissociation between implicit and explicit valence

uncovered in Andreatta et al. (2010) is specific to a physiological index of implicit valence, or, like in evaluative conditioning (Mallan et al., 2008) would also be present in a behavioral index of implicit valence.

Experiment 1

During forward conditioning, we hypothesized that startles elicited during the CS+ would be potentiated and that the CS+ would acquire negative explicit valence. During backward conditioning, we hypothesized that startles elicited during CS+ would be attenuated but that the CS+ would acquire negative explicit valence. Additionally, we expected the affective priming results to converge with the pattern of startle responding and the pre- and post- assessments of CS valence to converge with the online CS valence evaluations.

Method

Participants

Thirty-two undergraduate students (21 female) aged between 18 and 41 years ($M = 23.77$) volunteered participation in exchange for course credit or monetary compensation of AU\$15. Participants were randomly assigned to the backward first ($n = 16$) or forward first group ($n = 16$). One participant from the forward first group did not provide online CS valence evaluations and was excluded from the analyses.

Apparatus/Stimuli

The CSs were 600×450 pixel pictures of geometrical shapes (square, circle, triangle, diamond; black outlines on a white background). The pictures were presented on a 24 inch color LCD screen for 8 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 shapes used as CS+/CS-, and the two shapes used in the experiment (out of the possible four). The US was a 200 ms electrocutaneous stimulus, pulsed at 50 Hz and delivered by a Grass SD9 Stimulator to the participant's preferred forearm. Inter-trial intervals lasted 11 s, 13 s, or 15 s from CS offset to CS onset and were randomly varied throughout the experiment.

CS evaluations and an orbicularis oculi electromyogram (EMG) were recorded with a Biopac MP150 system, using acqKnowledge Version 4.1 at a sampling frequency of 1000 Hz. DMDX 5.3.4 software (Forster & Forster, 2003) was used to control the stimulus presentation and timing, record the pre- and post-experimental valence evaluations, and the reaction times and errors from the affective priming task. Trial-by-trial CS evaluations were recorded with a Biopac Variable Assessment Transducer (TSD115) with the anchors 0 (*very negative*) to 9 (*very positive*). Orbicularis oculi EMG was measured using three 4 mm Ag/AgCl electrodes, the first placed directly underneath the participant's left eye, the second below the corner of the left eye approximately 1 cm to the left of the first electrode, and the third (reference electrode) was placed in the middle of the participant's forehead. The electrodes were fitted with adhesive collars and filled with a standard electrode gel and impedance values were checked to ensure they were below 10 k Ω . Startle blinks were elicited with 105 dB bursts of white noise, lasting 50 ms with an instantaneous rise time. Startle bursts were generated by a custom built noise generator and presented through Sennheiser headphones. Startle probes were positioned either at 5 s or 7 s after CS onset during half of the CSs and at 6 s or 7 s before the next CS onset during half (22 total) of the inter-trial intervals. Orbicularis oculi EMG was recorded with a Biopac EMG100C amplifier (Gain: 5000; low pass 500 Hz, high pass 10 Hz).

Scoring and Response Definition

The raw EMG was passed through a 50 Hz notch filter and a second band pass filter (low pass 500 Hz, high pass 30 Hz) to reduce electro-magnetic noise and movement related artifacts, respectively. The filtered EMG was rectified and smoothed with a moving average based on 5 consecutive measurement points. Blink startle magnitude was defined as the maximum of the rectified and smoothed response curve occurring within 120 ms of the startle stimulus onset (Blumenthal et al., 2005). A trial was defined as a non-response trial if a response onset could not be visually identified within 20-60 ms after probe onset. A trial was defined as missing if a spontaneous or voluntary blink immediately preceded the startle probe onset or if the baseline EMG recorded 50 ms prior to probe onset was judged by visual inspection to be unstable. To reduce the impact of individual differences, raw magnitudes were transformed into T-scores using all startles measured during conditioning as the reference distribution. CS valence evaluations were scored as the largest positive or negative voltage deviation from a 1 s pre-CS baseline ('neutral' position) recorded during the 8 s CS presentation.

Procedure

Participants provided informed consent, washed underneath their left eye and on their forehead with a non-allergenic soap and were seated in a separate room adjacent to the control room. The EMG electrodes were attached underneath the left eye and on the forehead and the participants were presented with three startle bursts to habituate blink startle responding. The shock electrode was placed on the participants' preferred forearm and participants underwent a shock work-up procedure. Intensity was increased from 0 V in steps of 10 V until the participant reported feeling a sensation (usually described as a light tingle). The intensity was then increased in 5 V steps until the participant reported that the intensity was experienced as

‘unpleasant, but not painful’ and this intensity was used throughout the experiment. Participants then viewed the CS shapes and rated them on a 1 to 9 Likert scale (1= *unpleasant*, 9=*pleasant*) using the keyboard and completed an affective priming task, using the four possible CS shapes as primes. During affective priming, participants were asked to evaluate six pleasant (appealing, charming, desirable, favorable, nice, and enjoyable) and six unpleasant target words (annoying, disturbing, inferior, nasty, repulsive, and terrifying) as pleasant or unpleasant as quickly as possible while avoiding mistakes. On each trial, a shape prime was presented for 200 ms followed by a blank screen for 100 ms and the target word for 1s, or until a response was made. If the preceding CS prime is pleasant, participants should be faster to evaluate pleasant words (valence congruent) and slower to evaluate negative words (valence incongruent). If the preceding CS prime is unpleasant, participants should be slower to evaluate pleasant words (valence incongruent) and faster to evaluate negative words (valence congruent). 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 they would be presented with shapes, electrotactile stimuli, and noise bursts. They were asked to pay attention to the shapes and electrotactile stimuli, but to ignore the noise bursts. The participants were asked to operate the slider of the Biopac Variable Assessment Transducer with their preferred hand to indicate how pleasant/unpleasant they found each shape while it was on the screen, ensuring that the movement did not interfere with the physiological recordings and that the presence/absence of the US did not influence the evaluations. After making an evaluation, participants were asked to move the slider back to the neutral position before the next shape appeared. Participants were not informed about the CS-US contingencies.

The training procedure consisted of habituation and two conditioning phases (phase A and phase B). During habituation, in both groups, the CS+ and CS- were presented alone three times. During phase A, the forward first group received forward conditioning training, in which the CS+ was presented eight times with its offset coinciding with the onset of the US, and the CS- was presented eight times alone. The backward first group received backward conditioning training (8 CS+ and 8 CS-), in which the offset of the 200 ms US was followed by CS+ onset after a 100 ms delay, and the CS- was presented alone. During phase B (12 CS+ and 12 CS- each), the forward first group received backward conditioning training, whereas the backward first group received forward conditioning training. A 100 ms trace interval was used in backward conditioning to ensure that the sensation of the US did not overlap with the CS+, but was perceived at the onset (to be similar to the delay conditioning procedure used in the forward conditioning procedure)¹. Four additional CS+/CS- trials were added to phase B to ensure participants had sufficient opportunity to learn that the US timing had changed. The experimental design is displayed in Table 1. Startle probes were placed in 50% of the CS+/CS- presentations – habituation (2nd CS+/CS- presentation), phase A (2nd, 4th, 6th, and 8th CS+ presentations; 2nd, 4th, 6th, and 7th CS- presentations), and phase B (2nd, 4th, 6th, 8th, 10th and 12th CS+ presentations; 2nd, 4th, 6th, 7th, 10th, and 12th CS- presentations).² After the last conditioning trial, participants completed another rating task using the keyboard and a second priming task identical to the one completed before the experiment and were led into the experimenter room

¹ There is work suggesting that the relief effect starts within at least 3 s but may take 6 s to fully develop (see Andreatta, Mühlberger, & Pauli, 2016). We used a short interval because we believed it important not to confound the experiments by using two different conditioning procedures (delay in forward conditioning and (long) trace in backward conditioning). The startle probes were positioned 5 s and 7 s into the CS and should be adequately positioned to capture the relief effect.

² The startle probes were placed systematically throughout the CS+/CS- presentations, but as the trial sequence was pseudo-randomised, the probe position of startles during CSs was varied, and startles were also placed during the ITIs, it is unlikely that startle presentation was predictable.

for a post-experimental questionnaire. The questionnaire assessed contingency awareness by requiring participants to identify which shapes were presented during training, which shape was paired with the electrotactile stimulus, and the order in which the CS+ and the US were presented in phases A and B. Participants then rated the pleasantness of the electrotactile stimulus, the startle probes, and the CS shapes on a 7 point Likert scale (-3 [*very unpleasant*] to +3 [*very pleasant*]) before being debriefed and thanked.

Insert Table 1 about here

Statistical Analyses

Analysis of the evaluation times and error data from the affective priming tasks yielded no evidence of conditioning and therefore these results have not been reported but are available upon request from the corresponding author. All analyses were conducted with IBM SPSS Statistics 22 with a significance level of .05. Interactions have been followed-up with simple effect contrasts, and Pillai's trace statistics of the multivariate solution are reported.

Results

Preliminary Analyses

The means and standard deviations for the preliminary checks are presented in Table 2. A Pearson's chi-square test confirmed that the gender ratio did not differ between groups, $\chi^2(1) = 0.14, p = .710$. A series of independent samples t-tests revealed that the groups did not differ in age, startle magnitude during the inter-trial intervals, US intensity, or US valence, all $t < 1.08, p > .293$. The forward first group rated the startle probes as less pleasant than the backward first group, $t(30) = 2.37, p = .024$. Nine participants (forward first: 4, backward first: 5) could not

correctly report the experimental contingencies. When these participants are removed from the analyses the conclusions do not change and therefore data from the entire sample are reported.

Insert Table 2 about here

Habituation

Blink startle magnitudes and CS valence evaluations from habituation were subjected to separate 2 (Group: forward first, backward first) \times 2 (CS: CS+, CS-) factorial ANOVAs and the results are presented in Figures 1 and 2, respectively.

Blink Startle Magnitude. No main effects or interactions reached significance, all $F < 3.09$, $p > .089$, $\eta^2 < .100$.

Conditional Stimulus Valence. No main effects or interactions reached significance, all $F < 1.29$, $p > .267$, $\eta^2 < .043$.

Acquisition (Phase A and B)

Blink startle magnitudes and CS valence evaluations recorded during phases A and B were subjected to separate 2 (Group: forward first, backward first) \times 2 (CS: CS+, CS-) \times 2 (Phase: A, B) factorial ANOVAs and are presented in Figures 1 and 2, respectively.

Blink Startle Magnitude. A main effect of phase, $F(1, 29) = 14.64$, $p = .001$, $\eta^2 = .335$, was qualified by a CS \times Phase \times Group interaction, $F(1, 29) = 9.97$, $p = .004$, $\eta^2 = .256$. In the forward first group, startles elicited during CS+ and CS- did not differ during Phase A (forward conditioning), $F(1, 29) = 0.16$, $p = .691$, $\eta^2 = .006$, or Phase B (backward conditioning), $F(1, 29) = 1.67$, $p = .207$, $\eta^2 = .054$. In the backward first group, startles elicited during CS+ were smaller than startles elicited during CS- during Phase A (backward conditioning), $F(1, 29) =$

7.37, $p = .011$, $\eta^2 = .203$, but during Phase B (forward conditioning), startles elicited during CS+ were larger than startles elicited during CS-, $F(1, 29) = 8.54$, $p = .007$, $\eta^2 = .227$. To examine whether responding to CS+ and CS- differed between the groups, the analyses were re-run comparing group. Responding to CS+ did not differ between the groups during Phase A, $F(1, 29) = 0.63$, $p = .434$, $\eta^2 = .021$, but during Phase B, startles elicited during CS+ were larger in the backward first group than in the forward first group, $F(1, 29) = 5.51$, $p = .026$, $\eta^2 = .160$. Responding to the CS- did not differ between the groups during Phase A, $F(1, 29) = 2.91$, $p = .099$, $\eta^2 = .091$, or Phase B, $F(1, 29) = 1.29$, $p = .265$, $\eta^2 = .043$. The remaining main effects and interactions did not attain significance, all $F < 2.99$, $p > .094$, $\eta^2 < .094$.

Insert Figure 1 about here

Conditional Stimulus Valence. A marginal main effect of CS, $F(1, 29) = 3.91$, $p = .058$, $\eta^2 = .119$, revealed that across both phases and groups, CS+ was evaluated as marginally less pleasant than CS-. The remaining main effects and interactions did not attain significance, all $F < 2.43$, $p > .130$, $\eta^2 < .078$.

Insert Figure 2 about here

Pre and Post-experimental Pleasantness Ratings

The pleasantness ratings recorded before conditioning, after conditioning, and post-experimentally are presented in Figure 3 and were subjected to a 2 (Group: forward first, backward first) \times 2 (CS: CS+, CS-) \times 3 (Phase; before, after, post) factorial ANOVA. A main effect of CS, $F(1, 28) = 6.23$, $p = .019$, $\eta^2 = .182$, was moderated by a CS \times Phase interaction, $F(2, 27) = 5.85$, $p = .008$, $\eta^2 = .302$. Follow-up analyses revealed that before conditioning, ratings of CS+ and CS- did not differ, $F(1, 28) < 0.01$, $p = .973$, $\eta^2 < .001$, however after

conditioning, $F(1, 28) = 5.98, p = .021, \eta^2 = .176$, and post-experimentally, $F(1, 28) = 8.98, p = .006, \eta^2 = .243$, CS+ was rated as less pleasant than CS-. The remaining main effects and interactions did not attain significance, all $F < 2.68, p > .086, \eta^2 < .166$.

Insert Figure 3 about here

Discussion

In Experiment 1, we aimed to replicate the dissociation between startle modulation and explicit valence evaluations reported by Andreatta et al. (2010) using a trial-by-trial measure of CS valence. We examined forward and backward conditioning within groups, varying the order of conditioning type between groups. As predicted, during forward conditioning, startles elicited during CS+ were potentiated and during backward conditioning startles elicited during the CS+ were attenuated. Unexpectedly, however, although the overall pattern of results were similar in both groups differences were only significant in the backward first group. Replicating Andreatta et al. (2010), CS+ was evaluated as less pleasant than CS- in forward and backward conditioning, a difference which was marginal in the online ratings and significant in post-experimental ratings. This confirms that the dissociation between startle modulation and explicit valence evaluations in backward conditioning is reliable and evident regardless of the timing of CS valence measurement. Unexpectedly, however, differential CS evaluations were not evident in affective priming.

Although replicating past findings, the current results do not clarify why startle modulation and explicit valence evaluations dissociate. The dissociation could occur because implicit and explicit indices of valence selectively reflect activity of an impulsive and a reflective valence processing system, respectively (Bechara, 2005; Strack & Deutsch, 2004), but there are

also a number of alternative accounts. A long CS-US interval (8 s) was used in the current study and it is possible that different sections of the CS+ become associated with either US offset (pleasant relief) or US onset (unpleasant fear). The dissociation between startle modulation and CS valence ratings could occur because the measures are recorded during different sections of the 8 s CS. Alternatively, if participants realize that within the pseudorandom trial sequence, a CS+ presentation is more likely to follow a CS- presentation, they may allocate more attention to CS- as it is a better predictor of the next US during backward conditioning (see Wiens, Katkin, & Öhman, 2003, for a demonstration of trial order effects in pseudorandom trial sequences). The enhanced allocation of attention to CS- could enhance startle (see Lipp, Neumann, Pretorius, & McHugh, 2003, for an illustration of the effects of attention on blink startle modulation) providing an alternative interpretation for reduced startles during CS+ observed in backward conditioning.

Experiment 2

In Experiment 2, we examined whether early and late CS portions differ in valence during backward conditioning by presenting startle probes in both early and late sections of the CSs. We examined forward and backward conditioning between groups and hypothesized that in backward conditioning, startles elicited during the early sections of CS+ would be inhibited, while startles elicited during the late sections of CS+ would be potentiated. As in Experiment 1, we expected the CS+ to acquire negative explicit valence in both forward and backward conditioning. To assess whether the results may be driven by participants realizing that, in backward conditioning, the US is more likely to be presented after the CS- we included a manipulation check assessing participants' awareness of the trial order in the post-experimental questionnaire.

Method

Participants

Thirty-four undergraduate students (22 female) aged between 18 and 31 years ($M = 20.62$ years) volunteered participation in exchange for course credit or monetary compensation of AU\$15. Participants were randomly assigned to the forward conditioning group ($n = 17$) or the backward conditioning group ($n = 17$).

Apparatus/Stimuli

The conditional stimuli (CS) were 600×450 pixel pictures of geometrical shapes (square, diamond; black outlines on a white background). Counter-balancing was performed across participants, varying the nature of the first trial (CS+/CS-) and the shapes used as CS+/CS-. Two early (0.8s and 1.3s) and two late (5s, 7s) startle probe positions were used during the CSs and a probe was presented on each trial. The early probes were positioned to occur outside the pre-pulse inhibition time window of approximately 0.1s – 0.3s (Dawson, Schell, & Böhmelt, 1999; Neumann, Lipp, & Pretorius, 2004) and during time windows where emotional startle modulation has been reported in the picture viewing paradigm (Bradley, Cuthbert, & Lang, 1993). Startle probes were placed during half of the inter-trial intervals (17 total) either 6 s or 7 s before CS onset. As in Experiment 1 the inter-trial intervals lasted 11 s, 13 s, or 15 s, and were randomly varied throughout the experiment. The remainder of the apparatus and stimuli used were the same as for Experiment 1.

Scoring and Response Definition

Blink startle and CS valence evaluations were scored in the same manner as in Experiment 1.

Procedure

The conditioning procedure consisted of habituation, pre-acquisition, acquisition, and extinction phases. As in Experiment 1, three initial startle probes were given to habituate startle responses. During habituation, CS+ and CS- were presented alone four times, with each probe position utilized once. Before acquisition participants received one reinforced presentation of CS+ and one presentation of CS- alone (US presented 100 ms before CS+ onset or at CS+ offset for backward and forward groups, respectively). This additional phase was added to ensure that a trial in which the CS+ had been paired with the US had been presented before measures of startle and CS evaluation were obtained for analysis (the probe positions used in this phase were counter-balanced across participants). During acquisition, the forward conditioning group received eight presentations of CS+ with the offset of the CS+ coinciding with the onset of the US, and eight presentations of CS- alone; whereas the backward conditioning group received eight presentations of the 200 ms US followed by a 100 ms delay and the presentation of the CS+, and eight presentations of CS- alone. Each startle probe position was used once during the first four trials of acquisition for each CS and again during the last four trials of acquisition for each CS. During extinction, both groups received four unreinforced presentations of CS+ and CS-, and each probe position was used once. The experimental design for each group is displayed in Table 3. The post-experimental questionnaire included a question requiring participants to indicate whether they were able to predict when the next US would occur and if they had identified any patterns in the trial sequence. The ITIs, affective priming, the online CS valence evaluations, the pre-post valence evaluations, and the remainder of the procedure were the same as in Experiment 1.

Insert Table 3 about here

Statistical Analyses

As in Experiment 1, no evidence of conditioning was detected in affective priming and therefore these analyses have not been reported. The results are available upon request from the corresponding author. Statistical analyses were conducted as in Experiment 1.

Results

Preliminary Checks

The means, standard deviations, and comparisons for the preliminary checks are presented in Table 2. A Pearson's chi-square test confirmed that the gender ratio did not differ between groups, $\chi^2(1) = 2.06, p = .151$. A series of independent samples t-tests confirmed that groups did not differ in age, startle magnitude during the inter-trial intervals, US intensity, US valence, or rated startle valence, all $t < 1.59, p > .122$. Three participants (forward conditioning: 2, backward conditioning: 1) could not correctly report the experimental contingencies and eight participants (forward conditioning: 2, backward conditioning: 6) could verbalize the pattern in the pseudorandom trial sequence. When these participants were removed from the analyses the conclusions do not change and therefore results from the entire sample have been reported.

Main Analyses

Blink startle magnitudes recorded during habituation, acquisition and extinction were subjected to separate 2 (Group: forward, backward) \times 2 (CS: CS+, CS-) \times 2 (Probe position: early, late) factorial ANOVAs and are presented in Figure 4. The CS valence evaluations from the three phases were subjected to separate 2 (Group: forward, backward) \times 2 (CS: CS+, CS-) factorial ANOVAs and are shown in Figure 5.

Habituation.

Blink startle magnitude. No main effects or interactions reached significance, $F < 2.32$, $p > .138$, $\eta^2 < .068$.

Conditional stimulus valence. No main effects or interactions reached significance, all $F < 3.31$, $p > .078$, $\eta^2 < .094$.

Acquisition.

Blink startle magnitude. A CS \times Group interaction, $F(1, 31) = 16.75$, $p < .001$, $\eta^2 = .351$, revealed that, relative to CS-, startle during CS+ was potentiated in the forward conditioning group, $F(1, 31) = 7.43$, $p = .010$, $\eta^2 = .193$, but inhibited in the backward conditioning group, $F(1, 31) = 9.35$, $p = .005$, $\eta^2 = .232$. A comparison between the groups, revealed that startles elicited during the CS+ were larger in the forward conditioning group than in the backward conditioning group, $F(1, 31) = 18.75$, $p < .001$, $\eta^2 = .377$, but that startles elicited during CS- did not differ between the groups, $F(1, 31) = 1.84$, $p = .185$, $\eta^2 = .056$. A main effect of probe position, $F(1, 31) = 10.44$, $p = .003$, $\eta^2 = .252$, was qualified by a CS \times Probe Position interaction, $F(1, 31) = 16.31$, $p < .001$, $\eta^2 = .345$. Startle magnitude did not differ between early and late CS- probes, $F(1, 31) = 0.15$, $p = .704$, $\eta^2 = .005$, however, startle magnitude was smaller during early CS+ probes in comparison with late CS+ probes, $F(1, 31) = 34.76$, $p < .001$, $\eta^2 = .529$.

The CS \times Probe Position \times Group interaction, $F(1, 31) = 2.03$, $p = .164$, $\eta^2 = .062$, did not attain significance, however the pattern of results in Figure 4 suggests a difference between the groups. Due to our a-priori hypotheses (see Rosenthal & Rosnow, 1985) we performed follow-up analyses for this interaction, but they should be considered exploratory and interpreted

with care. In the forward conditioning group, startle magnitude did not differ between early CS+ and CS- probes, $F(1, 31) = 0.78, p = .385, \eta^2 = .024$, but startles elicited during CS+ were potentiated during the late probes, $F(1, 31) = 10.33, p = .003, \eta^2 = .250$. In the backward conditioning group, startles elicited during CS+ were inhibited during early probes, $F(1, 31) = 24.18, p < .001, \eta^2 = .438$, but did not differ between CS+ and CS- during late probes, $F(1, 31) = 0.02, p = .893, \eta^2 = .001$. To examine whether the groups differed in startle during CS+ and CS-, the follow-up analyses were also run comparing across groups. Startles elicited during both early, $F(1, 31) = 9.69, p = .004, \eta^2 = .238$, and late, $F(1, 31) = 10.90, p = .002, \eta^2 = .260$, CS+ probes were larger in the forward conditioning group than in the backward conditioning group. Startles elicited during early CS- probes were smaller in the backward conditioning group in comparison with the forward conditioning group, $F(1, 31) = 4.70, p = .038, \eta^2 = .132$, but did not differ between groups during late CS- probes, $F(1, 31) = 0.25, p = .620, \eta^2 = .008$. The remaining main effects and interactions did not attain significance, all $F < 3.33, p > .077, \eta^2 < .098$.

Insert Figure 4 about here

Conditional stimulus valence. Across groups, CS+ was rated as less pleasant than CS-, $F(1, 32) = 40.61, p < .001, \eta^2 = .559$. The remaining main effects and interactions did not attain significance, all $F < 0.39, p > .544, \eta^2 < .013$.

Insert Figure 5 about here

Extinction.

Blink startle magnitude. No main effects or interactions reached significance, all $F < 2.09, p > .158, \eta^2 < .062$.

Conditional stimulus valence. Across groups, CS+ was evaluated as less pleasant than CS-, $F(1, 32) = 39.44, p < .001, \eta^2 = .552$. The remaining main effects and interactions did not attain significance, all $F < 1.16, p > .290, \eta^2 < .036$.

Pre and Post-experimental Pleasantness Ratings

Main effects of CS, $F(1, 32) = 28.36, p < .001, \eta^2 = .470$, and ratings, $F(2, 31) = 11.56, p < .001, \eta^2 = .427$, were qualified by a CS \times Ratings interaction, $F(2, 31) = 39.86, p < .001, \eta^2 = .720$. Follow-up analyses revealed that ratings of CS+ and CS- did not differ before conditioning, $F(1, 32) = 2.24, p = .145, \eta^2 = .065$, but after conditioning, $F(1, 32) = 37.98, p < .001, \eta^2 = .543$, and post-experimentally, $F(1, 32) = 48.73, p < .001, \eta^2 = .604$, CS+ was rated as less pleasant than CS-. The remaining main effects and interactions did not attain significance, all $F < 2.21, p > .127, \eta^2 < .125$.

Insert Figure 6 about here

Discussion

In Experiment 2 we investigated whether the early and late CS sections could acquire different emotional value, a difference which could potentially account for the dissociation between startle modulation and explicit valence evaluations observed during backward fear conditioning. Startle modulation was assessed at early and late CS probe positions throughout both forward and backward conditioning. Replicating prior results, startle was inhibited during CS+ in backward conditioning and potentiated during CS+ in forward conditioning, but the CS+ acquired negative explicit valence during both forward and backward conditioning. An exploratory follow-up analysis suggested that startle was inhibited during CS+ in backward conditioning at early, but not late, probe positions, but that during forward conditioning, startle

was potentiated during CS+ at late, but not early probe positions. As these findings are exploratory, however, they should be interpreted with care and will require exploration in future studies. When we removed participants who were able to report that a pseudorandom trial sequence was used the results did not change, suggesting that startle inhibition during CS+ in backward conditioning was not driven by participants paying more attention to CS-. Similar to Experiment 1, we did not find evidence for affective priming in Experiment 2. It is possible that the addition of the startle probes, which can be considered a mild US by themselves and were presented in both CS+ and CS-, reduced the differential valence between the CS+ and CS-. Another possibility is that, as the priming task was taken after extinction training, conditioning effects were not obtained because differential implicit valence had extinguished. This interpretation would be consistent with the startle results in Experiment 2, but could not explain why differential priming scores were not obtained during Experiment 1, in which a reversal design was used.

General Discussion

Across two experiments, we examined the dissociation between startle modulation and explicit CS valence evaluations during backward conditioning reported by Andreatta et al. (2010). In Experiment 1 we aimed to replicate this dissociation using a within-participant design that employed a trial-by-trial measure of CS valence, permitting the concurrent assessment of CS evaluations and blink startle modulation. In Experiment 2 we examined the mechanisms underlying the dissociation by mapping the time course of startle modulation during the CSs. Consistent with results from evaluative conditioning (Mallan et al., 2008), participants evaluated CS+ as negative during both forward and backward conditioning. Replicating Andreatta et al. (2010), startles elicited during CS+ were potentiated during forward conditioning but attenuated

during backward conditioning in Experiment 1, an effect which unexpectedly was significant only in the backward first conditioning group. In Experiment 2, we replicated this pattern in a between-participant design and extended this by providing preliminary evidence that startle inhibition is most prominent early during the CS+ in backward conditioning, while startle potentiation is most prominent late during the CS+ in forward conditioning. These results suggest that in backward conditioning, the initial portion of the CS+ acquires positive valence, an effect that diminishes towards the end of the CS+ presentation, but this finding will require further exploration and should be interpreted with care as the findings are based on the exploratory follow-up of a non-significant interaction. Interestingly, the finding that startle was potentiated late, but not early, during CS+ in forward conditioning, replicates Weike, Schupp, and Hamm's (2008) finding that conditioned startle discrimination during forward conditioning is specific to late portions of the CS+.

If explicit and implicit valence indices are measured during different portions of the CSs then the dissociation between startle modulation and explicit valence evaluations could be explained by these different portions acquiring different valence because they are associated with either US offset or US onset, but some caveats should be considered. The average response latency for the provision of CS evaluations was about 2 s (Experiment 1: $M = 1.83$, $SD = 0.79$; Experiment 2: $M = 2.02$, $SD = 0.89$). If the positive affective response dissipates within this time the CS+ may be evaluated as negative because the evaluations are taken after this initial positive response. It is not clear, however, how long it takes for the relief effect to dissipate. There is evidence to suggest that it may only start approximately 3 s after the offset of the US and require 6 s to develop fully (see Andreatta et al., 2016; but also see Klumpers, Heitland, Oostin, Kenemans, & Baas, 2012 and Klumpers et al., 2010 for the time course of defensive responding

after threat offset). The observation of reduced startle early during CS+ in backward conditioning suggests that the relief effect was present in Experiment 2 within the first 1.5 s of US offset. It is possible that timing of the relief effect is subject to unknown moderators such as different design parameters. More work will be required to follow-up this difference and to examine the time course of startle modulation during a wider variety of time points during backward conditioning.

In Experiment 2, we found no difference in startle modulation during late (5 s and 7 s) CS probe positions in the backward conditioning group, but such differential startle modulation was present in Experiment 1. It is not clear why startle magnitude was still inhibited at the late probe positions during the CS+ in Experiment 1, but not in Experiment 2. The experiments differed in a number of methodological aspects that could account for the results, including overall probe density and the range of probe positions used. A comparison across experiments revealed that the extent of differential startle modulation seen in backward conditioning was considerably larger at the early probe positions in Experiment 2 than at the late probe positions of Experiment 1. Thus, the startle inhibition detected at the late CS+ probes in Experiment 1 could be the remnant of a bigger difference that was present early during the CS+. A startle probe was positioned in every trial during Experiment 2 and the higher density of probes (which can be considered a mild US by itself) could have dampened the positive valence of the CS+ during backward conditioning. This would have been less pronounced in Experiment 1 because only half of the CSs were probed.

It is also not clear why the CS+ was evaluated as negative during backward conditioning, in absence of significant startle potentiation. In Experiment 2, startle magnitude was numerically larger late during the CS+, but this difference was small and not significant. It is possible that

explicit CS evaluations provide a less differentiated measure of CS valence that is mainly affected by the pairing of CS and US whereas startle provides the opportunity to track subtle changes in the time course of CS valence that occur during the course of a single CS presentation, be it in forward or backward conditioning.

In Experiment 2, we did not find evidence for the notion that startle inhibition during CS+ in backward conditioning occurs because participants allocate more attention to the CS-. Enhanced allocation of attention can potentiate startle and may occur when participants learn that the presentation of a US is more likely after a CS- than after a CS+. This account, which is a variant of the threat-proximity hypothesis (Fanselow, 1994), seems unlikely for a number of reasons. First, it would predict that the difference in startle magnitude between CS- and CS+ should be largest towards CS offset, i.e., most prominent in startles measured at the late probe positions, but no such difference was observed during backward conditioning in Experiment 2. Moreover, such a difference should be evident in inter-trial interval startles that fell immediately before a CS+ or before a CS- in backward conditioning, but comparing the size of these responses did not provide any evidence in support of this proposal. Furthermore, removing participants who could verbalize that a pseudorandom trial sequence was used did not change the pattern of results. Participants' ability to verbalize the contingency was assessed after the experiment and it is possible that participants could predict the US during training but were not confident to verbalize this post-experimentally. Another potential explanation for the finding of smaller startles during CS+ in backward conditioning is that the presentation of the US itself inhibits subsequent startle responses. The ISI used in the current study was significantly shorter than the 6 s trace interval used by Andreatta et al. (2010), but this inhibition would need to still be present 1.5 s after the US offset and as inhibition of startle during CS+ in backward

conditioning was also evident at startle probes presented 5-7 s after CS onset in Experiment 1, an US interference account seems unlikely.

Across two experiments we were able to replicate the dissociation between startle modulation and explicit CS valence evaluations during backward conditioning. In both Experiments we examined CS valence concurrently with startle modulation and provided evidence that the dissociation between startle modulation and explicit CS valence is present during acquisition and not just post-experimentally. In Experiment 2, we provided some evidence that this dissociation could occur because the early, but not the late, CS+ portions become pleasant during backward conditioning. The results could suggest that backward conditioning may involve the acquisition of both positive and negative valence at different time points during the CS, but will require further exploration and the examination of a wider variety of startle probe positions.

Table 1.

Experimental design of Experiment 1

	Habituation	Phase A	Phase B
Forward First	3 CS+ (unreinforced)	8 CS+-US (forward)	12 US-CS+ (backward)
	3 CS-	8 CS-	12 CS-
Backward First	3 CS+ (unreinforced)	8 US-CS+ (backward)	12 CS+-US (forward)
	3 CS-	8 CS-	12 CS-

Table 3.

Experimental design of Experiment 2

	Habituation	Pre-Acquisition	Acquisition	Extinction
Forward Conditioning	4 CS+ (unreinforced)	1 +-US (forward)	8 CS+-US (forward)	4 CS+ (unreinforced)
	4 CS-	1 CS-	8 CS-	4 CS-
Backward Conditioning	4 CS+ (unreinforced)	1 US-CS+ (backward)	8 US-CS+ (backward)	4 CS+ (unreinforced)
	4 CS-	1 CS-	8 CS-	4 CS-

Table 2.

Means and Standard Deviations for the Preliminary Analyses Variables

	Experiment 1			Experiment 2		
	Forward First	Backward First	Comparison	Forward	Backward	Comparison
Gender Ratio (male:female)	5:11	6:10	$\chi^2(1) = 1.39, p = .710.$	8:9	4:13	$\chi^2(1) = 2.06, p = .151.$
Age	23.88 (4.56)	23.67 (5.92)	$t(29) = 0.11, p = .913$	21.24 (4.22)	20.00 (2.55)	$t(32) = 1.03, p = .309$
US Level (V)	31.44 (5.74)	31.38 (11.55)	$t(30) = 0.02, p = .985$	51.29 (16.89)	48.47 (10.86)	$t(32) = 0.58, p = .566$
US Valence	-1.63 (0.72)	-1.63 (1.20)	$t(30) < .001, p > .999$	-1.88 (0.60)	-1.74 (0.90)	$t(32) = -0.56, p = .580$
Startle Valence	-1.75 (0.77)	-0.81 (1.38)	$t(30) = -2.37, p = .024$	-1.35 (0.86)	-1.35 (0.61)	$t(32) < .001, p > .999$
Startle Magnitude in ITI	50.78 (1.33)	50.25 (1.48)	$t(30) = 1.07, p = .294$	48.78 (2.07)	49.92 (2.10)	$t(32) = -1.58, p = .123$

Figure Headings

Figure 1. Mean blink startle magnitudes recorded during habituation, phase A, and phase B of Experiment 1. (Error bars indicate standard errors of the mean; ** $p < .01$; * $p < .05$).

Figure 2. Mean conditional stimulus valence evaluations recorded during habituation, phase A, and phase B of Experiment 1. Across both conditioning phases (phase A and B), CS+ was evaluated as marginally less pleasant than CS-. (Error bars indicate standard errors of the mean).

Figure 3. Mean conditional stimulus valence ratings measured before habituation, after phase B, and post-experimentally in Experiment 1. CS+ was evaluated as less pleasant than CS- after phase B and post-experimentally. (Error bars indicate standard error of the mean; ** $p < .01$; * $p < .05$).

Figure 4. Mean blink startle magnitudes elicited at early and late probe positions during habituation, acquisition, and extinction phases of Experiment 2. (Error bars indicate standard error of the mean; *** $p < .001$; ** $p < .01$; * $p < .05$).

Figure 5. Mean conditional stimulus valence evaluations recorded during habituation, acquisition, and extinction phases of Experiment 2. Evaluations did not differ during habituation, but were less positive for CS+ than for CS- during acquisition and extinction in both groups. (Error bars indicate standard error of the mean).

Figure 6. Mean conditional stimulus valence ratings taken before habituation, after extinction, and post-experimentally in Experiment 2. (Error bars indicate standard error of the mean; ** $p < .01$; * $p < .05$).

References

- Andreatta, M., Mühlberger, A., & Pauli, P. (2016). When does pleasure start after the end of pain? The time course of relief. *Journal of Comparative Neurology*, *524*, 1653-1667. doi:10.1002/cne.23872
- Andreatta, M., Fendt, M., Mühlberger, A., Wieser, M., Imobersteg, S., Yarali, A., Gerber, B., & Pauli, P. (2012). Onset and offset of aversive events establish distinct memories requiring fear and reward networks. *Learning and Memory*, *19*, 518-526. doi:10.1101/lm.026864.112
- Andreatta, M., Mühlberger, A., Glotzbach-Schoon, E., & Pauli, P. (2013). Pain predictability reverses valence ratings of a relief-associated stimulus. *Front. Syst. Neurosci*, *7*, 10.3389. doi:10.3389/fnsys.2013.00053
- Andreatta, M., Mühlberger, A., Yarali, A., Gerber, B., & Pauli, P. (2010). A rift between implicit and explicit conditioned valence in human pain relief learning. *Proceedings of the Royal Society of London B: Biological Sciences*. doi:10.1098/rspb.2010.0103
- Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nature Neuroscience*, *8*, 1458-1463. doi:10.1038/nn1584
- 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*, 1-15. doi: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*, 976-986. doi:10.1016/S0006-3223(02)01546-9

- Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1993). Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology*, *30*, 541-545. doi:10.1111/j.1469-8986.1993.tb02079.x
- Dawson, M.E., Schell, A.M., & Böhmelt, A.H. (Eds.). (1999). Startle Modification: Implications for neuroscience, cognitive science, and clinical science. New York: 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
- Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review*, *1*, 429-438. doi:10.3758/bf03210947
- Fazio, R. H., & Olsen, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297-327. doi:10.1146/annurev.psych.54.101601.145225
- 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
- Gerber, B., Yarali, A., Diegelmann, S., Wotjak, C. T., Pauli, P., & Fendt, M. (2014). Pain-relief learning in flies, rats, and man: basic research and applied perspectives. *Learning & Memory*, *21*, 232-252. doi:10.1101/lm.032995.113
- Klumpers, F., Heitland, I., Oosting, R. S., Kenemans, J. L., & Baas, J. M. P. (2012). Genetic variation in serotonin transporter function affects human fear expression indexed by fear-

potentiated startle. *Biological psychology*, 89, 277-282.

doi:10.1016/j.biopsycho.2011.10.018

Klumpers, F., Raemaekers, M. A. H. L., Ruigrok, A. N. V., Hermans, E. J., Kenemans, J. L., & Baas, J. M. P. (2010). Prefrontal Mechanisms of Fear Reduction After Threat Offset. *Biological Psychiatry*, 68, 1031-1038. doi:10.1016/j.biopsych.2010.09.006

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological review*, 97, 377. doi:10.1037/0033-295X.97.3.377

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., Neumann, D. L., Pretorius, N. R., & McHugh, M. J. (2003). Attentional blink modulation during sustained and after discrete lead stimuli presented in three sensory modalities. *Psychophysiology*, 40, 285-290. doi:10.1111/1469-8986.00030

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

Mallan, K. M., Lipp, O. V., & Libera, M. (2008). Affect, attention, or anticipatory arousal? Human blink startle modulation in forward and backward affective conditioning. *International Journal of Psychophysiology*, 69, 9-17. doi:10.1016/j.ijpsycho.2008.02.005

Mohammadi, M., Bergado-Acosta, J., & Fendt, M. (2014). Relief learning is distinguished from safety learning by the requirement of the nucleus accumbens. *Behavioural Brain Research*, 272, 40-45. doi:10.1016/j.bbr.2014.06.053

Neumann, D. L., Lipp, O. V., & Pretorius, N. R. (2004). The effects of lead stimulus and reflex stimulus modality on modulation of the blink reflex at very short, short, and long lead intervals. *Perception & Psychophysics*, 66, 141-151. doi:10.3758/bf03194868

R. Rosenthal & R. Rosnow (Eds.), (1985). Contrast analysis: Focused comparisons in the analysis of variance. U.S.A: Cambridge University Press.

Solomon, R. L. (1980). The opponent-process theory of acquired motivation: The costs of pleasure and the benefits of pain. *American Psychologist*, 35, 691-712. doi:10.1037/0003-066X.35.8.691

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and social psychology review*, 8, 220-247.

doi:10.1207/s15327957pspr0803_1

Tanimoto, H., Heisenberg, M., & Gerber, B. (2004). Experimental psychology: event timing turns punishment to reward. *Nature*, 430, 983-983. doi:10.1038/430983a

Weike, A. I., Schupp, H. T., & Hamm, A. O. (2008). In dubio pro defensio: Initial activation of conditioned fear is not cue specific. *Behavioral Neuroscience*, 122, 685-696.

doi:10.1037/0735-7044.122.3.685

Wiens, S., Katkin, E. S., & Öhman, A. (2003). Effects of trial order and differential conditioning on acquisition of differential shock expectancy and skin conductance conditioning to

masked stimuli. *Psychophysiology*, *40*, 989-997. doi:10.1111/1469-8986.00117

Yarali, A., Niewalda, T., Chen, Y., Tanimoto, H., Duernagel, S., & Gerber, B. (2008). 'Pain relief' learning in fruit flies. *Animal Behaviour*, *76*, 1173-1185.

doi:10.1016/j.anbehav.2008.05.025