Relapse of Evaluative Learning – Evidence for Reinstatement, Renewal, but not Spontaneous Recovery, of Extinguished Evaluative Learning in a Picture-Picture Evaluative Conditioning Paradigm

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

[1] School of Psychology, Curtin University, Australia

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

Author Notes

Address for correspondence:

Camilla C. Luck, School of Psychology, Curtin University, GPO Box U1987 Perth WA 6845, Australia. E-mail: c.luck@curtin.edu.au

Data from Experiments 2, 3b, 4, 5, and 6 can be found at https://osf.io/hvak4/

Acknowledgements:

This work was supported by grants number DP180111869 and SR120300015 from the Australian Research Council.

Word count: 25434

Abstract

In evaluative conditioning, if one shape (conditional stimulus; CSp) is paired with pleasant unconditional stimulus (US) images and another (CSu) is paired with unpleasant US images differential CS valence and US expectancy develops, such that participants evaluate the CSp as more pleasant and more predictive of pleasant images than the CSu. This conditional CS valence and US expectancy can be reduced in an extinction procedure in which the CSs are repeatedly presented alone. We investigated whether evaluative and expectancy learning is subject to relapse (spontaneous recovery, reinstatement, and renewal) after extinction in a picture-picture evaluative conditioning paradigm. In Stream 1, after acquisition and extinction, the spontaneous recovery test was completed after a delay. During the spontaneous recovery test, conditional expectancy learning, but not conditional evaluative learning, returned. In Stream 2, the US pictures were presented in a random stream after extinction (reinstatement manipulation) which led to the return of conditional evaluative and expectancy learning. In Stream 3, after acquisition training in context A and extinction training in context B, conditional expectancy and evaluative learning returned when participants completed the renewal test in the acquisition context (context A; ABA renewal). Overall, the results suggest that conditional evaluative learning is subject to reinstatement and renewal, but not to spontaneous recovery, in a picture-picture evaluative conditioning paradigm.

Key words: evaluative conditioning; valence; extinction; relapse; reinstatement; ABA renewal; spontaneous recovery.

Everyday decisions, like the brand of coffee we drink, the car we drive, and who we choose to socialize with are influenced by a basic learning process known as evaluative conditioning. Evaluative conditioning – the learning of ‘likes’ and ‘dislikes’ – occurs when a neutral conditional stimulus (CS) becomes pleasant or unpleasant (i.e. acquires positive or negative valence) after repeated pairings with a positive or negative unconditional stimulus (US; for a review see De Houwer, Thomas, & Baeyens, 2001). For example, an unfamiliar person may be evaluated as more pleasant because they are associated with a good friend; or a new brand of coffee seems more favorable if it is endorsed by a liked celebrity. Evaluative conditioning is robust and occurs with a wide range of different stimuli, including pictures, people, sounds, and brands (see Hofmann, De Houwer, Perugini, Baeyens & Crombez, 2010). *Procedurally*, evaluative conditioning is a type of Pavlovian conditioning as the pairing of the CS and the US results in the CS eliciting a response that it did not elicit before conditioning (i.e. a conditional response). There is considerable debate, however, regarding whether the *mechanisms* underlying evaluative conditioning differ from those underling other types of Pavlovian conditioning.

Traditional Pavlovian conditioning procedures, such as fear conditioning and preparatory conditioning, are based on a signal learning system, where the presentation of the CS generates an expectation that the US will occur. Evaluative conditioning, on the other hand, has been proposed to be based on a referential learning system, where the CS activates a representation of the US, without generating an expectation that it will occur (known as the referential account; Baeyens, Eelen, Crombez, and Van den Bergh, 1992; see De Houwer et al., 2010 for a review). For example, the smell of sunscreen may make you think of the beach and activate positive feelings, without making you expect that the beach will appear. The function of an expectancy based system is to predict the occurrence of the US and therefore a reliable contingency between the CS and the US is essential. A referential system, on the other hand, is said to be independent of expectancy and only sensitive to co-occurrence (contiguity) between the CS and the US.

The referential account has been tested extensively by examining whether evaluative conditioning responds to experimental manipulations in a different manner to standard Pavlovian conditioning. Some experiments find a dissociation between evaluative conditioning and other Pavlovian conditioning paradigms (e.g. Hermans, Crombes, Vansteenwegen, Baeyens, & Eelen, 2002; Lipp, Mallan, Libera, & Tan, 2010; Luck & Lipp, 2015), but others find similar results in both paradigms (e.g. Lipp & Purkis, 2005). One of the most debated issues is whether evaluative conditioning is subject to extinction. As contingency and expectation are important for a signal learning system, presenting the CS alone during extinction should result in a decrease, and eventual elimination, of the conditional response. A referential learning system, on the other hand, should not respond to extinction training because the CS should not trigger an expectation that the US will occur, and therefore, no expectancy violation should take place during extinction. The referential model makes clear predictions regarding the influence of extinction on evaluative conditioning, but the pattern of results in the literature is not clear. Many early evaluative conditioning studies report that extinction does not reduce conditional valence (Baeyens, Crombez, Van den Bergh, & Elen, 1998; Baeyens, Díaz, & Ruiz, 2005), but a subsequent meta-analysis confirmed that, overall, conditional valence was reduced after extinction (Hoffman, De Houwer, Perugini, Baeyens & Crombez, 2010). This led to the widespread view that evaluative conditioning does extinguish, but at a slower rate to signal based Pavlovian conditioning. Debates still continue, however, with some studies reporting slower, but complete, extinction (Luck & Lipp, 2015), and still others reporting that conditional valence fails to extinguish (Engelhard, Leer, Lange, & Olatunji, 2014). These discrepancies likely suggest that unknown moderators are at play and that the role of extinction on evaluative conditioning is complex – leaving debates regarding whether a referential learning system governs evaluative learning unsettled.

Extinction learning is a necessary requirement for examining whether conditional responding can return after extinction, and therefore, debates regarding whether evaluative conditioning is subject to extinction may have detracted from a very important question – does evaluative conditioning relapse? Examining relapse of evaluative conditioning is important from a theoretical perspective as it provides another useful tool to examine whether the mechanisms underlying evaluative and expectancy learning differ. From a clinical perspective, examining relapse in evaluative conditioning may also provide key insights into the mechanisms underlying *fear* relapse and how to reduce it. During fear conditioning, a neutral, or innocuous stimulus (conditional stimulus; e.g. a picture of a shape) is followed by an aversive stimulus (unconditional stimulus; e.g. an electric shock). After repeated pairings, the CS begins to elicit conditional fear responding on its own (commonly indexed with measures of physiology). This conditional fear responding can be reduced with an extinction procedure in which the CS is presented in the absence of the US (Lipp, 2006), but extinction does not erase the original CS-US association, instead substantial evidence suggests that it creates a CS-noUS association that suppresses the original learning (see Bouton, 2002 for a discussion of this theory). If the CS-noUS association weakens, or the CS-US association is triggered, conditional responding will reemerge.

During fear conditioning, the CS acquires signal value and becomes a valid predictor of the US (expectancy conditioning), but it also acquires negative valence. This negative valence – the evaluative conditioning component – is a predictor of fear relapse rates (Dirkx, Hermans, Vansteenwegen & Baeyens, 2004; Zbozinek, Hermans, Prenoveau, Liao & Craske, 2015), leading to a new line of research aimed at reducing negative CS valence in order to reduce relapse (e.g. Dour, Brown & Craske, 2016; Luck & Lipp, 2018). If negative valence is targeted and reduced during fear extinction (or during exposure treatment for an anxiety disorder), it is important to know whether this reduction is likely to remain stable over time. If evaluative conditioning *is* subject to relapse, this relapse may mediate the return of fear seen in clinical settings.

Relapse of human fear learning has been researched using three different manipulations (for a review see Vervliet et al., 2013). Spontaneous recovery is observed when extinguished conditional responding returns after the mere passage of time (e.g. a delayed test after extinction). Reinstatement occurs when extinguished conditional responding returns after unsignalled presentations of the US; and renewal is observed when extinguished conditional responding returns to a CS that is presented in a context that differs from the extinction context. For example, if acquisition occurs in context A and extinction in context B, conditional responding may reoccur if the CS is presented in context A again (ABA renewal) or in a new context (context C; ABC renewal).

To date, the spontaneous recovery, reinstatement, and renewal of extinguished evaluative learning have not been demonstrated. Gawronski, Rydell, Vervliet, and De Houwer (2010) have demonstrated ABA and ABC renewal of evaluative responses after a *counterconditioning* procedure. In this design, an initial evaluative response was trained in context A, such that the target individual Bob was paired with either pleasant or unpleasant statements against one background screen color (i.e. blue screen). During phase 2, the background screen color changed (i.e. yellow screen; Context B) and Bob was paired with statements of the opposite valence.

Implicit evaluative responding (measured with an affect misattribution procedure; See Payne et al., 2005) acquired during context A returned when the CS was tested in context A again (ABA renewal) or in a novel context C (ABC renewal). Although these interesting findings demonstrate that implicit evaluative responses are modulated by context in evaluative learning, it is not clear whether renewal would occur in the same manner if extinction was used in phase 2 in place of counterconditioning (as in traditional renewal research). Aust, Haaf, and Stahl (2019) have shown that context modulates evaluative learning in their work examining dissociations between evaluative and expectancy learning using an extinction design. In this research, acquisition training took place on one background screen color (context A), extinction on a second color (context B), and then the participants were asked to evaluate CS valence on a third screen color (context C). After this novel context test, participants were asked to retrospectively rate the valence of the CSs (i.e. how pleasant did you find this CS during the first/second half?). Conditional valence was larger when participants were tested in the acquisition context and the novel context compared to the extinction context. These results suggest that context modulates evaluative responding and are consistent with ABC renewal but as the new context assessment took place *before* the retrospective extinction test, it is not clear whether these results reflect relapse of evaluative learning. To provide an unequivocal demonstration of relapse the acquisition, extinction, and the relapse tests need to occur at the end of the respective learning phase.

We examined whether evaluative conditioning would relapse after spontaneous recovery, reinstatement, and renewal manipulations in a picture-picture evaluative conditioning paradigm. We aimed to provide a comprehensive proof of concept and therefore we conducted 3 streams of research, comprising 7 Experiments in total and using between-participant and within-participant experimental designs. We examined each relapse induction in both designs, using neutral shapes as CSs and positive, negative, and neutral pictures as USs. In the between-participant design, the positive conditioning group was trained with pairings of one CS (valenced CS: CSv) with positive pictures and of a second CS (neutral CS: CSn) with neutral pictures. In the negative conditioning group, the valenced CS (CSv) was paired with unpleasant pictures. The negative conditioning group provides a comparison with differential fear conditioning, where one CS (CS+) is paired with an aversive US, and another (CS-) is presented alone. Including a positive conditioning group allowed us to examine whether acquisition, extinction, and relapse of positive valence differs from that of negative valence. The within-participant design involved the pairing of one CS (positive CS: CSp) with pleasant images and another CS (unpleasant CS: CSu) with unpleasant images. As both CSs are paired with valenced USs, this design may provide a stronger test of conditioning and relapse. Measures of US expectancy and CS valence were obtained throughout acquisition and extinction, and immediately after the relapse induction. Including a measure of US expectancy allowed us to determine whether the relapse induction was having the expected influence on signal based learning and has been recommended to examine whether evaluative conditioning and signal based conditioning dissociate (De Houwer et al., 2001). The conditioning design used here has been shown to lead to successful acquisition of conditional valence in both explicit and implicit measures (see Luck and Lipp, in revision).

In Stream 1 (Experiments 1 and 2), we examined whether evaluative conditioning was subject to spontaneous recovery by testing conditional responding following a delay after extinction. In Stream 2 (Experiments 3a, 3b, and 4) we examined whether evaluative conditioning is subject to reinstatement by presenting the valenced USs in a randomized stream after extinction. In Stream 3 (Experiments 5 and 6), we examined whether evaluative conditioning was subject to ABA renewal by changing the background screen color (i.e. acquisition occurred on color 1, extinction on color 2, and the renewal test on color 1) to manipulate context. We hypothesized that in all experiments the acquisition of evaluative learning would be observed with differential valence evaluations and expectancy ratings acquired between the CSv and CSn (between-participant designs) or the CSp and CSu (within-participant designs). We predicted that differential evaluations and expectancy ratings would reduce throughout extinction, but based on Hoffman et al. (2010), we expected that some differential valence evaluations might remain at the end of extinction. Based on the research conducted within the human fear conditioning literature, we predicted that in all cases the magnitude of differential conditional valence evaluations and expectancy ratings would increase in size after the relapse inductions.

Stream 1 – Spontaneous Recovery

Experiment 1 – Between Participant Conditioning Design

Method

Participants. 48 (32 female) undergraduate Curtin university students or community volunteers aged between 18 and 65 (*M* = 23.81, *SD* = 9.94) volunteered participation in exchange for course credit or monetary compensation or $15 AU. The research protocol was approved by the Curtin University Human Research Ethics Committee. An a-priori power analysis conducted using G-Power 3.1 revealed that 34 participants would be required to be 80% confident of detecting a large effect (f = .5) for the critical relapse interaction. This was conducted based on an ANOVA F test (within-between interaction; 2 groups [Positive Conditioning, Negative Conditioning]; 2 measurement points [Last Block of Extinction; Relapse Test]; Effect size specification: as in Cohen (1988); nonsphericity correction: 1) using the CS difference scores (CSv-CSn). Difference scores were used for the power calculations as a main effect of phase detected using the CS difference scores is statistically equivalent to a CS × Phase interaction and a Phase × Group interaction detected using the CS difference scores is statistically equivalent to a CS × Phase × Group interaction. Participants were randomly assigned to the positive (n = 24) or negative conditioning group (n = 24).

Apparatus/Stimuli. The experiment was conducted on PCs running Windows 7 in a group computer laboratory (maximum of 4 participants tested simultaneously). Participants were seated in separate cubicles in front of a 24 inch LED screen (resolution: 1920 × 1080; refresh rate: 120 Hz) and responses were recorded using a QWERTY keyboard. The experiment was programmed using DMDX software (version: 5.134; Forster & Forster, 2003). CSs were 900 × 675 pixel pictures of geometric shapes (circle, diamond, triangle, square; black outlines on a white background). CSs were presented centred in the middle of a black background for a duration of 2 s. The two shapes used during conditioning and the shapes used as CSv and CSn were counterbalanced across participants. USs were 6 pleasant (codes: 1460, 1710, 2154, 2340, 5825, 5833; mean valence rating = 7.92 [1 = *low pleasure*, 9 = *high pleasure*]; mean arousal rating = 5.22 [1 = *low arousal*, 9 = *high arousal*]; subject matter: happy family, animals, nature scenery), 6 unpleasant (codes: 9560, 9340, 9295, 9220, 2800, 2703; mean valence rating = 2.32; mean arousal rating = 4.93; subject matter: rubbish, starving children, cemetery scenes, sick animals), and 6 neutral images (codes: 7006, 7217, 7035, 7150, 7175, 7705; mean valence rating = 4.73; mean arousal rating = 2.26; subject matter: household items such as furniture and crockery) taken from the International Affective Picture System (IAPS; Lang, Bradley & Cuthbert, 2008). USs were presented centred in the middle of a black background for 2 s at the offset of the CS (delay conditioning; interstimulus interval: 2 s). The inter-trial interval was a 2 s black screen.

The unrelated filler task used as part of the spontaneous recovery manipulation was a Wikipedia article on the ring tailed lemur and a questionnaire consisting of 20 questions taken directly from the article (examples: ‘How is the ring-tailed lemur recognised?’; ‘What has reinforced the notion that lemurs are less intelligent than their simian cousins?’; ‘How do lemurs mark their territory?’). Participants were asked to rate CS and US valence on a 1-9 Likert scale with the anchors (1 = *unpleasant* and 9 = *pleasant*). US expectancy[1] was measured by asking participants to predict what type of pictures would follow the CSs on a 1-7 scale with the anchors (1 = *always neutral,* 2 = *often neutral,* 3 = *sometimes neutral,* 4 = *neither neutral nor pleasant (or unpleasant),* 5 = *sometimes pleasant (or unpleasant),* 6 = *often pleasant (or unpleasant),* 7 = *always pleasant (or unpleasant).* In the positive conditioning group the response options for 4-7 (valenced pictures) were ‘pleasant’, while, in the negative conditioning group the options were ‘unpleasant’.

Procedure. Participants provided informed consent, were assigned to a counterbalancing condition, and seated in a cubicle in front of a monitor. The experiment consisted of a CS valence baseline assessment, the acquisition and extinction phase, a 15 minute break in which they completed the unrelated filler task, the spontaneous recovery test, and a US valence assessment. During the baseline assessment, the CSv and CSn were presented with the valence rating scale (the order of CSv and CSn was randomised in all rating tasks). During acquisition, 4 training blocks, each consisting of 6 pairings between the CSv and valenced pictures (pleasant pictures in the positive conditioning group and unpleasant pictures in the negative conditioning group) and 6 pairings between CSn and neutral pictures, were presented. Each US picture was

[1] Please note: The US expectancy measure could be considered an online measure of contingency awareness as participants predict not only that a US will occur but also which US is likely to occur.

paired with the CSv/CSn once during each acquisition block and trial order was randomised. CS valence and US expectancy was assessed after each block, with both CSv and CSn evaluations always preceding the expectancy ratings. During extinction, 8 blocks of 12 random presentations of CSv and CSn alone (6 each) were presented, with CS valence and US expectancy assessments assessed after each block.

After the last extinction block, participants completed an unrelated filler task for 15 minutes. During this task, participants were asked to read the Wikipedia article on the ring tailed lemur and to answer as many questions as possible during the 15 minute period. After 15 minutes, the experimenter collected the questionnaires and started the spontaneous recovery test, which consisted of an assessment of CSv and CSn valence, followed by an assessment of CSv and CSn expectancy ratings. After the spontaneous recovery test phase, participants evaluated the valence of the US images (pleasant, neutral, and unpleasant), before providing demographic information, and a contingency assessment in which they were required to identify which shape had been paired with the valenced images and which shape had been paired with neutral images during the initial part of the experiment.

Data Preparation and Analyses. Frequentist statistics were conducted using IBM SPSS Statistics 25 with an alpha cut-off of .05. Phillai's trace statistics of the multivariate solution are reported for all analyses. The report of the results is based primarily on the frequentist analyses, however, to allow evidence for the null hypothesis to be quantified and for the reader's interest we have supplemented the frequentist statistics with Bayesian analyses conducted using JASP 0.9.2.0. For main effects, we report BF_{10} values from the model comparison and for interactions we report $BF_{inclusion}$ values from the effects analysis (across matched models). The $BF_{inclusion}$ (across matched models) compares models containing the effect of interest to equivalent models that have been stripped of the effect. This allows the individual interactions' contributions to be quantified without the lower order effects contributing to the strength of the model (for a rationale and discussion of this method – known as the Baws approach – see Mathôt, 2017). Follow-up analyses were conducted by comparing the effects of interest with paired and independent samples Bayesian *t*-tests (two-sided). For all Bayesian analyses we used the default settings defined in Jasp (ANOVAs: r scale fixed effects = 0.5, r scale random effects = 1, r scale covariates = 0.354, auto sample; *t*-tests: Cauchy prior scale: 0.707).

We report the BF_{10} values for all tests – values greater than 1 provide evidence for the alternative hypothesis and values less than one provide evidence for the null hypothesis. Bayes factors are interpretable as a graded measure of evidence, however, in general we have followed Jeffreys' (1961) conventions for interpreting the Bayes factors. BF_{10} values between 0.33 and 1, 0.33 and 0.1, and < 0.1 and are generally interpreted to provide anecdotal, substantial, and strong support for the null hypothesis, respectively. Similarly, BF_{10} values between 1 and 3, 3 and 10, and > 10 are generally interpreted to provide anecdotal, substantial, and strong support for the alternative hypothesis, respectively.

Analyses were re-run excluding participants who did not pass the post-experimental contingency assessment (non-verbalizers) to examine whether this exclusion results in a deviation from the general pattern of results. Results from the full sample are reported in all analyses but we have made note of any instance where exclusion of these participants changes the significance of a highest order omnibus effect or influences the relapse analyses. Raw data for Experiments 2, 3b, 4, 5, and 6 is available at https://osf.io/hvak4. We are unable to make the data from Experiments 1 and 3a publically available as this option was not part of our ethics process when these experiments were conducted and therefore participants did not give informed consent for their data to be made publically available.

Preliminary Checks. Separate Pearson's chi-square tests were used to check the male:female:other sex ratio and the contingency pass:fail ratio did not differ between the groups. An independent samples *t*-tests and a Bayesian independent samples *t*-tests were used to check that the groups did not differ on age. Baseline CS evaluations were subjected to a CS (CSv, CSn) × 2 Group (positive conditioning, negative conditioning) mixed-model factorial ANOVA to check that the baseline evaluations did not differ between the groups.

Acquisition and Extinction Analyses. CS valence evaluations and US expectancy ratings assessed during acquisition were analyzed with separate 2 CS (CSv, CSn) × 2 Group (positive conditioning, negative conditioning) × 4 Block (A1, A2, A3, A4) mixed-model factorial ANOVAs. The CS valence evaluations and US expectancy ratings measured during extinction were subjected to separate 2 CS (CSv, CSn) × 2 Group (positive conditioning, negative conditioning) × 9 Block (A4, E1, E2, E3, E4, E5, E6, E7, E8) mixed-model factorial ANOVAs. The last rating block of acquisition (A4) is included in the analysis to assess the effect of the first block of extinction training. The term *conditional valence* is used to refer to a pattern of differential valence between the CSv and CSn that is in the direction expected based on conditioning training (i.e. CSv less pleasant than CSn for negative conditioning and CSv more pleasant than CSn for positive conditioning). The term conditional expectancy is used to refer to a pattern of differential expectancy between the CSv and the CSn that is in the direction expected based on conditioning training (i.e. prediction scores are higher for CSv than for CSn regardless of group). CS difference scores (CSv – CSn) were subjected to the above ANOVAs (without the CS factor) to compare the size of conditional responding at the beginning and end of acquisition and extinction to assess the change in the strength of the conditional response across time. When CS difference scores are used to compare conditional responding across time to follow up CS × Phase interactions, the resulting comparison statistic (i.e. the main effect of phase) is equivalent to the CS × Phase interaction from the main ANOVA including CS as a factor. Similarly, when CS difference scores compare conditional responding between groups to follow up CS × Group interactions, the resulting comparison statistic (i.e. the main effect of group) is equivalent to the CS × Group interaction from the main ANOVA. To avoid redundancies, we have not reported these duplicate statistics. Group differences detected in the CS valence analyses will primarily reflect that conditional valence is pleasant in the positive conditioning group and unpleasant in the negative conditioning group. Therefore, to examine whether there was a difference in the *strength* of positive and negative evaluative learning the valence difference score at the end of acquisition (A4) was equated between the groups by reversing the sign (multiplying by -1) for the negative conditioning group only and this score was subjected to an independent samples *t*-test.

Relapse Analyses. CS valence evaluations and US expectancy ratings assessed after the last block of extinction and during the relapse test were subjected to separate 2 CS (CSv, CSn) × 2 Group (positive conditioning, negative conditioning) × 2 Phase (last block of extinction, relapse test) mixed-model factorial ANOVAs. Significant CS × Group × Phase interactions were followed up by examining whether conditional valence/expectancy increased in magnitude from after extinction to the relapse test in each group. Relapse is defined to have occurred if the conditional response increases in the direction established during acquisition (i.e. the CSv becomes more unpleasant relative to the CSn in the negative conditioning group and CSv becomes more pleasant relative to the CSn in the positive conditioning group).

Unconditional Stimulus Valence Analyses. The valence ratings for the pleasant, neutral, and unpleasant USs were averaged across individual pictures and the mean US evaluations were subjected to a 3 Valence (positive, neutral, negative) × 2 Group (positive conditioning, negative conditioning) factorial ANOVA.

Results

Preliminary Analyses. No pre-existing differences between the groups were uncovered in the preliminary analyses. The means, standard deviations, and comparison statistics are presented in Table 1. Three participants were not able to verbalize the experimental contingencies. The experimental program used here did not require participants to enter a response for the valence and expectancy questions before continuing with the experiment which resulted in a number of missing responses. During acquisition, two cases for CS valence[2] (positive conditioning: 1; negative conditioning: 1) and 7 cases for US expectancy (positive conditioning: 5; negative conditioning: 2) were lost. During extinction, three cases for CS valence (positive conditioning: 1; negative conditioning: 2) and 5 cases for US expectancy (positive conditioning: 4; negative conditioning: 1) were lost. During spontaneous recovery, 1 case for CS valence (positive conditioning group) and 2 cases for US expectancy (positive conditioning group) were lost.

Acquisition. The CS valence evaluations and US expectancy ratings measured during Experiment 1 are presented in Figures 1 and 2, respectively. Conditional valence and expectancy learning developed in both groups, confirming that the acquisition phase was successful.

---

[2] All participants provided data during the last block of acquisition and therefore the independent samples *t*-test comparing the equated valence difference score at the end of acquisition was conducted will all available data.

CS Valence. A main effect of block, $F(3, 42) = 4.12$, $p = .012$, $\eta p^2 = .227$, $BF_{10} = 0.18$, and a CS × Group interaction, $F(1, 44) = 40.86$, $p < .001$, $\eta p^2 = .481$, $BF_{inc} = 1.79 \times 10^{23}$, were moderated by a CS × Group × Block interaction, $F(3, 42) = 8.37$, $p < .001$, $\eta p^2 = .374$, $BF_{inc} = 2360.33$. In the positive conditioning group, conditional valence was present during[3] all blocks, all $F(1, 44) > 5.20$, $p < .028$, $\eta p^2 > .105$, BF_{10}: blocks 1 = 1.47, 2 = 21.08, 3 = 408.01, 4 = 2016.83, and increased from blocks 1 to 4, $p = .001$, $BF_{10} = 27.94$. In the negative conditioning group, conditional valence was not present during block 1, $F(1, 44) = 1.17$, $p = .286$, $\eta p^2 = .026$, $BF_{10} = 0.40$, but was during all other blocks, all $F(1, 44) > 5.61$, $p < .023$, $\eta p^2 > .114$, BF_{10}: blocks 2 = 1.61, 3 = 88.00, 4 = 730.60. The size of this difference increased from block 1 to 4, $p = .001$, $BF_{10} = 20.61$. The magnitude of conditional valence at the end of acquisition did not differ between the groups (negative conditioning group: $M = 2.29$, $SD = 2.22$; positive conditioning group: $M = 2.58$, $SD = 2.30$), $t(46) = 0.45$, $p = .657$, $BF_{10} = 0.31$. Although the CSn was paired with the same neutral pictures between groups, visual inspection of Figure 1 suggested the conditioning groups evaluated the CSn differently. The analyses confirmed this impression revealing that the negative conditioning group evaluated the CSn as more pleasant than the positive conditioning group during blocks 3, $BF_{10} = 2.85$, and 4, $BF_{10} = 126.20$, both $F(1, 44) > 5.80$, $p < .021$, $\eta p^2 > .116$. The remaining omnibus effects did not reach significance, all $F < 2.59$, $p > .115$, $\eta p^2 < .056$, BF < 0.72.

US Expectancy. Main effects of CS, $F(1, 39) = 461.09$, $p < .001$, $\eta p^2 = .922$, $BF_{10} = 6.68 \times 10^{102}$, block, $F(3, 37) = 3.20$, $p = .034$, $\eta p^2 = .206$, $BF_{10} = 0.02$, and group, $F(1, 39) = 8.75$, $p = .005$, $\eta p^2 = .183$, $BF_{10} = 0.53$, were moderated by a CS × Group interaction, $F(1, 39) = 11.27$, $p = .002$, $\eta p^2 = .224$, $BF_{inc} = 1.00 \times 10^6$, and a CS × Block interaction, $F(3, 37) = 7.01$, $p = .001$, ηp^2

[3] Please note: the corresponding CS valence evaluations and US expectancy ratings were assessed after completion of respective training block.

= .362, $BF_{inc} = 1.36 \times 10^6$. The CS × Group interaction revealed that conditional expectancy was present in both groups, both $F(1, 39) > 152.91$, $p < .001$, $\eta p^2 > .796$, $BF_{10} > 5.85 \times 10^5$, but the magnitude of this difference was larger in the negative conditioning group, $BF_{10} = 19.53$. The CS × Block interaction revealed that conditional expectancy was present during all blocks, all $F(1, 39) > 65.26$, $p < .001$, $\eta p^2 > .625$, $BF_{10} > 1.84\times 10^7$, but increased from block 1 to 4, $p < .001$, $BF_{10} = 712.50$. The remaining omnibus effects did not reach significance, all $F < 1.49$, $p > .235$, $\eta p^2 < .108$, BF < 0.10.

Extinction. The conditional valence and expectancy learning acquired during acquisition reduced throughout extinction in both groups.

CS Valence. A CS × Group interaction, $F(1, 43) = 22.58$, $p < .001$, $\eta p^2 = .344$, $BF_{inc} = 2.62 \times 10^{20}$, was moderated by a CS × Group × Block interaction, $F(8, 36) = 4.04$, $p = .002$, $\eta p^2 = .473$, $BF_{inc} = 612.44$. In the positive conditioning group, conditional valence was present after the last block of acquisition, $BF_{10} = 1530.56$, and during extinction blocks 1, 2, 4, 5, 6, and 7, all $F(1, 43) > 4.87$, $p < .033$, $\eta p^2 > .101$, BF_{10}: blocks 1 = 151.64, 2 = 2.51, 4 = 1.82, 5 = 1.89, 6 = 1.89, 7 = 0.74, but was marginal during blocks 3, $F(1, 43) = 3.13$, $p = .084$, $\eta p^2 = .068$, $BF_{10} = 1.32$, and 8, $F(1, 43) = 3.74$, $p = .060$, $\eta p^2 = .080$, $BF_{10} = 0.56$. In the negative conditioning group, conditional valence was present during the last block of acquisition, $BF_{10} = 432.77$, and during extinction blocks 1, $BF_{10} = 36.89$, and 2, $BF_{10} = 2.50$, all $F(1, 43) > 5.47$, $p < .025$, $\eta p^2 > .112$, was marginal during block 3, $F(1, 43) = 3.70$, $p = .061$, $\eta p^2 = .079$, $BF_{10} = 0.76$, and was not present during the remaining blocks, all $F(1, 43) < 2.83$, $p > .100$, $\eta p^2 < .062$, BF_{10}: blocks 4 = 1.38, 5 = 1.83, 6 = 1.01, 7 = 1.01, 8 = 0.40. Conditional valence decreased from after acquisition to after extinction in both the positive conditioning group, $BF_{10} = 1.93$, and the negative conditioning group, $BF_{10} = 123.12$, both $p < .012$. The difference in CSn evaluations between groups was still present during block 1, $F(1, 43) = 5.17$ $p = .028$, $\eta p^2 = .107$, $BF_{10} = 2.24$, and was marginal during blocks 4, $F(1, 43) = 4.04$, $p = .051$, $\eta p^2 = .086$, $BF_{10} = 1.46$, and 5, $F(1, 43) = 3.73$, $p = .060$, $\eta p^2 = .080$, $BF_{10} = 1.29$. The remaining omnibus effects did not reach significance, all $F < 1.29$, $p > .283$, $\eta p^2 < .222$, $BF < 1.61$.

US Expectancy. Main effects of CS, $F(1, 41) = 73.01$, $p < .001$, $\eta p^2 = .640$, $BF_{10} = 1.26 \times 10^{33}$, block, $F(8, 34) = 2.67$, $p = .021$, $\eta p^2 = .386$, $BF_{10} = 543.77$, and group, $F(1, 41) = 10.72$, $p = .002$, $\eta p^2 = .207$, $BF_{10} = 15.03$, and a CS × Block interaction, $F(8, 34) = 30.03$, $p < .001$, $\eta p^2 = .876$, $BF_{inc} = 2.08 \times 10^{45}$, were detected. The main effect of group revealed that overall expectancy ratings were lower (stronger prediction of neutral pictures) in the negative conditioning group than in the positive conditioning group. The CS × Block interaction revealed that conditional expectancy was present during the last block of acquisition, $BF_{10} = 3.90 \times 10^{21}$, and after extinction blocks 1, 2, 3, 4, 5, and 6, all $F(1, 41) > 6.56$, $p < .015$, $\eta p^2 > .137$, BF_{10}: blocks 1 = 1.24×10^9, 2 = 44.57, 3 = 123.27, 4 = 9.19, 5 = 7.03, 6 = 2.18, but not during blocks 7, $BF_{10} = 0.43$, or 8, $BF_{10} = 0.38$, both $F(1, 41) < 2.32$, $p > .135$, $\eta p^2 < .054$. Conditional expectancy reduced from after acquisition to after extinction, $p < .001$, $BF_{10} = 6.61 \times 10^{14}$. The remaining omnibus effects did not reach significance[4], all $F < 2.02$, $p > .074$, $\eta p^2 < .322$, $BF < 3.60$.

Spontaneous Recovery. Spontaneous recovery of expectancy conditioning was observed but spontaneous recovery of evaluative learning did not occur.

CS Valence. A main effect of phase, $F(1, 45) = 6.45$, $p = .015$, $\eta p^2 = .125$, $BF_{10} = 8.25$, revealed that overall evaluations became more pleasant from the last block of extinction to the

[4] This marginal effect is the CS × Block × Group interaction but BF provides evidence for the null hypothesis, $BF_{inc} = 0.33$.

spontaneous recovery test. A CS × Group interaction, $F(1, 45) = 5.05$, $p = .030$, $\eta p^2 = .101$, $BF_{inc} = 5.18$, revealed that conditional valence was present in the positive conditioning group, $F(1, 45) = 6.25$, $p = .016$, $\eta p^2 = .122$, $BF_{10} = 0.98$, but not in the negative conditioning group, $F(1, 45) = 0.43$, $p = .515$, $\eta p^2 = .009$, $BF_{10} = 0.45$. The CS × Group × Phase interaction did not attain significance, $F(1, 45) = 0.74$, $p = .395$, $\eta p^2 = .016$, $BF_{inc} = 0.32$, and no evidence for relapse was observed in the follow-up tests. The magnitude of conditional valence did not increase from after extinction to the spontaneous recovery test in the positive conditioning group, $F(1, 45) = 1.12$, $p = .296$, $\eta p^2 = .024$, $BF_{10} = 0.28$, or in the negative conditioning group, $F(1, 45) = 0.02$, $p = .883$, $\eta p^2 < .001$, $BF_{10} = 0.25$. The remaining omnibus effects did not reach significance, all $F < 1.77$, $p > .191$, $\eta p^2 < .038$, $BF < 0.42$.

US Expectancy. Main effects of CS, $F(1, 44) = 13.39$, $p = .001$, $\eta p^2 = .233$, $BF_{10} = 535.78$, phase, $F(1, 44) = 5.83$, $p = .020$, $\eta p^2 = .117$, $BF_{10} = 1.43$, and group, $F(1, 44) = 11.34$, $p = .002$, $\eta p^2 = .205$, $BF_{10} = 21.78$, and a CS × Phase interaction, $F(1, 44) = 16.06$, $p < .001$, $\eta p^2 = .267$, $BF_{inc} = 10.20$, were detected. Similar to the pattern present during extinction, the main effect of group revealed that overall expectancy scores were higher in the positive conditioning than in the negative conditioning group. The CS × Phase interaction provided evidence for spontaneous recovery. Conditional expectancy was not present after the last block of extinction, $F(1, 44) = 1.84$, $p = .182$, $\eta p^2 = .040$, $BF_{10} = 0.37$, but increased from after extinction to the relapse test, $BF_{10} = 128.17$, and was present during the spontaneous recovery assessment, $F(1, 44) = 19.74$, $p < .001$, $\eta p^2 = .310$, $BF_{10} = 417.69$. The remaining omnibus effects did not reach significance, all $F < .046$, $p > .501$, $\eta p^2 < .011$, $BF < 0.32$.

US Valence. The mean US valence evaluations are displayed in Table 2. Main effects of valence, $F(2, 42) = 561.04$, $p < .001$, $\eta p^2 = .964$, $BF_{10} = 2.54 \times 10^{61}$, and group, $F(1, 43) = 13.61$,

$p < .001$, $\eta p^2 = .240$, $BF_{10} = 0.31$, were moderated by a Valence × Group interaction, $F(2, 42) = 6.72$, $p = .003$, $\eta p^2 = .243$, $BF_{inc} = 43.30$. Both groups evaluated the positive pictures as more pleasant than the neutral pictures, both $p < .001$, $BF_{10} > 7.04 \times 10^8$, and the negative pictures, as more unpleasant than the neutral pictures, both $p < .001$, $BF_{10} > 2.66 \times 10^4$, but, while the valence of the positive, $F(1, 43) = 0.75$, $p = .392$, $\eta p^2 = .017$, $BF_{10} = 0.40$, and negative images, $F(1, 43) = 0.17$, $p = .685$, $\eta p^2 = .004$, $BF_{10} = 0.32$, did not differ between the groups, the negative conditioning group evaluated the neutral pictures as more pleasant than the positive conditioning group, $F(1, 43) = 16.96$, $p < .001$, $\eta p^2 = .283$, $BF_{10} = 138.44$.

Discussion

In Experiment 1, we examined whether conditional evaluative and expectancy conditioning would relapse after a spontaneous recovery manipulation in a between-participants conditioning design. After extinction, participants completed a 15 minute filler task before the spontaneous recovery assessment. Conditional valence and expectancy were successfully acquired throughout acquisition in both groups. We observed a substantial reduction in the magnitude of this conditional valence and expectancy from after acquisition to after extinction, confirming that the extinction procedure successfully reduced conditional learning. Unexpectedly, in the latter half of acquisition and the first block of extinction, evaluations of the CSn were more pleasant in the negative conditioning group than in the positive conditioning group. This is interesting as the training for the CSn was identical in both groups. Post-experimentally, the negative conditioning group also evaluated the neutral USs as more pleasant than the positive conditioning group and therefore it is possible that there was a pre-existing difference between the groups. As US valence was measured post-experimentally, however, it is also possible that this valence shift for the neutral US pictures occurs because the neutral pictures are perceived as less pleasant when they are presented alongside pleasant pictures but more pleasant when they are presented alongside unpleasant pictures.

Spontaneous recovery was observed for expectancy learning after the 15 minute delay, with the magnitude of conditional expectancy increasing from after extinction to the spontaneous recovery test. In contrast, we did not observe conditional relapse after spontaneous recovery in the CS valence measure. CSv and CSn evaluations became more pleasant from after extinction to the spontaneous recovery test in both groups. Strong support for the alternative hypothesis was obtained in the Bayesian analyses for the spontaneous recovery of expectancy learning but substantial evidence for the null was obtained in both groups for evaluative relapse. Observing strong relapse of conditional expectancy conditioning suggests that the 15 minute filler task was sufficient to induce relapse and that the absence of spontaneous recovery in the valence measure was not due to an insufficient delay. Nevertheless, we conducted the spontaneous recovery test one day after acquisition and extinction in Experiment 2 to replicate these results with a stronger spontaneous recovery manipulation.

Experiment 2 – Within Participant Conditioning Design

Method

Participants. 42 (26 female[5]) Curtin University students and community members aged between 18 and 62 (*M* = 24.98, *SD* = 8.14) volunteered participation in exchange for course credit or monetary compensation of $15 AU. The research protocol was approved by the Curtin University Human Research Ethics Committee. An a-priori power analysis conducted using G-Power 3.1 revealed that 34 participants would be required to be 80% confident of detecting a

[5] Two participants did not provide sex information

large effect (*f* = .5) for the critical relapse interaction. This was conducted based on an ANOVA *F* test (repeated measures-within factors; 1 group; 2 measurement points [Last Block of Extinction; Relapse Test]; Effect size specification: as in Cohen (1988); nonsphericity correction: 1) using the CS difference scores (CSp - CSu).

Apparatus/Stimuli. The experiment was programmed using Inquisit 4 software. The two shapes used during conditioning and the shapes used as the CSp and CSu were counterbalanced across participants. CSs were presented centred in the middle of a black background at a size of 50% of the participants' screen. USs were presented centred in the middle of a black background at a size of 60% of the participants' screen. US expectancy was measured by asking participants to predict what type of pictures would follow the CSs on a 1-7 scale with the anchors (1 = *always unpleasant*, 4 = *no pictures*, 7 = *always pleasant*). The remainder of the apparatus/stimuli were the same as for Experiment 1.

Procedure. Participants provided informed consent, were seated in an individual testing cubicle, and completed a baseline assessment of CSp and CSu valence. During acquisition, 4 training blocks, each consisting of 6 pairings between CSp and pleasant images and 6 pairings between CSu and unpleasant pictures, were presented. The CSp/CSu were paired with each pleasant/unpleasant US once during each training block and trial order was randomised. CS valence and US expectancy was measured after each block (valence evaluations always preceded expectancy ratings). During extinction, 8 blocks of 12 random presentations of CSp and CSu alone (6 each) were presented, with CS valence and US expectancy assessed after each block. After the last extinction measurement phase, participants provided demographic information, were asked to identify which shape was paired with the pleasant images and which shape was paired with the unpleasant images, and were asked to come back the next day. On Day 2, participants returned to the same testing cubicle for the spontaneous recovery test phase. This consisted of a rating of CSp and CSu valence and expectancy. Participants were then asked to evaluate the valence of the pleasant and unpleasant US pictures and a repeat of the contingency assessment was administered.

Data Preparation and Analyses. Reporting conventions and statistical programs were as in Experiment 1.

Preliminary Checks. Baseline CSp and CSu evaluations were subjected to a one-way ANOVA and a Bayesian paired sample t-test.

Acquisition and Extinction Analyses. CS valence evaluations and US expectancy ratings assessed during acquisition were analyzed with separate 2 CS (CSp, CSu) × 4 Block (A1, A2, A3, A4) repeated measures ANOVAs. CS valence evaluations and US expectancy ratings measured during extinction were subjected to separate 2 CS (CSp, CSu) × 9 Block (A4, E1, E2, E3, E4, E5, E6, E7, E8) repeated measures ANOVAs. The terms conditional valence and conditional expectancy refer to a difference between the CSp and CSu in the direction that would be expected for conditioning (i.e. CSp evaluated as more pleasant than CSu and CSp expectancy scores higher than CSu expectancy scores).

Relapse Analyses. CS valence evaluations and US expectancy ratings assessed after the last block of extinction and during the relapse test were subjected to separate 2 CS (CSp, CSu) × 2 Phase (last block of extinction, relapse test) repeated-measures ANOVAs. Significant CS × Phase interactions were followed up by examining whether conditional valence or expectancy increased in magnitude from after extinction to the relapse test. Relapse is defined to have occurred if the conditional response increases in the direction established during acquisition (i.e. the CSp becomes more unpleasant relative to the CSu).

Unconditional Stimulus Valence Analyses. The valence ratings for the pleasant and unpleasant USs were averaged across individual pictures and the mean US evaluations were subjected to a one-way ANOVA and a Bayesian paired sample t-test.

Results

Preliminary Analyses. CSp and CSu evaluations did not differ during the baseline assessment, $F(1, 41) = 0.56$, $p = .460$, $\eta p^2 = .013$, $BF_{10} = 0.22$. Twelve participants on Day 1 and 2 participants on Day 2 failed the contingency assessment. One participant did not complete the contingency questions on both days.

Acquisition. The CS valence evaluations and US expectancy ratings measured during Experiment 2 are presented in Figures 3 and 4, respectively. Conditional valence and expectancy conditioning were successfully acquired during acquisition.

CS Valence. Main effects of CS, $F(1, 41) = 69.67$, $p < .001$, $\eta p^2 = .630$, $BF_{10} = 1.68 \times 10^{44}$, and block, $F(3, 39) = 4.26$, $p = .011$, $\eta p^2 = .247$, $BF_{inc} = 0.02$, were moderated by a CS × Block interaction, $F(3, 39) = 10.56$, $p < .001$, $\eta p^2 = .448$, $BF_{inc} = 1113.47$. Conditional valence was present during all acquisition blocks, all $F(1, 38) > 17.59$, $p < .001$, $\eta p^2 > .300$, BF_{10}: block 1 = 173.90, all other > 1.30×10^6, and increased from block 1 to 4, $p = .001$, $BF_{10} = 37.73$.

US Expectancy. A main effect of CS, $F(1, 41) = 255.52$, $p < .001$, $\eta p^2 = .862$, $BF_{10} = 1.56 \times 10^{99}$, was moderated by a CS × Block interaction, $F(3, 39) = 7.47$, $p < .001$, $\eta p^2 = .365$, $BF_{inc} = 4257.24$. Conditional expectancy was present during all acquisition blocks, all $F(1, 41) > 51.09$, $p < .001$, $\eta p^2 > .554$, $BF_{10} > 1.25 \times 10^6$, and increased from blocks 1 to 4, $p < .001$, $BF_{10} = 629.30$. The main effect of block did not reach significance, $F(3, 39) = 0.82$, $p = .493$, $\eta p^2 = .059$, $BF_{10} = 0.01$.

Extinction. The conditional valence and expectancy learning acquired during acquisition reduced throughout the extinction phase.

CS Valence. A main effect of CS, $F(1, 41) = 17.12$, $p < .001$, $\eta p^2 = .295$, $BF_{10} = 3.83 \times 10^{22}$, was moderated by a CS × Block interaction, $F(8, 34) = 4.43$, $p = .001$, $\eta p^2 = .511$, $BF_{inc} = 2.54 \times 10^{15}$. Conditional valence was present during the last block of acquisition, $BF_{10} = 1.31 \times 10^{6}$, and during extinction blocks 1, 2, 3, 4, and 7, all $F(1, 41) > 4.14$, $p < .049$, $\eta p^2 > .091$, BF_{10}: blocks 1 = 2052.98, 2 = 11.51, 3 = 1.24, 4 = 1.08, 7 = 1.58, was marginal during block 8, $F(1, 41) = 3.30$, $p = .077$, $\eta p^2 = .075$, $BF_{10} = 0.75$, and was not present during blocks 5 and 6, both $F(1, 41) < 2.17$, $p > .149$, $\eta p^2 < .051$, BF_{10}: blocks 5 = 0.27, 6 = 0.45. Conditional valence decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 4.29 \times 10^{4}$. The main effect of block did not reach significance, $F(8, 34) = 1.40$, $p = .233$, $\eta p^2 = .248$, $BF_{10} = 0.00$.

US Expectancy. A main effect of CS, $F(1, 41) = 35.18$, $p < .001$, $\eta p^2 = .462$, $BF_{10} = 6.22 \times 10^{38}$, was moderated by a CS × Block interaction, $F(8, 34) = 28.26$, $p < .001$, $\eta p^2 = .869$, $BF_{inc} = 6.94 \times 10^{61}$. Conditional expectancy was present after the last block of acquisition, $BF_{10} = 9.12 \times 10^{20}$, and during extinction blocks 1, 2, 3, 4, 5, and 7, all $F(1, 41) > 4.97$, $p < .032$, $\eta p^2 > .108$, BF_{10}: blocks 1 = 2066.05, 2 = 22.47, 3 = 2.50, 4 = 2.16, 5 = 3.32, 7 = 1.54, was marginal during block 8, $F(1, 41) = 4.04$, $p = .051$, $\eta p^2 = .090$, $BF_{10} = 1.03$, and was not present during block 6, $F(1, 41) = 2.12$, $p = .153$, $\eta p^2 = .049$, $BF_{10} = 0.44$. Conditional expectancy decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 1.26 \times 10^{15}$. The main effect of block did not reach significance, $F(8, 34) = 0.98$, $p = .472$, $\eta p^2 = .187$, $BF_{10} = 2.32 \times 10^{-4}$

Spontaneous Recovery. Spontaneous recovery of expectancy conditioning occurred, but spontaneous recovery of evaluative learning did not.

CS Valence. A main effect of CS, $F(1, 41) = 5.44$, $p = .025$, $\eta p^2 = .117$, $BF_{10} = 19.17$, revealed that CSp was evaluated as more pleasant than CSu. A main effect of phase, $F(1, 41) = 5.58$, $p = .023$, $\eta p^2 = .120$, $BF_{10} = 1.11$, revealed that evaluations were more pleasant during the spontaneous recovery test than during the last block of extinction. The CS × Phase interaction, $F(1, 41) = 0.70$, $p = .408$, $\eta p^2 = .017$, $BF_{inc} = 0.25$, did not reach significance. In the Bayesian follow-up analyses, the magnitude of the *conditional* valence did not change from after extinction to the spontaneous recovery test, $BF_{10} = 0.23$. Thus, spontaneous recovery did not occur.

US Expectancy. A main effect of CS, $F(1, 41) = 23.58$, $p < .001$, $\eta p^2 = .365$, $BF_{10} = 1.15 \times 10^8$, was moderated by a CS × Phase interaction, $F(1, 41) = 19.32$, $p < .001$, $\eta p^2 = .320$, $BF_{inc} = 963.97$, confirming that relapse of expectancy learning occurred. After extinction, the pattern of conditional expectancy attained marginal significance, $F(1, 41) = 4.04$, $p = .051$, $\eta p^2 = .090$, $BF_{10} = 1.03$, but the magnitude of conditional expectancy increased from after extinction to the relapse test, $BF_{10} = 307.30$, and was present during the spontaneous recovery assessment, $F(1, 41) = 31.03$, $p < .001$, $\eta p^2 = .431$, $BF_{10} = 9946.53$. The main effect of phase did not reach significance, $F(1, 41) = 0.23$, $p = .637$, $\eta p^2 = .005$, $BF_{10} = 0.17$.

US Valence. Pleasant USs ($M = 7.88$, $SD = 0.78$) were evaluated as more pleasant than unpleasant USs ($M = 2.03$, $SD = 0.91$), $F(1, 41) = 719.06$, $p < .001$, $\eta p^2 = .946$, $BF_{10} = 2.16 \times 10^{24}$.

Discussion

In Experiment 2, we examined whether conditional valence and expectancy conditioning would relapse after a time delay using a within-participants conditioning design. Participants completed acquisition and extinction training on Day 1 and the spontaneous recovery test in the same laboratory on Day 2. Conditional valence and expectancy were acquired during acquisition and significantly reduced throughout extinction. Spontaneous recovery of expectancy learning was observed, with the magnitude of conditional expectancy increasing from after extinction on Day 1 to the spontaneous recovery test on Day 2. Replicating the results of Experiment 1, we did not observe relapse of evaluative learning after the time delay. Differential valence did not re-emerge, instead evaluations of *both* CSp and CSu became more positive from the end of extinction to the spontaneous recovery test. This overall increase in pleasantness from Day 1 to Day 2 is similar to that detected in the negative conditioning group of Experiment 1 and is a qualitatively different pattern to the hypothesized relapse effect. We obtained substantial evidence for the null hypothesis in the Bayesian analyses and as the pattern of results is not in the expected direction, the failure to observe spontaneous recovery is not likely to be due to a problem with inadequate power. This failure is also very unlikely to be due to the manipulation itself as the time delay was sufficient to induce spontaneous recovery of expectancy conditioning. Together with the results of Experiment 1, this experiment provides strong evidence that evaluative conditioning is not subject to spontaneous recovery.

Stream 2 – Reinstatement

Experiment 3a – Between Participant Conditioning Design

Method

Participants. 48 (32 female) undergraduate Curtin university students or community volunteers aged between 18 and 65 (*M* = 23.81, *SD* = 9.94) volunteered participation in exchange for course credit of monetary compensation or $15 AU. The research protocol was approved by the Curtin University Human Research Ethics Committee. Sample size was calculated as in Experiment 1. Participants were randomly assigned to the positive (n = 24) or negative conditioning group (n = 24).

Apparatus/Stimuli and Procedure. After the last extinction block, the 6 valenced US pictures (pleasant in positive conditioning group and unpleasant in negative conditioning group) were presented one at a time in a randomised stream separated by 2 s intervals (reinstatement manipulation). After the reinstatement manipulation, participants were asked to evaluate CS valence and US expectancy again. The remainder of the apparatus/stimuli and procedure was the same as Experiment 1.

Data Preparation and Analyses. As in Experiment 1.

Results

Preliminary Analyses. No differences between the groups were detected in the preliminary analyses. The means, standard deviations, and comparison statistics are presented in Table 1. Five participants were unable to verbalize the experimental contingencies. Missing responses resulted in the loss of 2 cases for CS valence[6] (positive conditioning: 1; negative conditioning: 1) and 9 cases for US expectancy (positive conditioning: 5; negative conditioning: 4) during acquisition; 6 cases for CS valence (positive conditioning: 5; negative conditioning: 1) and 3 cases for US expectancy (positive conditioning: 3; negative conditioning: 0) during extinction; and 1 case for CS valence (negative conditioning group) and 1 case for US expectancy (positive conditioning group) during reinstatement.

Acquisition. The CS valence evaluations and US expectancy ratings measured during Experiment 3a are presented in Figures 5 and 6, respectively. Overall, the analyses reveal that

[6] All participants provided data during the last block of acquisition and therefore the independent samples t-test comparing the equated valence difference score at the end of acquisition was conducted will all available data.

both conditional valence and expectancy were acquired during acquisition in both conditioning groups confirming that the acquisition phase was successful.

CS Valence. A main effect of block, $F(3, 42) = 3.37$, $p = .027$, $\eta p^2 = .194$, $BF_{10} = 0.19$, a marginal Block × Group interaction, $F(3, 42) = 2.80$, $p = .052$, $\eta p^2 = .167$, $BF_{inc} = 0.13$, and a CS × Group interaction, $F(1, 44) = 15.74$, $p < .001$, $\eta p^2 = .263$, $BF_{inc} = 8.01 \times 10^{10}$, were moderated by a CS × Group × Block interaction, $F(3, 42) = 4.89$, $p = .005$, $\eta p^2 = .259$, $BF_{inc} = 10.79$. In the positive conditioning group, conditional valence was not present during block 1, $F(1, 44) = 0.11$, $p = .746$, $\eta p^2 = .002$, $BF_{10} = 0.23$, was marginal during block 2, $F(1, 44) = 2.90$, $p = .096$, $\eta p^2 = .062$, $BF_{10} = 2.46$, and significant during blocks 3, $BF_{10} = 5.79$, and 4, $BF_{10} = 23.63$, both $F(1, 44) > 5.43$, $p < .025$, $\eta p^2 > .109$. Conditional valence increased in size from block 1 to 4, $p = .011$, $BF_{10} = 4.66$. In the negative conditioning group, conditional valence was not present during block 1, $F(1, 44) = 1.70$, $p = .199$, $\eta p^2 = .037$, $BF_{10} = 0.43$, but was present during all other blocks, all $F(1, 44) > 9.75$, $p < .004$, $\eta p^2 > .181$, BF_{10} : blocks 2 = 2.92, 3 = 6.04, 4 = 8.32, and increased in magnitude from block 1 to 4, $p = .009$, $BF_{10} = 3.45$. The magnitude of conditional valence at the end of acquisition did not differ between the groups (positive: $M = 1.50$, $SD = 2.06$; negative: $M = 2.25$, $SD = 3.27$), $t(38.79) = 0.95$, $p = .348$, $BF_{10} = 0.42$. Similar to the results of Experiment 1, the negative conditioning group evaluated the CSn as more pleasant than the positive conditioning group during blocks 3, $BF_{10} = 1.88$, and 4, $BF_{10} = 1.77$, both $F(1, 44) > 4.55$, $p < .039$, $\eta p^2 > .093$. The remaining omnibus effects did not reach significance, all $F < 0.75$, $p > .393$, $\eta p^2 < .017$, BF < 0.37.

US Expectancy. A main effect of CS, $F(1, 37) = 269.02$, $p < .001$, $\eta p^2 = .879$, $BF_{10} = 1.42 \times 10^{73}$, a main effect of group, $F(1, 37) = 5.09$, $p = .030$, $\eta p^2 = .121$, $BF_{10} = 0.36$, a CS × Group interaction, $F(1, 37) = 6.96$, $p = .012$, $\eta p^2 = .158$, $BF_{inc} = 1839.16$, and a CS × Block interaction, $F(3, 35) = 16.83$, $p < .001$, $\eta p^2 = .591$, $BF_{inc} = 1.49 \times 10^7$, were detected. The CS × Group interaction revealed that conditional expectancy was present during acquisition in both groups, both $F(1, 37) > 92.36$, $p < .001$, $\eta p^2 > .713$, $BF_{10} > 1.09 \times 10^5$, but the magnitude of this difference was larger in the negative conditioning group, $BF_{10} = 4.29$. The CS × Block interaction confirmed that conditional expectancy was present during all blocks, all $F(1, 37) > 39.70$, $p < .001$, $\eta p^2 > .517$, $BF_{10} > 7.46 \times 10^4$, but increased from block 1 to 4, $p < .001$, $BF_{10} = 1.17 \times 10^6$. The remaining omnibus effects did not reach significance, all $F < 1.67$, $p > .192$, $\eta p^2 < .125$, BF < 0.11.

Extinction. Overall, extinction learning did occur, with the analyses revealing that the conditional valence and expectancy acquired during acquisition reduced in both conditioning groups.

CS Valence. A CS × Group interaction, $F(1, 40) = 6.40$, $p = .015$, $\eta p^2 = .138$, $BF_{inc} = 1.30 \times 10^5$, and a marginal Group × Block interaction, $F(8, 33) = 2.00$, $p = .078$, $\eta p^2 = .327$, $BF_{inc} = 320.45$, were moderated by a CS × Group × Block interaction, $F(8, 33) = 3.84$, $p = .003$, $\eta p^2 = .482$, $BF_{inc} = 110.34$. In the positive conditioning group, conditional valence was present after the last block of acquisition, $F(1, 40) = 7.18$, $p = .011$, $\eta p^2 = .152$, $BF_{10} = 11.68$, and during block 3, $F(1, 40) = 4.78$, $p = .035$, $\eta p^2 = .107$, $BF_{10} = 1.02$, was marginal during blocks 1, $F(1, 40) = 3.31$, $p = .076$, $\eta p^2 = .076$, $BF_{10} = 1.09$, and 2, $F(1, 40) = 3.00$, $p = .091$, $\eta p^2 = .070$, $BF_{10} = 0.79$, and was not present during the remaining blocks, $F(1, 40) < 2.69$, $p > .109$, $\eta p^2 < .063$, BF_{10}: blocks 4 = 0.45, 5 = 0.47, 6 = 1.14, 7 = 0.27, 8 = 0.29. Conditional valence reduced from after acquisition to after extinction, $p = .010$, $BF_{10} = 16.72$. In the negative conditioning group, conditional valence was present after the last block of acquisition, $F(1, 40) = 12.26$, $p = .001$, $\eta p^2 = .235$, $BF_{10} = 8.32$, was marginal during block 1, $F(1, 40) = 3.34$, $p = .075$, $\eta p^2 = .077$, $BF_{10} =$

0.81, and was not present during the remaining blocks, all $F(1, 40) < 2.69$, $p > .109$, $\eta p^2 < .063$, BF_{10}: blocks 2 = 0.24, 3 = 0.47, 4 = 0.32, 5 = 0.22, 6 = 0.56, 7 = 0.25, 8 = 0.22. Conditional valence reduced from after acquisition to after extinction, $p < .001$, $BF_{10} = 18.21$. The negative conditioning group evaluated the CSn as more pleasant than the positive conditioning group during extinction blocks 1, 3, 5, 7, and 8, all $F(1, 40) > 4.36$, $p < .044$, $\eta p^2 > .098$, BF_{10}: blocks 1 = 2.91, 3 = 6.64, 5 = 2.26, 7 = 1.75, 8 = 1.66. The remaining main effects and interactions did not reach significance, all $F < 2.30$, $p > .137$, $\eta p^2 < .055$, BF < 0.83.

US Expectancy. A main effect of CS, $F(1, 43) = 84.55$, $p < .001$, $\eta p^2 = .663$, $BF_{10} = 4.49 \times 10^{52}$, a main effect of block, $F(8, 36) = 3.89$, $p = .002$, $\eta p^2 = .464$, $BF_{10} = 9.50 \times 10^4$, a CS × Group interaction, $F(1, 43) = 9.50$, $p = .004$, $\eta p^2 = .181$, $BF_{inc} = 3.45 \times 10^8$, and a CS × Block interaction, $F(8, 36) = 21.79$, $p < .001$, $\eta p^2 = .829$, $BF_{inc} = 5.61 \times 10^{35}$, were detected. The CS × Group interaction revealed that conditional expectancy was present in the positive conditioning, $BF_{10} = 835.90$, and the negative conditioning group, $BF_{10} = 2.02 \times 10^5$, both $F(1, 43) > 17.51$, $p < .001$, $\eta p^2 > .289$, but the magnitude of this was larger in the negative conditioning group, $BF_{10} = 11.04$. The CS × Block interaction confirmed that conditional expectancy was present after acquisition, $BF_{10} = 1.06 \times 10^{19}$, and during all extinction blocks, all $F(1, 43) > 7.21$, $p < .011$, $\eta p^2 > .143$, BF_{10}: blocks 1 = 1.06×10^9, 2 = 5617.76, 3 = 4603.31, 4 = 234.32, 5 = 108.13, 6 = 91.72, 7 = 4.25, 8 = 14.33, but the magnitude of this decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 5.42 \times 10^{13}$. The remaining omnibus effects did not reach significance, all $F < 2.04$, $p > .160$, $\eta p^2 < .046$, BF < 0.50.

Reinstatement. The analyses reveal that reinstatement of expectancy learning occurred in both groups but while, reinstatement of evaluative learning was present in the negative conditioning group, it was not evident in the positive conditioning group.

CS Valence. A main effect of group, $F(1, 45) = 5.05$, $p = .030$, $\eta p^2 = .101$, $BF_{10} = 2.35$, a CS × Group interaction, $F(1, 45) = 5.44$, $p = .024$, $\eta p^2 = .108$, $BF_{inc} = 5.67$, and a Group × Phase interaction, $F(1, 45) = 4.86$, $p = .033$, $\eta p^2 = .097$, $BF_{inc} = 3.09$, were moderated by a CS × Group × Phase interaction, $F(1, 45) = 10.15$, $p = .003$, $\eta p^2 = .184$, $BF_{inc} = 1.05$. Conditional valence was not present after the last block of extinction in the positive, $F(1, 45) = 0.96$, $p = .332$, $\eta p^2 = .021$, $BF_{10} = 0.37$, or the negative conditioning group, $F(1, 45) = 0.11$, $p = .740$, $\eta p^2 = .002$, $BF_{10} = 0.23$. After reinstatement, however, conditional valence was present in the negative conditioning group, $F(1, 45) = 7.09$, $p = .011$, $\eta p^2 = .136$, $BF_{10} = 1.43$, but not in the positive conditioning group, $F(1, 45) = 1.82$, $p = .184$, $\eta p^2 = .039$, $BF_{10} = 1.18$. The magnitude of differential valence increased from after extinction to the reinstatement test in the negative conditioning group, $F(1, 45) = 11.42$, $p = .002$, $\eta p^2 = .202$, $BF_{10} = 4.68$, but did not change in the positive conditioning group, $F(1, 45) = 1.22$, $p = .276$, $\eta p^2 = .026$, $BF_{10} = 0.56$. The remaining main effects and interactions did not reach significance, all $F < 2.71$, $p > .107$, $\eta p^2 < .057$, $BF < 0.32$.

US Expectancy. Main effects of CS, $F(1, 45) = 49.82$, $p < .001$, $\eta p^2 = .525$, $BF_{10} = 7.50 \times 10^7$, and phase, $F(1, 45) = 21.97$, $p < .001$, $\eta p^2 = .328$, $BF_{10} = 100.99$, a CS × Group interaction, $F(1, 45) = 11.84$, $p = .001$, $\eta p^2 = .208$, $BF_{inc} = 160.91$, and a CS × Phase interaction, $F(1, 45) = 25.59$, $p < .001$, $\eta p^2 = .362$, $BF_{inc} = 182.95$, were moderated by a CS × Group × Phase interaction, $F(1, 45) = 4.50$, $p = .039$, $\eta p^2 = .091$, $BF_{inc} = 0.90$. In the positive conditioning group, conditional expectancy was not present after the last block of extinction, $F(1, 45) = 1.17$, $p = .286$, $\eta p^2 = .025$, $BF_{10} = 1.19$, but was during the reinstatement test, $F(1, 45) = 8.37$, $p = .006$, $\eta p^2 = .157$, $BF_{10} = 42.36$, and increased from after extinction to reinstatement[7], $F(1, 45) = 4.23$, $p =$

[7] This difference is marginal in the sample of contingency verbalizers, $F(1, 40) = 3.88$, $p = .056$, $\eta p^2 = .088$, $BF_{10} = 2.25$.

.046, $\eta p^2 = .086$, $BF_{10} = 3.26$. In the negative conditioning group, conditional expectancy was present after the last block of extinction, $F(1, 45) = 14.71$, $p < .001$, $\eta p^2 = .246$, $BF_{10} = 6.53$, and after reinstatement, $F(1, 45) = 65.49$, $p < .001$, $\eta p^2 = .593$, $BF_{10} = 2.68 \times 10^4$, but the magnitude of this conditional expectancy increased from after extinction to reinstatement, $F(1, 45) = 26.33$, $p < .001$, $\eta p^2 = .369$, $BF_{10} = 141.42$. The remaining omnibus effects did not reach significance, all $F < 2.93$, $p > .094$, $\eta p^2 < .062$, BF < 0.82.

US Valence. The mean US valence evaluations are displayed in Table 2. Main effects of valence, $F(2, 45) = 873.09$, $p < .001$, $\eta p^2 = .975$, $BF_{10} = 7.81 \times 10^{60}$, and group, $F(1, 46) = 25.35$, $p < .001$, $\eta p^2 = .355$, $BF_{10} = 0.47$, were moderated by a Valence × Group interaction, $F(2, 45) = 5.18$, $p = .009$, $\eta p^2 = .187$, $BF_{inc} = 69.49$. Both groups evaluated the positive pictures as more pleasant than the neutral pictures and the negative pictures as more unpleasant than the neutral pictures, all $p < .001$, $BF_{10} > 7.74 \times 10^5$, however, while, the valence of the negative images, $F(1, 46) = 0.94$, $p = .337$, $\eta p^2 = .020$, $BF_{10} = 0.42$, did not differ between the groups, the negative conditioning group evaluated the neutral pictures, $F(1, 46) = 19.05$, $p < .001$, $\eta p^2 = .293$, $BF_{10} = 293.10$, as more pleasant than the positive conditioning group. Similarly, the negative conditioning group also evaluated the positive pictures as marginally more pleasant than the positive conditioning group, $F(1, 46) = 3.33$, $p = .075$, $\eta p^2 = .067$, $BF_{10} = 1.09$.

Discussion

In Experiment 3a, we examined whether evaluative and expectancy conditioning would relapse after reinstatement. After extinction, the valenced US images were presented one at a time in a randomized stream and CS valence and US expectancy was tested. Conditional valence and expectancy learning were successfully acquired in both groups. Similarly, both groups showed a reduction in the size of conditional valence and expectancy throughout extinction. As in Experiment 1, evaluations of the CSn during the latter half of acquisition were more pleasant in the negative than the positive conditioning group. This difference was also present during extinction, where we observed a fast and complete extinction of differential CS valence evaluations in both groups but overall more pleasant evaluations of the CSn in the negative conditioning group than in the positive conditioning group. Interestingly, CSv valence extinguished to the level of the CSn in both groups, suggesting that the CSv extinction was anchored to the CSn. Replicating the US valence results from Experiment 1, participants in the negative conditioning group evaluated the neutral US pictures as more pleasant than participants in the positive conditioning group – suggesting that the shift in CSn valence is likely driven by this change in US valence. The difference in neutral US valence could occur because the valence of the neutral pictures has intrinsically changed or it could occur because of the manner in which participants report the US valence has changed between the groups (i.e. it is a response bias).

We observed reinstatement of expectancy conditioning in both groups, with the magnitude of conditional expectancy ratings increasing after the stream of US images. We also observed reinstatement of evaluative conditioning in the negative conditioning group, where conditional valence was not present after extinction but re-emerged after reinstatement. Unexpectedly, reinstatement was not detected in the positive conditioning group. It is possible that reinstatement of evaluative conditioning only occurs after negative conditioning but as conditional valence for the positive conditioning group was numerically larger after reinstatement and we did not receive strong evidence for the null or for the alternative hypothesis in the Bayesian analyses, it seems more likely that the failure to find reinstatement in the positive conditioning group is due to a problem with experimental power. As our power analyses was conducted based on finding a large effect for the overall omnibus interaction, our experiment is underpowered to detect the effect of reinstatement individually in the conditioning groups. To confirm whether evaluative learning is subject to reinstatement after positive conditioning we decided to re-run the experiment with a larger sample size.

Experiment 3b – Between Participant Conditioning Design (Replication)

Experiment 3b was conducted to examine whether evaluative conditioning is subject to reinstatement after positive conditioning. To further explore the change in valence of the neutral US images we included a novel set of neutral US images into the rating task at the end of the experiment.

Method

Participants. Two hundred and seventy-one participants from the United States (108[8] female) aged between 22 and 72 (*M* = 36.58, *SD* = 10.86) were recruited via Amazon's Mechanical Turk (Buhrmester, Kwang, & Gosling, 2011). Participants volunteered participation in exchange for $4.50 US**.** The research protocol was approved by the Curtin University Human Research Ethics Committee. An a-priori power analysis was conducted using the parameters described in Experiment 1 but based on detecting a medium effect (*f* = .25) for the critical relapse interaction, which revealed that 128 participants would be required. However, an additional power analysis based on an ANOVA *F* test (*f* = .25; repeated measures, within factors; 1 group [Positive/Negative Conditioning]; 2 measurement points [Last Block of Extinction; Relapse Test]; Effect size specification: as in Cohen (1988); nonsphericity correction: 1) revealed that 128 participants would be required per group (256 total) to be 80% confident of detecting a

[8] Six participants did not provide demographic information.

significant effect in the relapse follow-up analyses. Participants were randomly assigned to the positive (n = 134) or the negative (n = 137) conditioning group.

Apparatus/Stimuli

The experiment was programmed using Inquisit 4 software, hosted on the Inquisit millisecond server, and run on Amazon Mechanical Turk using the Turkprime platform (Litman, Robinson, & Abberbock, 2016). The CSs were presented centred at a size of 50% of the participants' screen. USs were presented centred at a size of 60% of the participants screen. The novel US pictures were 6 neutral IAPS images (codes: 7004, 7009, 7030, 7050, 7235, 7950; mean valence rating = 4.83; mean arousal rating = 2.52; subject matter: household items such as furniture and crockery). Participants were asked to rate CS and US valence on a 1-9 Likert scale with the anchors (1 = *unpleasant,* 5 = *neutral,* and 9 = *pleasant*). US expectancy was measured by asking participants to predict what type of pictures would follow the CSs on a 1-7 scale with the anchors (1 = *neutral pictures will always follow, 4 = no pictures will follow,* 7 = *pleasant (or unpleasant) pictures will always follow).* The response option for 7 was 'pleasant' in the positive conditioning group and 'unpleasant' in the negative conditioning group. All other aspects of the apparatus/stimuli were the same as Experiment 1. The reinstatement manipulation was presented as in Experiment 3a.

Procedure. After participants had evaluated all of the USs used during acquisition they were asked to evaluate the valence of the 6 novel neutral images. All other aspects of the procedure were the same as in Experiment 3a.

Data Preparation and Analyses. The valence ratings for the pleasant, neutral, novel neutral and unpleasant USs were averaged across individual pictures and the mean US evaluations were subjected to a 4 Valence (positive, neutral, novel neutral, negative) × 2 Group (positive conditioning, negative conditioning) factorial ANOVA. All other data preparation and analyses were performed as in Experiment 1.

Results

Preliminary Analyses. There was a marginally higher fail to pass ratio in the positive conditioning group. No other differences between the groups were detected in the preliminary analyses. The means, standard deviations, and comparison statistics are presented in Table 1. Sixty-six participants were unable to verbalize the experimental contingencies and 3 participants did not answer the contingency question.

Acquisition. The CS valence evaluations and US expectancy ratings recorded during Experiment 3b are presented in Figures 7 and 8, respectively. Conditional valence and expectancy conditioning were successfully acquired in both groups.

CS Valence. Main effects of CS, $F(1, 269) = 13.01$, $p < .001$, $\eta p^2 = .046$, $BF_{10} = 6.83 \times 10^5$, block, $F(3, 267) = 4.97$, $p = .002$, $\eta p^2 = .053$, $BF_{10} = 0.08$, and group, $F(1, 269) = 74.28$, $p < .001$, $\eta p^2 = .216$, $BF_{10} = 7.99 \times 10^{12}$, a CS × Group interaction, $F(1, 269) = 244.24$, $p < .001$, $\eta p^2 = .476$, $BF_{inc} = 1.58 \times 10^{152}$, and a Block × Group interaction, $F(3, 267) = 25.77$, $p < .001$, $\eta p^2 = .225$, $BF_{inc} = 1.99 \times 10^{14}$, were moderated by a CS × Group × Block interaction, $F(3, 267) = 34.23$, $p < .001$, $\eta p^2 = .278$, $BF_{inc} = 1.09 \times 10^{14}$. In both groups, conditional valence was present during all blocks of acquisition, all $F(1, 269) > 12.81$, $p < .001$, $\eta p^2 > .045$, BF_{10}: positive conditioning block 1 = 112.10, all other blocks > 1.26×10^8; negative conditioning all blocks > 1.56×10^7, and increased in size from block 1 to 4, both $p < .001$, $BF_{10} > 2.47 \times 10^7$. The magnitude of conditional valence at the end of acquisition was larger in the negative conditioning group ($M = 2.70$, $SD = 2.51$) than in the positive conditioning group ($M = 1.78$, $SD = 2.16$), $t(269) = 3.22$, $p = .001$, $BF_{10} = 17.20$. The negative conditioning group evaluated the

CSn as more pleasant than the positive conditioning group during all blocks, all $F(1, 269) > 4.08$, $p < .045$, $\eta p^2 > .014$. BF_{10}: blocks 1 = 6.41, 2 = 0.92, 3 = 10.28, 4 = 2.97. The CS × Block interaction did not reach significance, $F(3, 267) = 1.26$, $p = .288$, $\eta p^2 = .014$, $BF_{inc} = 0.01$.

US Expectancy. A main effect of CS, $F(1, 269) = 551.44$, $p < .001$, $\eta p^2 = .672$, $BF_{10} = 5.81 \times 10^{325}$, a main effect of group, $F(1, 269) = 6.41$, $p = .012$, $\eta p^2 = .023$, $BF_{10} = 0.52$, a CS × Group interaction[9], $F(1, 269) = 8.12$, $p = .005$, $\eta p^2 = .029$, $BF_{inc} = 2.28 \times 10^6$, and a CS × Block interaction, $F(3, 267) = 34.46$, $p < .001$, $\eta p^2 = .279$, $BF_{inc} = 3.03 \times 10^{18}$, were detected. The CS × Group interaction revealed that conditional expectancy was present in both groups, both $F(1, 269) > 210.54$, $p < .001$, $\eta p^2 > .439$, $BF_{10} > 3.16 \times 10^{23}$, but the magnitude of this conditional expectancy was larger in the negative conditioning group, $BF_{10} = 6.07$. The CS × Block interaction confirmed that conditional expectancy was present during all blocks, all $F(1, 269) > 154.19$, $p < .001$, $\eta p^2 > .364$, $BF_{10} > 1.25 \times 10^{25}$, but the size of this increased from block 1 to 4, $p < .001$, $BF_{10} = 4.26 \times 10^{17}$. The remaining omnibus effects did not reach significance, all $F < 1.90$, $p > .130$, $\eta p^2 < .021$, $BF < 0.04$.

Extinction. The conditional valence and expectancy conditioning acquired throughout acquisition reduced throughout extinction in both groups.

CS Valence. Main effects of CS, $F(1, 269) = 5.74$, $p = .017$, $\eta p^2 = .021$, $BF_{10} = 9277.49$, and group, $F(1, 269) = 7.40$, $p = .007$, $\eta p^2 = .027$, $BF_{10} = 4.31$, and CS × Group, $F(1, 269) = 103.89$, $p < .001$, $\eta p^2 = .279$, $BF_{inc} = 1.11 \times 10^{98}$, Block × Group, $F(8, 262) = 17.82$, $p < .001$, $\eta p^2 = .352$, $BF_{inc} = 5.94 \times 10^{43}$, and CS × Block, $F(8, 262) = 2.55$, $p = .011$, $\eta p^2 = .072$, $BF_{inc} = 0.03$, interactions were moderated by a CS × Group × Block interaction, $F(8, 262) = 24.23$, $p <$

[9] This interaction only attains marginal significance in the contingency verbalizers, $F(1, 200) = 3.24$, $p = .074$, $\eta p^2 = .016$, $BF_{inc} = 43.26$, but the BF_{inc} value suggests it is a reliable effect.

.001, $\eta p^2 = .425$, $BF_{inc} = 1.50 \times 10^{72}$. In the positive conditioning group, conditional valence was present after the last block of acquisition, $BF_{10} = 6.23 \times 10^{13}$, and during extinction blocks 1, 2, 3, 4, 5, and 8, all $F(1, 269) > 5.02$, $p < .026$, $\eta p^2 > .018$, BF_{10}: blocks 1 = 1.18×10^4, 2 = 43.30, 3 = 709.40, 4 = 42.20, 5 = 2.69, 8 = 5.52, was marginal during block 7, $F(1, 269) = 3.74$, $p = .054$, $\eta p^2 = .014$, $BF_{10} = 0.89$, and did not attain significance during block 6, $F(1, 269) = 1.75$, $p = .187$, $\eta p^2 = .006$, $BF_{10} = 0.25$. Conditional valence decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 1.90 \times 10^9$. In the negative conditioning group, conditional valence was present after the last block of acquisition, $BF_{10} = 2.92 \times 10^{21}$, and during all extinction blocks, all $F(1, 269) > 8.74$, $p < .004$, $\eta p^2 > .031$, BF_{10}: blocks 1 = 8.72×10^6, 2 = 4.40×10^6, 3 = 3272.68, 4 = 13.34, 5 = 2.54, 6 = 681.31, 7 = 25.24, 8 = 150.77. The magnitude of this difference decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 5.39 \times 10^{13}$. The negative conditioning group evaluated the CSn as more pleasant than the positive conditioning group during blocks 1, $BF_{10} = 17.93$, and 2, $BF_{10} = 1.10$, all $F(1, 269) > 4.46$, $p < .036$, $\eta p^2 > .016$. The main effect of block did not reach significance, all $F(8, 262) = 0.99$, $p = .443$, $\eta p^2 = .029$, $BF_{10} = 1.72 \times 10^{-6}$.

US Expectancy. Main effects of CS, $F(1, 269) = 152.09$, $p < .001$, $\eta p^2 = .361$, $BF_{10} = 2.52 \times 10^{115}$, and group, $F(1, 269) = 5.98$, $p = .015$, $\eta p^2 = .022$, $BF_{10} = 1.41$, were moderated by a marginal CS × Group interaction[10], $F(1, 269) = 3.59$, $p = .059$, $\eta p^2 = .013$, $BF_{inc} = 436.01$, and a CS × Block interaction, $F(8, 262) = 69.65$, $p < .001$, $\eta p^2 = .680$, $BF_{inc} = \infty$. Conditional expectancy was present in both the positive, $BF_{10} = 5.01 \times 10^8$, and the negative conditioning groups, $BF_{10} = 4.76 \times 10^{24}$, both $F(1, 269) > 53.88$, $p < .001$, $\eta p^2 > .166$, but although the marginal CS × Group interaction suggests the groups differ in the size of conditional expectancy,

[10] This interaction was not significant in the contingency verbalizers, $F(1, 200) = 0.60$, $p = .441$, $\eta p^2 = .003$, $BF_{inc} = 0.38$.

the Bayesian analyses were indecisive, BF_{10} = 0.73. The CS × Block interaction confirmed that conditional expectancy was present after the last block of acquisition, $BF_{10} = 3.25 \times 10^{65}$, and during extinction blocks 1, 2, 3, 4, and 5, all $F(1, 269) > 9.75$, $p < .002$, $\eta p^2 > .034$, BF_{10}: blocks 1 = 4.92×10^{10}, 2 = 6.66×10^{6}, 3 = 6090.92, 4 = 8.02, 5 = 24.08. During blocks 6, $F(1, 269) = 3.84$, $p = .051$, $\eta p^2 = .014$, $BF_{10} = 0.46$, 7, $F(1, 269) = 4.07$, $p = .045$, $\eta p^2 = .015$, $BF_{10} = 0.51$, and 8, $F(1, 269) = 2.86$, $p = .092$, $\eta p^2 = .011$, $BF_{10} = 0.28$, the difference between the CSv and CSn was significant or marginal but the BF_{10} provided evidence for the null hypothesis. Conditional expectancy reduced from after acquisition to after extinction, $p < .001$, $BF_{10} = 1.39 \times 10^{63}$. The remaining omnibus effects did not reach significance, all $F < 1.33$, $p > .232$, $\eta p^2 < .039$, BF < 19.56.

Reinstatement. Conditional expectancy learning relapsed after the reinstatement manipulation and relapse of evaluative learning was observed in the negative conditioning group. Evidence for relapse of evaluative learning was also observed in the positive conditioning group but the relapse effect in this group only attained marginal significance.

CS Valence. A main effect of group, $F(1, 269) = 13.70$, $p < .001$, $\eta p^2 = .048$, BF_{10} = 80.95, a CS × Group interaction, $F(1, 269) = 50.53$, $p < .001$, $\eta p^2 = .158$, $BF_{inc} = 1.84 \times 10^{13}$, and a Group × Phase interaction, $F(1, 269) = 36.41$, $p < .001$, $\eta p^2 = .119$, $BF_{inc} = 8.50 \times 10^{6}$, were moderated by a CS × Group × Phase interaction, $F(1, 269) = 12.85$, $p < .001$, $\eta p^2 = .046$, BF_{inc} = 4.23. In the positive conditioning group, conditional valence was present after the last block of extinction, $F(1, 269) = 7.62$, $p = .006$, $\eta p^2 = .028$, $BF_{10} = 5.52$, and after reinstatement, $F(1, 269) = 13.06$, $p < .001$, $\eta p^2 = .046$, $BF_{10} = 134.03$. The magnitude of this conditional valence increased marginally after reinstatement, $F(1, 269) = 2.91$, $p = .089$, $\eta p^2 = .011$, $BF_{10} = 0.43$. When participants who failed the contingency assessment were removed the relapse effect was strengthened, $F(1, 200) = 3.83$, $p = .052$, $\eta p^2 = .019$, $BF_{10} = 1.77$. In the negative conditioning group, conditional valence was present after the last block of extinction, $F(1, 269) = 17.55$, $p < .001$, $\eta p^2 = .061$, $BF_{10} = 150.77$, and after reinstatement, $F(1, 269) = 38.67$, $p < .001$, $\eta p^2 = .126$, $BF_{10} = 2.15 \times 10^5$. The magnitude of this conditional valence increased after reinstatement, $F(1, 269) = 11.38$, $p = .001$, $\eta p^2 = .041$, $BF_{10} = 14.96$. The remaining main effects and interactions did not reach significance, all $F < 2.91$, $p > .089$, $\eta p^2 < .011$, BF < 0.47.

US Expectancy. Main effects of CS, $F(1, 269) = 87.53$, $p < .001$, $\eta p^2 = .246$, $BF_{10} = 1.68 \times 10^{19}$, phase, $F(1, 269) = 37.80$, $p < .001$, $\eta p^2 = .123$, $BF_{10} = 6725.78$, and group, $F(1, 269) = 8.38$, $p = .004$, $\eta p^2 = .030$, $BF_{10} = 3.94$, a Phase × Group interaction, $F(1, 269) = 5.28$, $p = .022$, $\eta p^2 = .019$, $BF_{inc} = 0.70$, and a CS × Phase interaction, $F(1, 269) = 87.47$, $p < .001$, $\eta p^2 = .245$, 5.50×10^{15}, were detected. The Phase × Group interaction reflects that during the last block of extinction, $F(1, 269) = 1.99$, $p = .160$, $\eta p^2 = .007$, $BF_{10} = 0.34$, expectancy scores did not differ between the groups, but after reinstatement, expectancy scores were higher in the pleasant conditioning group, $F(1, 269) = 10.58$, $p = .001$, $\eta p^2 = .038$, $BF_{10} = 19.00$. These scores are collapsed across CSv and CSn (rather than conditional expectancy scores i.e. CSv – CSn) and therefore this interaction is not informative about conditioning or relapse. The CS × Phase interaction provides evidence for relapse after the reinstatement manipulation. Conditional expectancy was marginal during the last block of extinction, $F(1, 269) = 2.86$, $p = .092$, $\eta p^2 = .011$, $BF_{10} = 0.28$, but increased from after extinction to reinstatement, $BF_{10} = 1.49 \times 10^{15}$, and was present during the reinstatement test, $F(1, 269) = 108.07$, $p < .001$, $\eta p^2 = .287$, $BF_{10} = 2.33 \times 10^{18}$. The remaining omnibus effects did not reach significance, all $F < 3.08$, $p > .080$, $\eta p^2 < .012$, BF < 0.70.

US Valence. The mean US valence evaluations are displayed in Table 2. A main effect of valence, $F(3, 267) = 539.68$, $p < .001$, $\eta p^2 = .858$, $BF_{10} = 1.02 \times 10^{346}$, confirmed that unpleasant USs were evaluated as less pleasant than neutral USs, novel neutral pictures and that positive USs were evaluated as more pleasant than neutral USs and novel neutral pictures, all $p < .001$, $BF_{10} > 2.01 \times 10^{95}$. Evaluations of neutral USs and novel neutral pictures did not differ, $p = .559$, $BF_{10} = 0.08$. The remaining omnibus effects did not reach significance[11], all $F < 1.75$, $p > .159$, $\eta p^2 < .020$, BF < 0.09.

Discussion

In Experiment 3b we examined whether positive conditional valence would be sensitive to reinstatement in a larger sample. Conditional valence and expectancy learning were acquired throughout acquisition in both groups, and this learning reduced during extinction. As in Experiment 3a, relapse of conditional expectancy learning was observed after reinstatement. Also consistent with the results of Experiment 3a, there was a significant increase in conditional valence evaluations after the reinstatement manipulation in the negative conditioning group. In the positive conditioning group, there was evidence that reinstatement occurred but the effect was marginal and smaller than that for negative conditioning. Anecdotal support for the alternative hypothesis was obtained in the Bayesian analyses only in those participants who were able to verbalize the experimental contingencies. Conditional valence was stronger in the negative conditioning group at the end of acquisition and this difference in the strength of the acquisition learning likely accounts for the difference in the size of reinstatement. As in Experiments 1 and 3a, the negative conditioning group evaluated the CSn as more pleasant

[11] In the sample of participants who passed the contingency assessment, the Valence × Group interaction, $F(3, 198) = 3.54$, $p = .016$, $\eta p^2 = .051$, $BF_{inc} = 0.05$, attains significance but the BF_{inc} provides support for the null hypothesis.

during acquisition. In contrast with the results of Experiment 3a, however, this difference was less pronounced during extinction and the pattern of US valence did not differ between the conditioning groups.

Experiment 4 – Within Participant Conditioning Design

Method

Participants. 141 participants from the United States (66[12] female) aged between 19 and 79 (*M* = 36.95, *SD* = 11.05) were recruited via Amazon's Mechanical Turk. Participants volunteered participation in exchange for $4.50 US. The research protocol was approved by the Curtin University Human Research Ethics Committee. An a-priori power analysis was conducted using the parameters described in Experiment 2 but based on detecting a medium effect (f = .25) for the critical relapse interaction, which revealed that 128 participants would be required.

Apparatus/Stimuli and Procedure. After the last extinction measurement phase, the 12 valenced US pictures (6 pleasant and 6 unpleasant) were presented separated by 2 s intervals one at a time in a randomized stream (reinstatement manipulation) and participants were asked to evaluate CS valence and US expectancy again. After the reinstatement assessment, participants evaluated US valence, provided demographic information, and completed the contingency assessment. The remaining apparatus/stimuli and procedure were conducted as in Experiment 2.

Data Preparation and Analyses. The data preparation and analyses were conducted as in Experiment 2.

Results

[12] One participants did not provide age information.

Preliminary Analyses. CSp and CSu evaluations did not differ during the baseline assessment, $F(1, 143) = 0.95$, $p = .332$, $\eta p^2 = .007$, $BF_{10} = 0.15$. Sixteen participants were unable to verbalize the experimental contingencies.

Acquisition. The CS valence evaluations and US expectancy ratings measured during Experiment 4 are presented in Figures 9 and 10, respectively. Conditional evaluative and expectancy conditioning were successfully acquired during acquisition.

CS Valence. Main effects of CS, $F(1, 143) = 293.09$, $p < .001$, $\eta p^2 = .672$, $BF_{10} = 6.43 \times 10^{177}$, and block, $F(3, 141) = 3.65$, $p = .014$, $\eta p^2 = .072$, $BF_{10} = 0.01$, were moderated by a CS × Block interaction, $F(3, 141) = 43.41$, $p < .001$, $\eta p^2 = .480$, $BF_{inc} = 1.04 \times 10^{20}$. Conditional valence was present during all acquisition blocks, all $F(1, 143) > 72.56$, $p < .001$, $\eta p^2 > .336$, $BF_{10} > 3.21 \times 10^{11}$, and the size of this difference increased from block 1 to 4, $p < .001$, $BF_{10} = 5.55 \times 10^{18}$.

US Expectancy. A main effect of CS, $F(1, 143) = 673.59$, $p < .001$, $\eta p^2 = .825$, $BF_{10} = 3.07 \times 10^{265}$, was moderated by a CS × Block interaction, $F(3, 141) = 22.11$, $p < .001$, $\eta p^2 = .320$, $BF_{inc} = 8.74 \times 10^{19}$. Conditional expectancy was present during all acquisition blocks, all $F(1, 143) > 95.20$, $p < .001$, $\eta p^2 > .399$, $BF_{10} > 3.68 \times 10^{14}$, and the size of this difference increased from block 1 to 4, $p < .001$, $BF_{10} = 7.82\times 10^{8}$. The main effect of block did not reach significance, $F(3, 141) = 1.02$, $p = .385$, $\eta p^2 = .021$, $BF_{10} = 0.00$.

Extinction. The conditional valence and expectancy learning acquired during acquisition reduced throughout extinction.

CS Valence. Main effects of CS, $F(1, 143) = 101.85$, $p < .001$, $\eta p^2 = .416$, $BF_{10} = 2.26 \times 10^{152}$, and block, $F(8, 136) = 2.05$, $p = .046$, $\eta p^2 = .107$, $BF_{10} = 3.16 \times 10^{-5}$, were moderated by a

CS × Block interaction, $F(8, 136) = 24.00$, $p < .001$, $\eta p^2 = .585$, $BF_{inc} = 3.39 \times 10^{63}$. Conditional valence was present during the last block of acquisition, $BF_{10} = 4.74 \times 10^{33}$, and during all extinction blocks, all $F(1, 143) > 25.14$, $p < .001$, $\eta p^2 > .149$, $BF_{10} > 8084.99$, but reduced in size from acquisition to after extinction, $p < .001$, $BF_{10} = 1.14 \times 10^{21}$.

US Expectancy. Main effects of CS, $F(1, 143) = 122.86$, $p < .001$, $\eta p^2 = .462$, $BF_{10} = 9.35 \times 10^{131}$, and block, $F(8, 136) = 2.49$, $p = .015$, $\eta p^2 = .128$, $BF_{10} = 5.82 \times 10^{-6}$, were moderated by a CS × Block interaction, $F(8, 136) = 36.35$, $p < .001$, $\eta p^2 = .681$, $BF_{inc} = 3.73 \times 10^{152}$. Conditional expectancy was present after acquisition, $BF_{10} = 3.30 \times 10^{49}$, and during all extinction blocks, all $F(1, 143) > 12.40$, $p < .001$, $\eta p^2 > .079$, BF_{10} blocks 1 = 6.15×10^{11}, 2 = 1.61×10^{6}, 3 = 1.17×10^{4}, 4 = 5528.35, 5 = 791.24, 6 = 77.60, 7 = 247.57, 8 = 31.83, but the size of this difference decreased from the after acquisition to after extinction, $p < .001$, $BF_{10} = 1.05 \times 10^{32}$.

Reinstatement. Both conditional valence and expectancy conditioning relapsed after the reinstatement manipulation.

CS Valence. A main effect of CS, $F(1, 143) = 59.60$, $p < .001$, $\eta p^2 = .294$, $BF_{10} = 5.88 \times 10^{24}$, was moderated by a CS × Phase interaction, $F(1, 143) = 19.98$, $p < .001$, $\eta p^2 = .123$, $BF_{inc} = 22.16$. This interaction confirmed that reinstatement of evaluative learning occurred. Conditional valence was present after extinction, $F(1, 143) = 29.63$, $p < .001$, $\eta p^2 = .172$, $BF_{10} = 5.14 \times 10^{4}$, and during the reinstatement test, $F(1, 143) = 68.12$, $p < .001$, $\eta p^2 = .323$, $BF_{10} = 7.38 \times 10^{10}$, but conditional valence was larger during the reinstatement test, $BF_{10} = 903.30$. The main effect of phase did not reach significance, $F(1, 143) = 0.53$, $p = .468$, $\eta p^2 = .004$, $BF_{10} = 0.10$.

US Expectancy. A main effect of CS, $F(1, 143) = 44.87$, $p < .001$, $\eta p^2 = .239$, $BF_{10} = 2.71 \times 10^{20}$, was moderated by a CS × Phase interaction, $F(1, 143) = 29.08$, $p < .001$, $\eta p^2 = .169$, $BF_{inc} = 9.33 \times 10^4$. This interaction confirmed that relapse of expectancy learning occurred. Conditional expectancy was present after extinction, $F(1, 143) = 12.40$, $p = .001$, $\eta p^2 = .080$, $BF_{10} = 31.83$, and after reinstatement, $F(1, 143) = 54.27$, $p < .001$, $\eta p^2 = .275$, $BF_{10} = 6.18 \times 10^8$, but the magnitude of this difference was larger after reinstatement, $BF_{10} = 4.11 \times 10^4$. The main effect of phase did not reach significance, $F(1, 143) = 0.46$, $p = .501$, $\eta p^2 = .003$, $BF_{10} = 0.01$.

US Valence. Pleasant USs ($M = 8.06$, $SD = 1.04$) were evaluated as more pleasant than unpleasant USs ($M = 1.78$, $SD = 0.93$), $F(1, 143) = 1977.39$, $p < .001$, $\eta p^2 = .933$, $BF_{10} = 6.91 \times 10^{81}$.

Discussion

In Experiment 4 we examined whether evaluative and expectancy conditioning would relapse after a reinstatement manipulation in a within participants conditioning design. Conditional valence and expectancy learning were successfully acquired during acquisition and this learning reduced throughout extinction. We observed relapse after the reinstatement manipulation for both evaluative and expectancy conditioning, with an increase in the size of both conditional CS valence and US expectancy ratings after the presentation of the US images during reinstatement. The results replicate Experiment 3a and 3b and overall provide strong evidence that both evaluative and expectancy conditioning are subject to reinstatement.

Stream 3 – Renewal

Experiment 5 – Between Participant Conditioning Design

Method

Participants. Three hundred and forty-four participants from the United States (149[13] female) aged between 18 and 73 (*M* = 35.97, *SD* = 10.34) were recruited via Amazon's Mechanical Turk. Participants volunteered participation in exchange for $4.50 US. The research protocol was approved by the Curtin University Human Research Ethics Committee. Power was determined as in Experiment 3b, however, we aimed to recruit additional participants to ensure an adequate sample of participants who passed the manipulation check (i.e. noticed the background screen color change during the experiment). Participants were randomly assigned to the positive (*n* = 167) or negative conditioning group (*n* = 177).

Apparatus/Stimuli. An ABA renewal procedure was used with grey (RGB: 120, 120, 120) and black (RGB: 000, 000, 000) background screen colors serving as the context manipulation. Background screen colour has been used to successfully induce renewal in previous fear relapse studies (see Dibbets, Havermans, & Arntz, 2008; and Dibbets, Havermans, & Arntz, 2012). The remainder of the apparatus/stimuli were the same as for Experiment 1.

Procedure. Participants completed the baseline valence assessment and acquisition training (including the inter-trial interval and the valence and expectancy assessments) on screen color 1 and extinction training and the extinction assessments on color 2. After the last extinction rating session, the screen color immediately changed back to color 1 and participants completed the renewal valence and expectancy assessment on color 1. Assignment of black/grey to color 1/color 2 was counterbalanced between participants. The post-experimental questionnaire included a 'yes' or 'no' question assessing whether participants noticed the screen color change during the experiment. The remainder of the procedure was the same as Experiment 3b.

Data Preparation and Analyses. A Pearson's chi-square test was performed to check

[13] One participant did not identify as male or female and four participants did not provide sex information. Ten participants did not provide age information.

that the ratio of participants who reported noticing/not noticing the screen color change did not differ between the groups. The remainder of the data preparation and analyses were conducted as in Experiment 3b.

Results

Preliminary Analyses. No differences between the conditioning groups were uncovered in the preliminary analyses. The means, standard deviations, and comparison statistics are presented in Table 1. Sixty-one participants were unable to verbalize the experimental contingencies and 3 participants did not answer the contingency question. Eighty-four participants (24.4%) reported that they did not notice the screen color change and three participants did not answer the screen color question. We have made note of any instance where excluding participants who did not pass the manipulation check affects the relapse results.

Acquisition. The CS valence evaluations and US expectancy ratings measured during Experiment 5 are presented in Figures 11 and 12, respectively. Conditional evaluative and expectancy conditioning were acquired in both groups.

CS Valence. Main effects of CS, $F(1, 342) = 15.29$, $p < .001$, $\eta p^2 = .043$, $BF_{10} = 2.50 \times 10^8$, block, $F(3, 340) = 4.98$, $p = .002$, $\eta p^2 = .042$, $BF_{10} = 0.04$, and group, $F(1, 342) = 111.85$, $p < .001$, $\eta p^2 = .246$, $BF_{10} = 3.12 \times 10^{19}$, and CS × Group, $F(1, 342) = 270.03$, $p < .001$, $\eta p^2 = .441$, $BF_{inc} = 4.92 \times 10^{164}$, Block × Group, $F(3, 340) = 36.15$, $p < .001$, $\eta p^2 = .242$, $BF_{inc} = 1.60 \times 10^{14}$, and CS × Block interactions, $F(3, 340) = 4.31$, $p = .005$, $\eta p^2 = .037$, $BF_{inc} = 1.65$, were moderated by a CS × Group × Block interaction, $F(3, 340) = 49.23$, $p < .001$, $\eta p^2 = .303$, $BF_{inc} = 7.79 \times 10^{25}$. In both groups, conditional valence was present during all blocks of acquisition, all $F(1, 342) > 10.70$, $p < .002$, $\eta p^2 > .030$, BF_{10}: positive conditioning: block 1 = 55.35, all other > 1.41×10^{12}; negative conditioning: block 1 = 1443.00, all other > 1.59×10^{17}, and increased from block 1 to 4, both $p < .001$, $BF_{10} > 2.71 \times 10^7$. The magnitude of conditional valence at the end of acquisition was larger in the negative conditioning group ($M = 2.95$, $SD = 2.86$) than in the positive conditioning group ($M = 1.72$, $SD = 2.39$), $t(336.90) = 4.34$, $p < .001$, $BF_{10} = 755.40$. The negative conditioning group evaluated the CSn as more pleasant than the positive conditioning group during blocks 2, 3, and 4, all $F(1, 342) > 4.89$, $p < .028$, $\eta p^2 > .014$, BF_{10}: blocks 2 = 1.24, 3 = 28.94, 4 = 463.65.

US Expectancy. A main effect of CS, $F(1, 342) = 755.37$, $p < .001$, $\eta p^2 = .688$, $BF_{10} = 4.76 \times 10^{405}$, was moderated by a CS × Block interaction, $F(3, 340) = 54.07$, $p < .001$, $\eta p^2 = .323$, $BF_{inc} = 8.83 \times 10^{34}$. Conditional expectancy was present during all blocks, all $F(1, 342) > 150.78$, $p < .001$, $\eta p^2 > .305$, $BF_{10} > 7.06 \times 10^{25}$, and increased from block 1 to 4, $p < .001$, $BF_{10} = 4.17 \times 10^{25}$. The remaining omnibus effects did not reach significance, all $F < 2.43$, $p > .120$, $\eta p^2 < .008$, BF < 6.50.

Extinction. The evaluative and expectancy conditioning acquired during acquisition reduced throughout extinction in both groups.

CS Valence. Main effects of CS, $F(1, 342) = 10.16$, $p = .002$, $\eta p^2 = .029$, $BF_{10} = 5.51 \times 10^{11}$, block, $F(8, 335) = 4.05$, $p < .001$, $\eta p^2 = .088$, $BF_{10} = 153.62$, and group, $F(1, 342) = 7.69$, $p = .006$, $\eta p^2 = .022$, $BF_{10} = 3.91$, and CS × Group, $F(1, 342) = 67.32$, $p < .001$, $\eta p^2 = .164$, $BF_{inc} = 1.23 \times 10^{81}$, Block × Group, $F(8, 335) = 20.33$, $p < .001$, $\eta p^2 = .327$, $BF_{inc} = 5.24 \times 10^{44}$, and CS × Block interactions, $F(8, 355) = 2.62$, $p = .009$, $\eta p^2 = .059$, $BF_{inc} = 0.06$, were moderated by a CS × Group × Block interaction, $F(8, 335) = 30.46$, $p < .001$, $\eta p^2 = .421$, $BF_{inc} = 2.61 \times 10^{78}$. In the positive conditioning group, conditional valence was present after the last block of acquisition, $BF_{10} = 8.82 \times 10^{13}$, and during extinction blocks 1, 2, and 3, all $F(1, 342) > 5.01$, $p < .026$, $\eta p^2 > .014$, BF_{10}: blocks 1 = 4994.81, 2 = 1.24, 3 = 1.65, was marginal in block 4,

$F(1, 342) = 3.55$, $p = .060$, $\eta p^2 = .010$, $BF_{10} = 0.65$, and was not present during the remaining blocks, all $F(1, 342) < 1.02$, $p > .314$, $\eta p^2 < .003$, BF_{10} blocks 5 = 0.16, 6 = 0.09, 7 = 0.10, 8 = 0.13. The magnitude of this difference decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 4.81 \times 10^{12}$. In the negative conditioning group, conditional valence was present after the last block of acquisition, $BF_{10} = 2.56 \times 10^{26}$, and during all extinction blocks, all $F(1, 342) > 10.96$, $p < .002$, $\eta p^2 > .031$, BF_{10}: blocks 1 = 6.64×10^5, 2 = 2.70×10^5, 3 = 591.22, 4 = 581.22, 5 = 1356.62, 6 = 7.95, 7 = 101.95, 8 = 73.27. The magnitude of this difference decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 1.92 \times 10^{20}$. The negative conditioning group evaluated the CSn as more pleasant than the positive conditioning group during blocks 1, 4, 5 and 8, all $F(1, 342) > 4.18$, $p < .042$, $\eta p^2 > .012$, BF_{10}: blocks 1 = 1.01, 4 = 1.78, 5 = 0.88, 8 = 1.32.

US Expectancy. Main effects of CS, $F(1, 342) = 192.61$, $p < .001$, $\eta p^2 = .360$, $BF_{10} = 1.38 \times 10^{136}$, and block, $F(8, 335) = 2.11$, $p = .034$, $\eta p^2 = .048$, $BF_{10} = 1.30 \times 10^{-5}$, and a marginal main effect of group, $F(1, 342) = 3.64$, $p = .057$, $\eta p^2 = .011$, $BF_{10} = 0.41$, were moderated by a CS × Group interaction[14], $F(1, 342) = 5.93$, $p = .015$, $\eta p^2 = .017$, $BF_{inc} = 5.14 \times 10^4$, and a CS × Block interaction, $F(8, 335) = 97.89$, $p < .001$, $\eta p^2 = .700$, $BF_{inc} = \infty$. The CS × Group interaction revealed that conditional expectancy was present in the positive conditioning group, $BF_{10} = 8.34 \times 10^{12}$, and the negative conditioning group, $BF_{10} = 7.50 \times 10^{17}$, both all $F(1, 342) > 63.61$, $p < .001$, $\eta p^2 > .156$, but was larger in the negative conditioning group, $BF_{10} = 2.02$. The CS × Block interaction revealed that conditional expectancy was present after the last block of acquisition, $BF_{10} = 1.73 \times 10^{88}$, and during extinction blocks 1, 2, 3, 4, 5, 6, and 7, all $F(1, 342) > 7.15$, $p <$

[14] This interaction only attains marginal significance in the sample of contingency verbalizers, $F(1, 278) = 3.13$, $p = .078$, $\eta p^2 = .011$, $BF_{inc} = 50.13$, but the BF_{inc} suggests that it is a reliable effect.

.008, $\eta p^2 > .020$, BF_{10}: blocks 1 = 6.96×10^{11}, 2 = 3.74×10^{6}, 3 = 7918.68, 4 = 6.77, 5 = 64.21, 6 = 13.92, 7 = 2.30. Conditional expectancy was not present during block 8, $F(1, 342) = 1.98$, $p = .161$, $\eta p^2 = .006$, $BF_{10} = 0.17$, and reduced from after acquisition to after extinction, $p < .001$, $BF_{10} = 2.21 \times 10^{83}$. The remaining omnibus effects did not reach significance, all $F < 1.07$, $p > .384$, $\eta p^2 < .025$, BF < 0.00.

Renewal. Renewal of expectancy conditioning was observed in both groups. Renewal of negative evaluative conditioning was observed in the frequentist analyses but the BF_{10} was indecisive. Renewal of positive evaluative conditioning was observed but only in the participants who passed the contingency assessment and manipulation check.

CS Valence. A main effect of CS, $F(1, 342) = 5.56$, $p = .019$, $\eta p^2 = .016$, $BF_{10} = 32.59$, and a CS × Group interaction, $F(1, 342) = 17.99$, $p < .001$, $\eta p^2 = .050$, $BF_{inc} = 1.04 \times 10^{7}$, were moderated by a CS × Group × Phase interaction, $F(1, 342) = 5.81$, $p = .016$, $\eta p^2 = .017$, $BF_{inc} = 0.32$. In the positive conditioning group, conditional valence was not present after the last block of extinction, $F(1, 342) = 0.69$, $p = .406$, $\eta p^2 = .002$, $BF_{10} = 0.13$, or during the renewal test, $F(1, 342) = 2.37$, $p = .124$, $\eta p^2 = .007$, $BF_{10} = 0.37$, and the magnitude of conditional valence did not change from after extinction to the renewal test, $F(1, 342) = 1.63$, $p = .202$, $\eta p^2 = .005$, $BF_{10} = 0.22$. Positive renewal was present, however, in the sample of participants who passed the contingency assessment and reported that they noticed the screen color change. In these participants, conditional valence was not present after the last block of extinction, $F(1, 212) = 0.72$, $p = .398$, $\eta p^2 = .003$, $BF_{10} = 0.15$, but was during the renewal test, $F(1, 212) = 4.29$, $p = .040$, $\eta p^2 = .020$, $BF_{10} = 1.14$. In this sample, the magnitude of the conditional valence increased from after extinction to the renewal test, $F(1, 212) = 4.60$, $p = .033$, $\eta p^2 = .021$, $BF_{10} = 3.59$. In the negative conditioning group, conditional valence was present after the last block of extinction, $F(1, 342) = 16.36$, $p < .001$, $\eta p^2 = .046$, $BF_{10} = 73.27$, and during the renewal test, $F(1, 342) = 22.26$, $p < .001$, $\eta p^2 = .061$, $BF_{10} = 524.75$. The magnitude of this conditional valence increased during renewal, $F(1, 342) = 4.60$, $p = .033$, $\eta p^2 = .013$, but the $BF_{10} = 0.60$ for this increase provided anecdotal evidence for the null hypothesis. When the participants who failed the contingency assessment and did not notice the screen color change were removed, the BF_{10} was strengthened, $BF_{10} = 1.095$. The remaining omnibus effects did not reach significance, all $F < 0.40$, $p > .530$, $\eta p^2 < .002$, $BF < 0.16$.

US Expectancy. A main effect of CS, $F(1, 342) = 50.61$, $p < .001$, $\eta p^2 = .129$, $BF_{10} = 4.03 \times 10^{15}$, was moderated by a CS × Phase interaction, $F(1, 342) = 54.75$, $p < .001$, $\eta p^2 = .138$, $BF_{inc} = 7.62 \times 10^{10}$, which confirmed that renewal occurred. Conditional expectancy was not present during the last block of extinction, $F(1, 342) = 1.98$, $p = .161$, $\eta p^2 = .006$, $BF_{10} = 0.17$, but increased from after extinction to the renewal test, $BF_{10} = 4.55 \times 10^9$, and was significant during the renewal test, $F(1, 342) = 68.21$, $p < .001$, $\eta p^2 = .166$, $BF_{10} = 1.70 \times 10^{12}$. The remaining omnibus effects did not reach significance, all $F < 1.94$, $p > .164$, $\eta p^2 < .006$, $BF < 0.20$.

US Valence. The mean US valence evaluations are displayed in Table 2. A main effect of valence, $F(3, 340) = 1221.84$, $p < .001$, $\eta p^2 = .915$, $BF_{10} = 2.79 \times 10^{526}$, was moderated by a Valence × Group interaction, $F(3, 340) = 8.83$, $p < .001$, $\eta p^2 = .072$, $BF_{inc} = 56.80$. Evaluations of pleasant, $BF_{10} = 0.20$, and unpleasant USs, $BF_{10} = 0.34$, did not differ between the conditioning groups, both $F(1, 342) < 2.21$, $p > .138$, $\eta p^2 < .007$, but the negative conditioning group evaluated both the neutral USs, $BF_{10} = 38.74$, and the novel neutral pictures, $BF_{10} = 11.04$, as more pleasant than the positive conditioning group, both $F(1, 342) > 9.53$, $p < .003$, $\eta p^2 > .027$. Both groups evaluated the unpleasant USs as less pleasant than the neutral USs and novel neutral pictures, and the pleasant USs as more pleasant than the neutral USs and the novel neutral pictures, all $p < .001$, all $BF_{10} > 3.25 \times 10^{62}$. Interestingly, however, evaluations of the neutral USs and the novel neutral pictures did not differ in the positive conditioning group, $p = .395$, $BF_{10} = 0.11$, but in the negative conditioning group, evaluations of neutral USs were more pleasant than of novel neutral pictures, $p = .015$, $BF_{10} = 8.14$. The main effect of group did not reach significance, $F(1, 342) = 3.28$, $p = .071$, $\eta p^2 = .009$, $BF_{10} = 0.10$.

Discussion

In Experiment 5 we examined whether conditional evaluative and expectancy conditioning would relapse after an ABA renewal manipulation in a between participants conditioning design. Participants completed acquisition training on color 1 and extinction training on color 2, before returning to the acquisition screen color for the relapse assessment. Conditional evaluative and expectancy learning developed during acquisition and reduced during extinction in both groups. We observed renewal of expectancy conditioning after the screen color change with conditional expectancy increasing from after extinction to the relapse assessment. We also observed renewal of evaluative conditioning but the effect was less robust. There was no evidence for renewal of positive evaluative conditioning in the full sample but renewal did occur in participants who passed the contingency assessment and manipulation check. In these participants, conditional valence increased from after extinction to the renewal test and the BF_{10} provided substantial evidence for the alternative hypothesis. In the negative conditioning group, conditional valence significantly increased from after extinction to the renewal test, but the BF_{10} was indecisive.

The results provide evidence that renewal of evaluative conditioning does occur but the effect was small and not robust. It is possible that these small effects occur because our renewal manipulation was not very strong. We avoided using salient colors during the renewal manipulation as unpublished pilot work suggested that more salient background colors can influence evaluative conditioning but as 25% of our sample reported that they did not notice the screen color change it is possible that the black/grey color change used here was too weak. Although, the screen color manipulation was sufficient to induce renewal of expectancy learning, it is possible that changing to more salient screen colors would produce stronger renewal and therefore blue and grey were used as the background screen colors in Experiment 6.

Experiment 6 – Within Participant Conditioning Design

Method

Participants. One hundred and fifty-eight participants from the United States (67[15] female) aged between 19 and 65 (*M* = 36.19, *SD* = 10.42) were recruited via Amazon's Mechanical Turk. Participants volunteered participation in exchange for $4.50 US. The research protocol was approved by the Curtin University Human Research Ethics Committee. Power was determined as in Experiment 4, however, we aimed to recruit additional participants to ensure an adequate sample of participants who noticed the screen color change.

Apparatus/Stimuli. An ABA renewal procedure was used with grey (rgb: 120, 120, 120) and blue (rbg: 65, 140, 210) background screen colors serving as the context manipulation. The remainder of the apparatus/stimuli were the same as for Experiment 2.

Procedure. The ABA renewal manipulation and renewal manipulation check were performed as in Experiment 5. The remainder of the procedure was the same as Experiment 2.

Data Preparation and Analyses. The data preparation and analyses were conducted as

[15] One participant did not provide sex information. Four participants did not provide age information.

in Experiment 2.

Results

Preliminary Analyses. CSp and CSu evaluations did not differ during the baseline assessment, $F(1, 157) = 0.01$, $p = .930$, $\eta p^2 < .001$, $BF_{10} = 0.09$. Seventeen participants were unable to verbalize the experimental contingencies and one participant did not answer the contingency question. Nineteen participants reported that they did not notice the screen color change during the experiment and one participant did not answer the screen color question. The pattern of relapse results does not change when participants who did not notice the screen color change are excluded.

Acquisition. The CS valence evaluations and US expectancy ratings measured during Experiment 6 are presented in Figures 13 and 14, respectively. Conditional evaluative and expectancy conditioning was successfully acquired during acquisition.

CS Valence. Main effects of CS, $F(1, 157) = 286.42$, $p < .001$, $\eta p^2 = .646$, $BF_{10} = 2.55 \times 10^{187}$, and block, $F(3, 155) = 11.04$, $p < .001$, $\eta p^2 = .176$, $BF_{10} = 0.02$, were moderated by a CS × Block interaction, $F(3, 155) = 41.16$, $p < .001$, $\eta p^2 = .443$, $BF_{inc} = 2.43 \times 10^{17}$. Conditional valence was present during all blocks, all $F(1, 157) > 84.40$ $p < .001$, $\eta p^2 > .349$, $BF_{10} > 2.52 \times 10^{13}$, but increased from block 1 to 4, $p < .001$, $BF_{10} = 4.01 \times 10^{18}$.

US Expectancy. A main effect of CS, $F(1, 157) = 712.39$, $p < .001$, $\eta p^2 = .819$, $BF_{10} = 3.46 \times 10^{315}$, was moderated by a CS × Block interaction, $F(3, 155) = 12.04$, $p < .001$, $\eta p^2 = .189$, $BF_{inc} = 1.28 \times 10^6$. Conditional expectancy was present during all acquisition blocks, all $F(1, 143) > 256.75$, $p < .001$, $\eta p^2 > .620$, $BF_{10} > 3.76 \times 10^{31,}$ and increased from block 1 to 4, $p < .001$, $BF_{10} = 1.04 \times 10^5$. The main effect of block did not reach significance, $F(3, 155) = 0.93$, $p = .426$, $\eta p^2 = .018$, $BF_{10} = 0.00$.

Extinction. The conditional valence evaluations and US expectancy ratings acquired during acquisition reduced throughout extinction.

CS Valence. A main effect of CS, $F(1, 157) = 100.17$, $p < .001$, $\eta p^2 = .390$, $BF_{10} = 1.08 \times 10^{150}$, was moderated by a CS × Block interaction, $F(8, 150) = 27.50$, $p < .001$, $\eta p^2 = .595$, $BF_{inc} = 6.16 \times 10^{72}$. Conditional valence was present during the last block of acquisition and during all extinction blocks, all $F(1, 157) > 30.02$, $p < .001$, $\eta p^2 > .160$, $BF_{10} > 6.57 \times 10^4$, but decreased from after acquisition to after extinction, $p < .001$, $BF_{10} = 7.89 \times 10^{25}$. The main effect of block did not reach significance, $F(8, 150) = 1.93$, $p = .060$, $\eta p^2 = .093$, $BF_{10} = 2.22 \times 10^{-5}$.

US Expectancy. A main effect of CS, $F(1, 157) = 118.38$, $p < .001$, $\eta p^2 = .430$, $BF_{10} = 4.20 \times 10^{137}$, was moderated by a CS × Block interaction, $F(8, 150) = 49.42$, $p < .001$, $\eta p^2 = .725$, $BF_{inc} = 1.09 \times 10^{139}$. Conditional expectancy was present during the last block of acquisition, $BF_{10} = 4.93 \times 10^{50}$, and during all extinction blocks, all $F(1, 157) > 10.36$, $p < .002$, $\eta p^2 > .061$, BF_{10}: blocks 1, 2, 3, 4, 5, 7 > 5342.94, 6 = 12.29, 8 = 1411.69, but decreased from the after acquisition to after extinction, $p < .001$, $BF_{10} = 4.96 \times 10^{34}$. The main effect of block did not reach significance, $F(8, 150) = 1.74$, $p = .093$, $\eta p^2 = .085$, $BF_{10} = 5.54 \times 10^{-6}$.

Renewal. Conditional evaluative and expectancy conditioning increased during renewal confirming that relapse occurred.

CS Valence. A main effect of CS, $F(1, 157) = 46.24$, $p < .001$, $\eta p^2 = .228$, $BF_{10} = 1.37 \times 10^{19}$, was moderated by a CS × Phase interaction[16], $F(1, 157) = 10.85$, $p = .001$, $\eta p^2 = .065$, $BF_{inc} = 2.11$. This interaction confirmed that renewal of evaluative conditioning occurred. Conditional valence was present after extinction, $F(1, 157) = 30.03$, $p < .001$, $\eta p^2 = .161$, $BF_{10} = 6.57 \times 10^4$,

[16] In the participants who noticed the screen color change and passed the contingency assessment the $BF_{inc} = 3.66$.

and during the renewal test, $F(1, 157) = 44.90$, $p < .001$, $\eta p^2 = .222$, $BF_{10} = 2.44 \times 10^7$, but conditional valence was larger during the renewal test, $BF_{10} = 15.41$. The main effect of phase did not reach significance, $F(1, 157) = 1.07$, $p = .302$, $\eta p^2 = .007$, $BF_{10} = 0.01$.

US Expectancy. A main effect of CS, $F(1, 157) = 37.95$, $p < .001$, $\eta p^2 = .195$, $BF_{10} = 2.32 \times 10^{16}$, was moderated by a CS × Phase interaction, $F(1, 157) = 11.87$, $p = .001$, $\eta p^2 = .070$, $BF_{inc} = 17.82$. This interaction confirmed that relapse of expectancy learning occurred. Conditional expectancy was present after extinction, $F(1, 157) = 20.96$, $p < .001$, $\eta p^2 = .118$, $BF_{10} = 1411.69$, and during the renewal test, $F(1, 157) = 34.71$, $p < .001$, $\eta p^2 = .181$, $BF_{10} = 4.46 \times 10^5$, but the magnitude of this difference was larger during renewal, the $BF_{10} = 24.52$. The main effect of phase did not reach significance, $F(1, 157) = 0.01$, $p = .945$, $\eta p^2 < .001$, $BF_{10} = 0.09$.

US Valence. Pleasant USs ($M = 8.22$, $SD = 0.83$) were evaluated as more pleasant than unpleasant USs ($M = 1.83$, $SD = 1.32$), $F(1, 157) = 1798.93$, $p < .001$, $\eta p^2 = .920$, $BF_{10} = 1.32 \times 10^{84}$.

Discussion

In Experiment 6 we examined whether evaluative and expectancy conditioning would relapse after an ABA renewal manipulation in a within participants conditioning design. To increase the salience of the renewal manipulation we changed the background screen colors to grey and blue which resulted in a reduction in the number of participants who failed the manipulation check (12.6% in Experiment 6 vs. 24.4% in Experiment 5). Conditional valence and expectancy learning developed throughout acquisition and reduced during extinction. Robust renewal of both evaluative and expectancy conditioning was observed with conditional valence and expectancy ratings increasing after the screen color change. Strong support for the renewal was obtained for both evaluative and expectancy relapse in the Bayesian analyses. These data provide strong evidence that renewal does occur in evaluative learning and suggest that the weaker results obtained in Experiment 5 were likely due to the weaker relapse manipulation used.

General Discussion

Experimental Aims and Summary

Across three streams of research, we examined whether evaluative conditioning would relapse after spontaneous recovery (Stream 1), reinstatement (Stream 2), and renewal (Stream 3). All relapse manipulations occurred after extinction training and were followed by a test of CS valence and US expectancy (relapse test). Spontaneous recovery was tested by conducting the relapse test after a 15 minute delay (Experiment 1) or on the following day (Experiment 2). Reinstatement was induced by presenting the valenced US pictures unsignalled in a randomized stream. Renewal was induced by conducting the relapse test on the acquisition screen color (color A) after extinction training on a different screen color (color B). We examined each relapse induction in a between-participant and a within-participant conditioning design to ensure the findings were robust across procedural variations and to allow us to replicate the findings within the current manuscript. Including a measure of US expectancy allowed us to confirm that the relapse manipulation was having the expected influence and to examine evaluative conditioning and expectancy conditioning within the same experimental paradigm. This is preferable to comparing evaluative conditioning and expectancy conditioning across different paradigms as procedural variations between studies could influence the results and create artificial dissociations. In all experiments, we expected that conditional CS valence evaluations and US expectancy ratings would be acquired throughout acquisition. We predicted that these conditional evaluations would reduce throughout extinction, but based on Hoffman et al. (2010) we expected that some differential valence evaluations could remain at the end of extinction. After the relapse inductions, we hypothesized that differential conditional valence evaluations and US expectancy ratings would increase.

Relapse of Evaluative and Expectancy Conditioning

In all experiments we observed a reduction of differential CS evaluations and US expectancy ratings during extinction, suggesting that an inhibitory association was formed during extinction training. We also observed spontaneous recovery, reinstatement, and renewal of expectancy learning in all experiments – such that conditional US expectancy ratings (i.e. the difference between the CSv and CSn or the CSp and CSu) increased from after extinction to the relapse test. These findings replicate results from the fear conditioning literature and suggest that the experimental manipulations were having the intended effect in all cases.

Overall, we found strong evidence that evaluative conditioning does relapse after extinction, but unlike for US expectancy, the results differed across relapse inductions. In Stream 1, evaluative conditioning did not relapse after the spontaneous recovery manipulation. This finding was consistent across two experiments using different conditioning designs and is very unlikely to occur because of a problem with the manipulation. In Experiment 1, the spontaneous recovery assessment took place 15 minutes after extinction, and while this delay is relatively short, it was sufficient to induce relapse of expectancy conditioning. In Experiment 2, we replicated the results with a spontaneous recovery assessment that took place one day after extinction. Although the sample size in Stream 1 was smaller than in the other two streams, we think it is very unlikely that inadequate experimental power is responsible for the results. The pattern observed in both spontaneous recovery assessments was different to that observed in the other research streams. In both conditioning designs, all CSs became more pleasant after the time delay – including those that had been paired with negative USs during acquisition. This overall increase in pleasantness for all CSs is a *qualitatively* different pattern to that predicted by our hypothesis. Furthermore, the Bayesian analyses provided substantial evidence for the null hypothesis in both conditioning groups of Experiment 1 and in Experiment 2.

In Stream 2, evaluative conditioning consistently relapsed after the reinstatement manipulation. In Experiment 3a, relapse of evaluative conditioning was present after negative conditioning, but not after positive conditioning. In Experiment 3b, there was evidence for relapse of evaluative conditioning after both negative and positive conditioning – but while relapse of negative learning was strong, relapse of positive learning was less robust – with marginal relapse seen only in the participants who could verbalize the experimental contingencies. The extent of evaluative learning during the acquisition phase of Experiments 3b was larger for the negative conditioning group than for the positive conditioning group and therefore it seems likely that weaker relapse observed after positive reinstatement is the result of weaker learning in the positive conditioning group. Replicating the results obtained in Experiments 3a and 3b, reinstatement of evaluative conditioning was observed in Experiment 4 using a within-participants conditioning design.

In Stream 3, evaluative learning relapsed after renewal in both the between participant and the within participant conditioning design but the results were more robust in the within participant design. In Experiment 5, renewal of negative evaluative conditioning was observed in the full sample but this was very small and only detected using the frequentist analyses. Renewal of positive evaluative learning was not observed in the full sample but was present in participants who noticed the screen color change and could verbalize the experimental contingencies and was supported in both the frequentist and Bayesian analyses. In Experiment 6, using the within-participants design, a larger renewal effect was detected – likely because the renewal manipulation was strengthened by making the change between background screen colors more salient.

Taken together, the results suggest that evaluative conditioning is subject to reinstatement and renewal, but not spontaneous recovery. Observing relapse of evaluative conditioning suggests that evaluative learning is mediated by a signal based learning mechanism but the dissociations across relapse inductions suggest that there may be subtle differences in the underlying mechanisms. The current findings are in line with Bouton’s (2002) theory of relapse which proposes that extinction does not erase the original CS-US association learned during acquisition, but instead creates a new context specific extinction association (CS-noUS) which inhibits the original association. Bouton proposed that reinstatement and spontaneous recovery are special cases of renewal. In reinstatement, the presentation of the US acts as an acquisition context reminder and in spontaneous recovery the delayed test occurs in a new *temporal* context. The dissociation observed across relapse inductions in the current research is also consistent with this model. Both reinstatement and renewal involve an active reminder of acquisition – either via the presentation of the US or the return to the acquisition context. Conversely, spontaneous recovery does not involve the presentation of any stimuli or a return to the acquisition context (i.e. there is no explicit manipulation). As it is not possible to return to the acquisition temporal context, any change in the passage of time would result in the formation of a new temporal context (i.e. temporal context C). Therefore, while, reinstatement and renewal trigger the actual *acquisition* context, spontaneous recovery would trigger relapse because of the departure from the extinction context to a *new* context C. This account explains differences between reinstatement, renewal, and spontaneous recovery for evaluative learning, but does not explain why these discrepancies were not evident in expectancy learning.

The observation of differences across relapse inductions for evaluative learning but not for expectancy learning is also consistent with the temporal integration hypothesis that was first proposed by Lipp et al. (2010) and further developed and tested by Aust, Stahl, and Haaf (2019). The temporal integration hypothesis suggests that expectancy learning and evaluative learning are governed by a single learning system but differ in the manner in which recent and past events are weighted in order to affect behaviour. Whereas expectancy conditioning is weighted heavily towards the more recent past to permit quick adaptation to changes and prediction and preparation for future events, evaluative conditioning is weighted more heavily towards experiences in the more distant past. For example, when making decisions about evaluations and expectancies a different proportion of the learning history may be taken into account – i.e. participants may consider the last 100 trials when making evaluative decisions but only the last 10 trials when making predictions. This differential weighting of past events may explain the differential influence of relapse inductions on evaluative and expectancy learning. Both reinstatement and renewal provide participants with a reminder of ‘past experience’ to integrate into upcoming evaluative judgements (i.e. the stream of USs and return to the acquisition context). The mere passage of time used to induce spontaneous recovery, however, does not emphasise participants’ ‘past experience’ which took place prior to extinction. The change in temporal context may be sufficient to trigger relapse of expectancy learning as it is heavily weighted towards predicting the future, but this change may not be sufficient to trigger relapse of evaluative learning as it does not relate to experiences that occurred in the more distant past and the training received during extinction will exert a stronger effect.

Positive vs. Negative Conditioning

Across Experiments, differences between positive and negative conditioning emerged. Both evaluative and expectancy conditioning effects were consistently stronger in the negative conditioning group and there was some evidence that negative conditioning took longer to extinguish than positive conditioning. The conditioning groups also evaluated the CSn differently – despite the CSn having the same training history between groups. This shift was present throughout acquisition and was often also present during early extinction. The neutral US images were evaluated as less pleasant after positive conditioning than after negative conditioning in all experiments except 3b, which suggests that the difference in CSn valence was driven by a relative shift in the valence of the neutral US pictures. Neutral pictures are likely perceived as relatively more unpleasant when they are presented with pleasant pictures during positive conditioning and relatively more pleasant when they are presented with unpleasant pictures during negative conditioning. While, it is possible that the valence of the neutral US images has intrinsically changed, the valence shift could occur because the manner in which participants report the US valence has changed – for instance, the mechanism may be a temporary framing effect or a more general response bias.

In the case of a temporary framing effect, the presentation of the valenced images during acquisition may temporarily anchor participants' ratings for the remainder of the experimental session. For example, if participants are initially shown highly graphic negative images (i.e. pictures of mutilated bodies) they may evaluate more mildly negative images (i.e. pictures of rubbish) as more pleasant than they would if they were shown the same mildly negative images alongside pleasant pictures – but, conceivably, in both cases a given picture of rubbish has a 'true' valence that has not changed. The same mechanism may be responsible for the differential evaluation of neutral USs in the current studies – the 'true' valence of the neutral USs has not intrinsically changed but the whole category ‘neutral’ is now anchored in the evaluative judgements relative to what else was presented during acquisition (i.e. pleasant or unpleasant). On the other hand, if the shift is caused by a general response bias, the valence shift may simply reflect a tendency for participants to use the end points of the 1-9 valence scale during acquisition (i.e. during positive conditioning: CSv = 9 and CSn = 1 and during negative conditioning: CSv = 1 and CSn = 9). This should also result in an exaggeration of pleasant US valence in the positive conditioning group and unpleasant US valence in the negative conditioning group but this may not be detected in a small sample due to ceiling/floor effects.

In Experiments 3b and 5 we added novel neutral pictures during the final valence test in an attempt to disentangle these mechanisms. If the valence of the neutral images has intrinsically changed the valence shift should only be present for the neutral USs and not the novel neutral pictures. If, however, the valence shift reflects a framing effect then the valence of both neutral USs and novel neutral pictures should change between groups, with no differences between the neutral categories. If the valence shift reflects a general response bias, it should be present for both neutral USs and novel neutral pictures and should also exaggerate evaluations of the US category used during conditioning.

To maximize power and detect the most reliable differences, we ran a combined analysis on the US valence results obtained in Experiments 3b and 5. A Valence × Group interaction, $F(3, 611) = 9.86$, $p < .001$, $\eta p^2 = .046$, $BF_{inc} = 17.14$, revealed that neutral USs, $F(1, 613) = 7.12$, $p = .008$, $\eta p^2 = .011$, $BF_{10} = 2.86$, and novel neutral pictures, $F(1, 613) = 3.99$, $p = .046$, $\eta p^2 = .006$, $BF_{10} = 0.63$, were evaluated as more pleasant in the negative conditioning group than in the positive conditioning group. Furthermore, while evaluations of positive USs did not differ between the groups, $F(1, 613) = 0.20$, $p = .658$, $\eta p^2 < .001$, $BF_{10} = 0.01$, evaluations of unpleasant USs, $F(1, 613) = 3.85$, p =.050, $\eta p^2 = .006$, $BF_{10} = 0.59$, were marginally more unpleasant in the negative conditioning group than in the positive conditioning group. Evaluations of neutral pictures and novel neutral pictures did not differ in the positive conditioning group, $p = .516$, $BF_{10} = 0.08$, but neutral USs were evaluated as more pleasant than novel pictures in the negative conditioning group, $p = .016$, $BF_{10} = 2.52$.

Based on these results, a response bias explanation seems unlikely. Positive US valence was not exaggerated in the positive conditioning group and the negative conditioning group evaluated the unpleasant USs as *less* pleasant than the positive conditioning group. Consistent with a general framing effect, the negative conditioning group evaluated the valence of the novel images as less pleasant than the positive conditioning group – although the BF for this effect was indecisive. The observation that the negative conditioning group evaluated the neutral USs as more pleasant than the novel neutral pictures, however, suggests that more than a general framing effect is at play. More research will be required to identify the exact mechanisms underlying this effect and to explore why differences between the neutral USs and novel neutral pictures are present only in the negative conditioning group.

Extinction of Evaluative Learning

Across studies, differential valence was still evident at the end of extinction but we did observe a substantial reduction in the size of conditional valence from after acquisition to after extinction. When the data are combined across all studies, approximately 80% of the conditional response was eliminated during the extinction procedure (78% for the within participant designs: conditional valence after acquisition = 5.09 vs. conditional valence after extinction = 1.10; 87% for positive conditioning: 1.79 vs. 0.24; and 84% for negative conditioning: 2.77 vs. 0.43). On average, conditional expectancy learning was reduced by 95% from after acquisition to after extinction (89% for within designs: 4.91 vs. 0.55; 99% for positive conditioning: 3.82 vs. 0.05; and 95% for negative conditioning: 4.75 vs. 0.26). Although, more of the conditional response seemed to remain for evaluative conditioning, the extinction effects were substantial.

The large extinction effects may occur because the inclusion of an expectancy measure biases participants to base their evaluative judgements on their recent expectancy ratings. Similar extinction effects, however, have been observed in the same within-participants design used here when participants do not rate US expectancy during acquisition and extinction (see Luck and Lipp, in revision). It is also possible that rating CS valence multiple times during conditioning exaggerates extinction. Gawronski, Gast, and De Houwer (2014) argue that reductions in evaluative learning after extinction are due to changes in cognitive processes due to being asked to rate valence multiple times and not due to changes in the actual evaluative *representations.* In a within-participants design, the authors found a reduction in conditional valence from after acquisition to after extinction if participants rated CS valence both after acquisition and after extinction. In a between-participants design, however, when one group of participants completed CS evaluations and affective priming after acquisition and another group completed the same measures after extinction, the authors did not observe a difference between groups in either measure. Gawronski et al. (2014) argue that the within-participant reduction from post-acquisition to post-extinction occurs because of changes in judgements and not evaluative representations but this interpretation assumes that the results of the between-participant comparisons reflect the true evaluative representations. Instead, failing to find a difference in the between-participant study could occur because participants in the acquisition and extinction group use the scale similarly to indicate a relative preference for one stimulus over another the first time they are asked. For instance, to communicate a relative preference for the CSp over the

CSu on a 1-9 pleasantness scale participants may provide judgements of ‘3’ and ‘7’, respectively – but a judgement of ‘3’ and ‘7’ may not correspond to identical evaluative representations in the different groups. Judgments on a pleasantness scale are very likely influenced by a number of aspects of the experiment including the other stimuli the participants view and how many times the participant has been asked to rate the stimuli. A single judgment (i.e. ‘3’ or ‘5’ or ‘7’) is very unlikely to correspond to an exact amount of valence and therefore it is important to be cautious in interpreting valence judgments across groups as a concrete change in evaluative representation (or lack of) when other aspects of the procedure differ (i.e. the groups have been asked to evaluate stimuli a different number of times). This issue is difficult to resolve as implicit measures of CS valence should not be biased by these default response tendencies but suffer from low reliability and may not be sensitive enough to detect small evaluative changes after extinction (see Gawronski & De Houwer, 2014). It is possible that the extinction effects detected in our investigation occur because asking participants to rate valence multiple times leads to changes in evaluative judgements and not evaluative representations but more research will be required to settle this question in the future. In the absence of further evidence, however, it seems most parsimonious, to us at least, to conclude that participants used their evaluative representations to make their evaluative judgments.

Implications for Theories of Evaluative Learning

The current findings pose a challenge for both single-process and dual process theories of evaluative and expectancy learning. The clear extinction effects in this investigation as well as our observations of reinstatement and renewal suggest that evaluative learning is mediated by a signal-based learning system and are in line with single-process models proposing that evaluative and expectancy learning have the same underlying mechanisms. The differential effect of spontaneous recovery on evaluative and expectancy learning, however, is consistent with theories proposing distinct learning systems. This dissociation, however, can be accommodated within single-process models which make additional assumptions regarding transfer or judgement strategies, such as the temporal integration hypothesis (see Lipp et al., 2010 and Aust, Stahl, & Haaf, 2019). More research will be required to investigate the mechanisms underlying evaluative learning and the current investigation was not intended to settle debates of this nature. Instead, our findings demonstrate for the first time that relapse of evaluative learning does occur. Future investigations of relapse in evaluative conditioning will be useful to examine dissociations or similarities between evaluative and expectancy learning.

Limitations and Future Directions

It is not possible within the current set of experiments to exclude that our results are due to demand characteristics or to exclude the possibility that the evaluative relapse we observed was due to judgment related processes rather than changes in evaluative representations – if these are indeed different. A demand characteristics explanation, however, is not consistent with the dissociation between evaluative and expectancy learning in Experiments 1 and 2. In other work, we have demonstrated that the general conditioning procedure used here leads to conditional evaluative learning using an implicit measure (see Luck and Lipp, in revision) but demonstrating evaluative relapse using an implicit measure will be an important direction for future work. This will not be without challenge, however, as the effects of relapse are small and may only be present during the first few trials of the relapse test. Ideally, the implicit measure would be administered after extinction and then again after the relapse induction in a within-participant design or in a between-participants design comparing the evaluative conditioning after extinction and after the relapse induction in separate groups. As relapse is transient, however, it is likely to extinguish within the first few trials of the implicit test. Affective priming, for instance, involves the presentation of the CSs repeatedly as primes. The repeated presentation of the CSs could be considered akin to extinction training or a retraining of the CSs in a 50-50 evaluative training schedule as the CS primes are paired with pleasant and unpleasant target words. Blink startle, a brain-stem reflex that is sensitive to stimulus valence and not under conscious control, may provide an alternative implicit measure to detect evaluative relapse (Lipp, 2006). Emotional modulation of startle responding, however, is subject to boundary conditions, including stimulus arousal and may require the use of high arousal USs, such as pleasant and unpleasant sounds. Measuring startle during evaluative conditioning will also require considerable procedural modifications, including the use of longer CS durations (6-8 s) and the addition of aversive startle probes (50 ms 105 dB blasts of white noise) to the CSs. This is a very important direction for future work, but it is not clear whether these modifications will influence evaluative learning or the relapse process itself. Future research should also examine whether the process of evaluative relapse differs when a wider range of CSs and USs is used during conditioning and explore whether the congruency, or match, between the CSs and USs influences relapse. For example, conditioning with stimuli that have a logical match (i.e. neutral faces and positive/negative character words) may influence the acquisition of evaluative learning and possibly modulate evaluative relapse.

Implications for Fear Conditioning Research

The current findings have relevance for fear conditioning and anxiety relapse prevention research. Negative valence is acquired to a CS during fear acquisition – or to the feared object in the case of clinical phobias – and this negative valence has been shown to predict fear relapse (see Hermans et al., 2004; Zbozinek et al., 2015). The current research suggests that, if this negative valence is reduced during treatment, this reduction will also be subject to relapse. Future research will be required to investigate evaluative relapse within fear conditioning to examine whether it also occurs within this paradigm. If so, then it will be interesting to investigate whether interventions that have been shown to reduce fear relapse would have similar levels of success on evaluative relapse.

Conclusions and Future Directions for Evaluative Relapse

We have demonstrated for the first time that relapse of evaluative conditioning occurs reliably and have established that evaluative relapse is responsive to reinstatement and renewal, but not spontaneous recovery. This proof of concept study has laid the ground for an exciting new direction of research in evaluative conditioning. Future work will be required to replicate evaluative relapse using an implicit measure, to determine why evaluative learning is not subject to spontaneous recovery, to investigate differences between positive and negative relapse, to investigate how different conditioning procedures influence relapse of evaluative learning, and to examine whether manipulations which have been shown to reduce relapse in fear conditioning will also reduce relapse of evaluative learning. Relapse of evaluative learning provides a new tool to examine the mechanisms underlying evaluative conditioning and to investigate dissociations between evaluative and expectancy learning. We hope this proof of concept stimulates an exciting and fruitful new direction of research aimed at understanding evaluative conditioning.

References

Aust, F., Haaf, J. M., & Stahl, C. (2019). A Memory-Based Judgment Account of Expectancy-Liking Dissociations in Evaluative Conditioning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/xlm0000600

Baeyens, F., Crombez, G., Van den Bergh, O., & Eelen, P. (1988). Once in contact always in contact: Evaluative conditioning is resistant to extinction. *Advances in Behaviour Research and Therapy, 10*(4), 179-199. doi: 10.1016/0146-6402(88)90014-8

Baeyens, F., Díaz, E., & Ruiz, G. (2005). Resistance to extinction of human evaluative conditioning using a between-subjects design. *Cognition and Emotion*, *19*(2), 245-268. doi:10.1080/02699930441000300

Baeyens, F., Eelen, P., Crombez, G., & van den Bergh, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behaviour Research and Therapy, 30*(2), 133-142. doi:http://dx.doi.org/10.1016/0005-7967(92)90136-5

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

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science, 6*, 3-5. doi:10.1177/1745691610393980

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*(6), 853-869. doi:10.1037/0033-2909.127.6.853

Dibbets, P., Havermans, R., & Arntz, A. (2008). All we need is a cue to remember: The effect of an extinction cue on renewal. *Behaviour Research and Therapy, 46*(9), 1070-1077. doi:https://doi.org/10.1016/j.brat.2008.05.007

Dibbets, P., Havermans, R. C., & Arntz, A. R. (2012). Partial presentation of an extinction cue enhances renewal in an ABA paradigm. *The Open Psychology Journal, 5*, 11-19.

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2004). Reinstatement of extinguished conditioned responses and negative stimulus valence as a pathway to return of fear in humans. *Learning & Memory, 11*(5), 549-554.

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

Engelhard, I. M., Leer, A., Lange, E., & Olatunji, B. O. (2014). Shaking That Icky Feeling: Effects of Extinction and Counterconditioning on Disgust-Related Evaluative Learning. *Behavior Therapy, 45*(5), 708-719. doi:10.1016/j.beth.2014.04.003

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

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

Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus Contextualization in Automatic Evaluation. *Journal of Experimental Psychology: General, 139*, 683-701. doi:10.1037/a0020315

Hermans, D., Crombez, G., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2002). Expectancy-learning and evaluative learning in human classical conditioning: Differential effects of extinction. In S. P. Shohov (Ed.), *Advances in psychology research*, Vol. 12, pp. 17-40). Hauppauge, NY, US: Nova Science Publishers.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*(3), 390-421. doi:10.1037/a0018916

Jeffreys, H. (1961). *Theory of Probability* (3 ed.). Oxford, UK: Oxford University Press.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical Report A-8.* University of Florida, Gainesville, FL.

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., Mallan, K. M., Libera, M., & Tan, M. (2010). The effects of verbal instruction on affective and expectancy learning. *Behaviour Research and Therapy, 48*(3), 203-209. doi:http://dx.doi.org/10.1016/j.brat.2009.11.002

Lipp, O. V., & Purkis, H. M. (2005). No support for dual process accounts of human affective learning in simple Pavlovian conditioning. *Cognition and Emotion, 19*(2), 269-282. doi:10.1080/02699930441000319

Luck, C. C., & Lipp, O. V. (2015). A potential pathway to the relapse of fear? Conditioned negative stimulus evaluation (but not physiological responding) resists instructed extinction. *Behaviour Research and Therapy, 66*(0), 18-31. doi:10.1016/j.brat.2015.01.001

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

Luck, C., & Lipp, O. V. (2019, September 13). Relapse of Evaluative Learning – Evidence for Reinstatement, Renewal, but not Spontaneous Recovery, of Extinguished Evaluative Learning in a Picture-Picture Evaluative Conditioning Paradigm. Retrieved from osf.io/hvak4

Luck, C. C., & Lipp, O. V. (in revision). Measuring unconditional stimulus expectancy during evaluative conditioning strengthens explicit, but not implicit, conditional stimulus valence.

Mathôt, S. (2017) Bayes like a Baws: Interpreting Bayesian repeated measures in JASP. Retrieved from https://www.cogsci.nl/blog/interpreting-bayesian-repeated-measures-in-jasp

O'Brien, F., & Cousineau, D. (2014). Representing Error bars in within-subject designs in typical software packages. Tutorials in Quantitative Methods for Psychology, 10, 56-67. doi:10.20982/tqmp.10.1.p056

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An Inkblot for Attitudes: Affect Misattribution as Implicit Measurement. *Journal of Personality and Social Psychology, 89*, 277-293. doi:10.1037/0022-3514.89.3.277

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

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*(4), 654-667.

Table 1

Means, Standard Deviations, and Test Statistics for variables assessed in the preliminary analyses of Experiments 1, 3a, 3b, and 5.

	Positive Conditioning	Negative Conditioning	Comparison
Experiment 1 Sex Ratio (Male:Female:Other)	2:22:0	4:20:0	$\chi^2(1) = 0.76$, $p = .383$
Contingency Ratio (Fail:Pass)	2:22	1:23	$\chi^2(1) = 0.36$, $p = .551$
Age	22.50 (10.18)	22.58 (5.12)	$t(46) = 0.36$, $p = .972$, $BF_{10} = 0.29$
Baseline Valence CSv CSn	 6.42 (1.68) (5.89 (1.85)	 6.55 (2.22) 6.32 (2.23)	$F < 3.99$, $p > .052$, $\eta p^2 < .093$, $BF < 1.11$.
Experiment 3a			
Sex Ratio (Male:Female:Other)	10:14:0	6:18:0	$\chi^2(1) = 1.50$, $p = .221$
Contingency Ratio (Fail:Pass)	3:21	2:22	$\chi^2(1) = 0.22$, $p = .637$
Age	23.29 (10.56)	24.33 (9.47)	$t(46) = 0.36$, $p = .721$, $BF_{10} = 0.30$
Baseline Valence CSv CSn	 6.45 (2.14) 6.70 (2.06)	 6.72 (2.65) 6.94 (1.83)	all $F < 0.70$, $p > .408$, $\eta p^2 < .020$, $BF < 0.48$.
Experiment 3b			
Sex Ratio (Male:Female:Other)	79:51:0	78:57:0	$\chi^2(1) = 0.25$, $p = .620$
Contingency Ratio (Fail:Pass)	39:92	27:110	$\chi^2(1) = 3.65$, $p = .056$

Age	35.55 (9.69)	37.55 (11.82)	$t(255.76) = 1.50$, $p = .135$, $BF_{10} = 0.39$.
Baseline Valence CSv CSn	 5.82 (1.67) 5.84 (1.60)	 5.84 (1.50) 5.86 (1.65)	all $F < 0.45$, $p > .832$, $\eta p^2 < .001$, BF < 0.19.
Experiment 5			
Sex Ratio (Male:Female:Other)	99:68:0	91:81:1	$\chi^2(1) = 2.37$, $p = .306$
Contingency Ratio (Fail:Pass)	31:136	30:144	$\chi^2(1) = 0.10$, $p = .750$
Noticed Screen Change Ratio (Pass:Fail)	130:37	127:47	$\chi^2(1) = 1.08$, $p = .298$
Age	35.75 (10.26)	36.18 (10.45)	$t(332) = 0.38$, $p = .703$, $BF_{10} = 0.13$.
Baseline Valence CSv CSn	 5.74 (1.50) 5.85 (1.52)	 5.89 (1.42) 5.93 (1.33)	all $F < 1.47$, $p > .226$, $\eta p^2 < .005$, BF < 0.25.

Table 2

Means and Standard Deviations for the Unconditional Stimuli in Experiments 1, 3a, 3b, and 5.

	Positive Pictures	Neutral Pictures	Novel Pictures	Negative Pictures
Experiment 1				
Positive Conditioning	8.12 (0.78)	4.02 (1.20)	N/A	1.84 (0.62)
Negative Conditioning	8.32 (0.75)	5.35 (0.96)	N/A	1.93 (0.82)
Experiment 3a				
Positive Conditioning	7.96 (0.95)	3.92 (1.25)	N/A	1.64 (0.54)
Negative Conditioning	8.38 (0.63)	5.52 (1.27)	N/A	1.83 (0.78)
Experiment 3b				
Positive Conditioning	8.12 (0.94)	5.29 (1.00)	5.29 (0.95)	2.32 (1.97)
Negative Conditioning	8.07 (0.91)	5.29 (0.75)	5.26 (0.88)	2.04 (1.61)
Experiment 5				
Positive Conditioning	8.00 (0.87)	5.02 (0.88)	4.98 (0.84)	2.06 (1.35)
Negative Conditioning	8.10 (0.85)	5.37 (0.95)	5.26 (0.85)	1.84 (1.38)

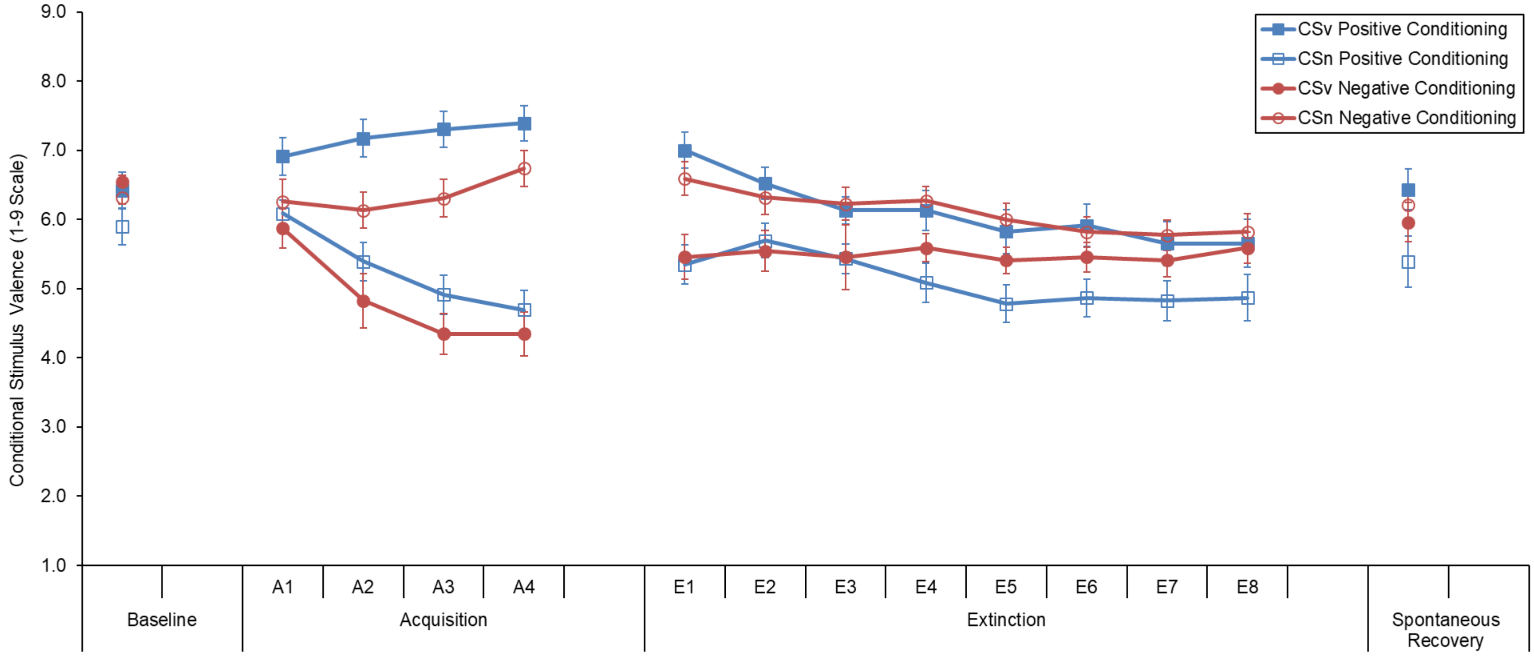

Figure 1. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the spontaneous recovery test of Experiment 1. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

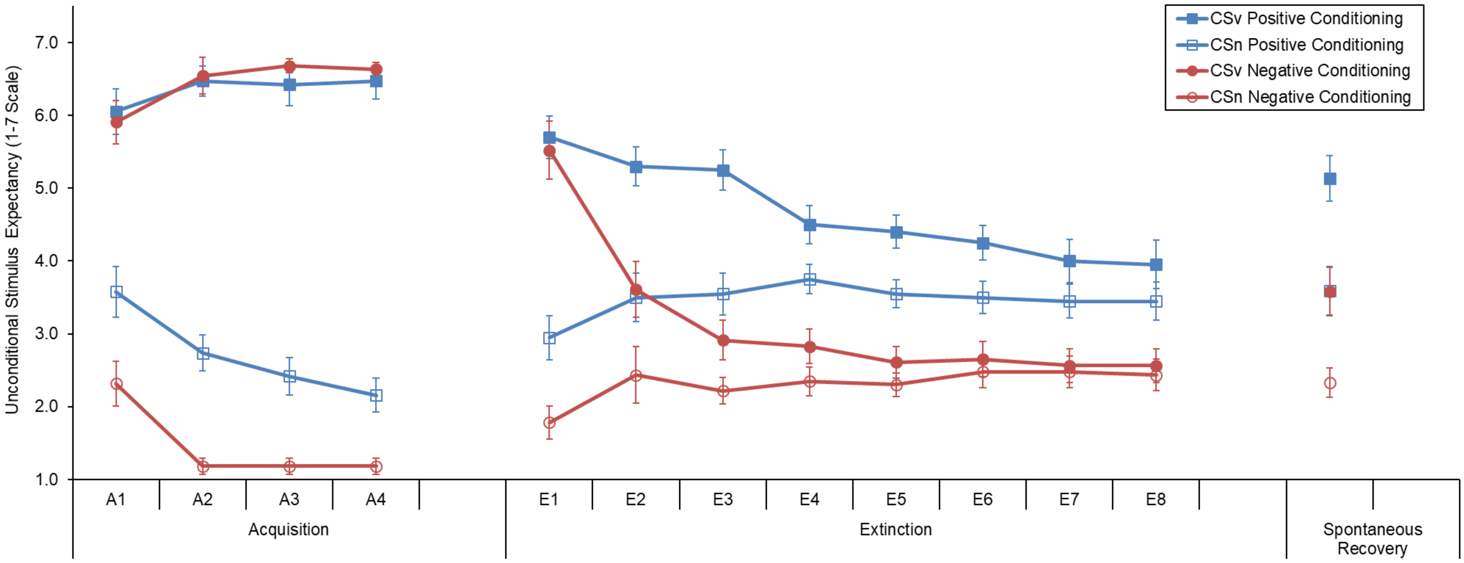

Figure 2. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the spontaneous recovery test of Experiment 1. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Ratings lower than 4 reflect a prediction of neutral images, ratings higher than 4 reflect a prediction of pleasant/unpleasant images, and a rating of 4 reflects a prediction neither pleasant nor unpleasant images. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).

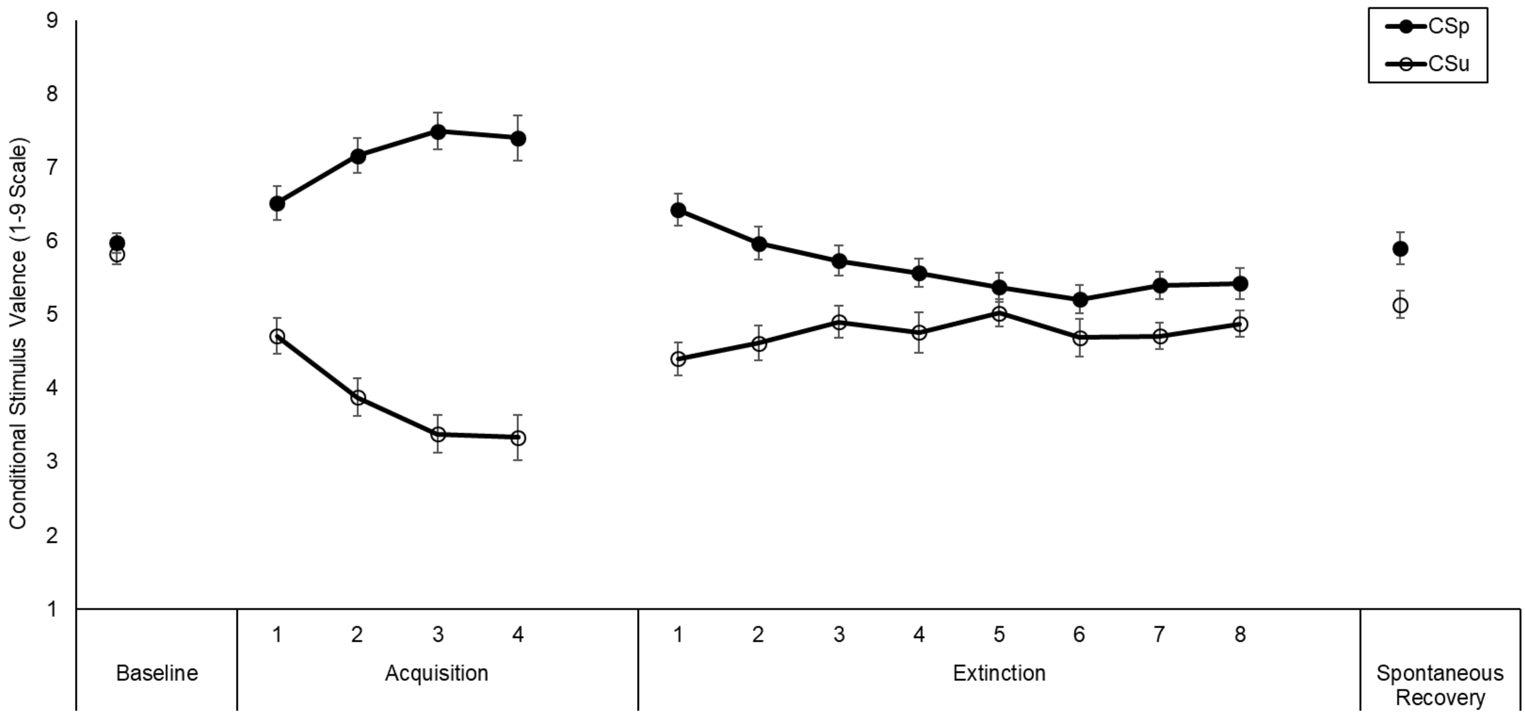

Figure 3. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the spontaneous recovery test of Experiment 2. During acquisition, the CSp was paired with pleasant images and the CSu was paired with unpleasant images. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).

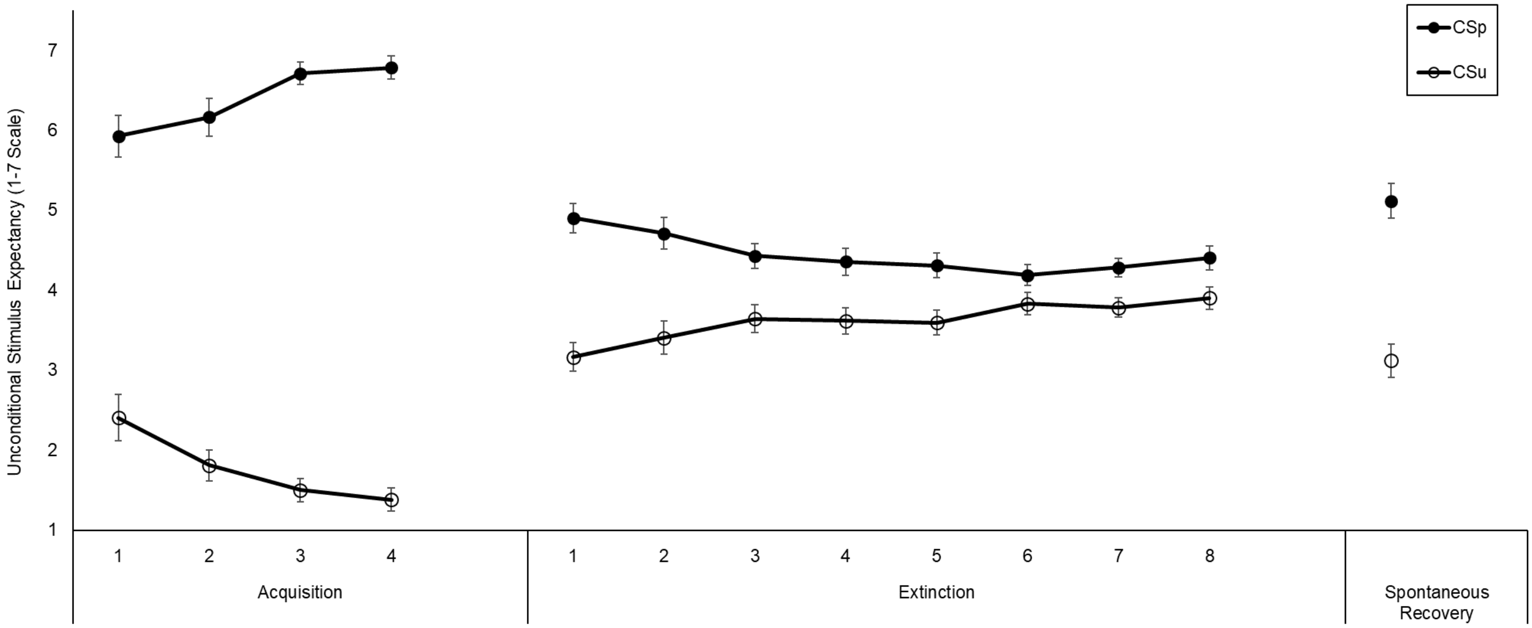

Figure 4. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the spontaneous recovery test of Experiment 2. During acquisition, the CSp was paired with pleasant images and the CSu was paired with unpleasant images. Ratings lower than 4 reflect a prediction of unpleasant images, ratings higher than 4 reflect a prediction of pleasant images, and a rating of 4 reflects a prediction of no images. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

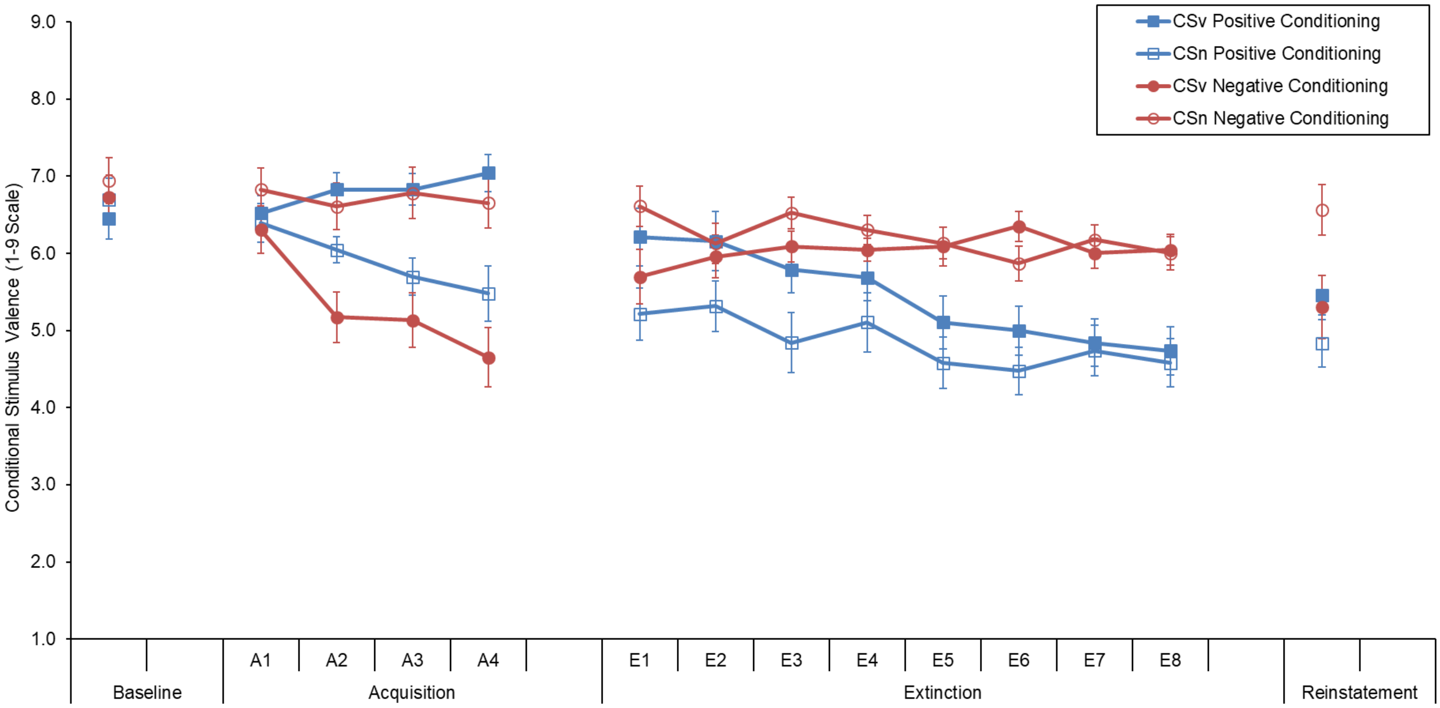

Figure 5. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the reinstatement test of Experiment 3a. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).

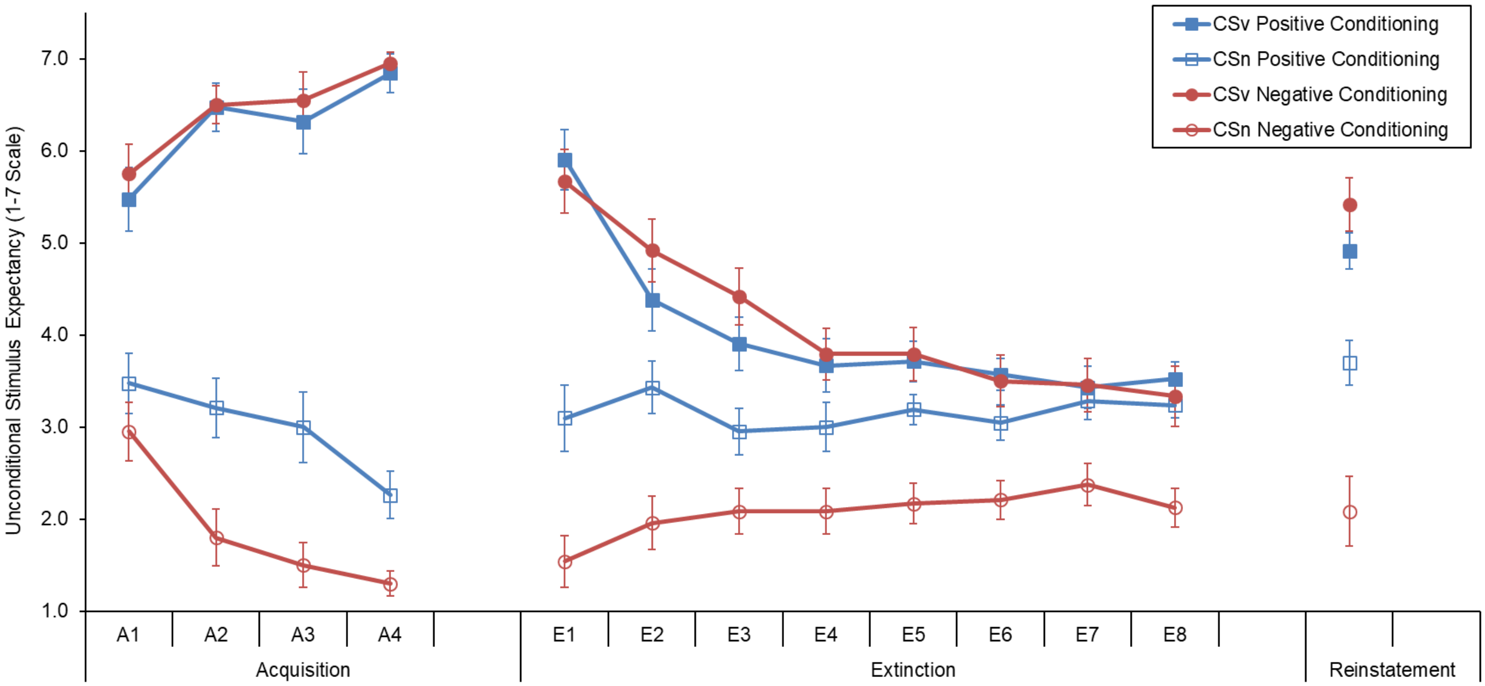

Figure 6. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the reinstatement test of Experiment 3a. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Ratings lower than 4 reflect a prediction of neutral images, ratings higher than 4 reflect a prediction of pleasant/unpleasant images, and a rating of 4 reflects a prediction of neither pleasant nor unpleasant images. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

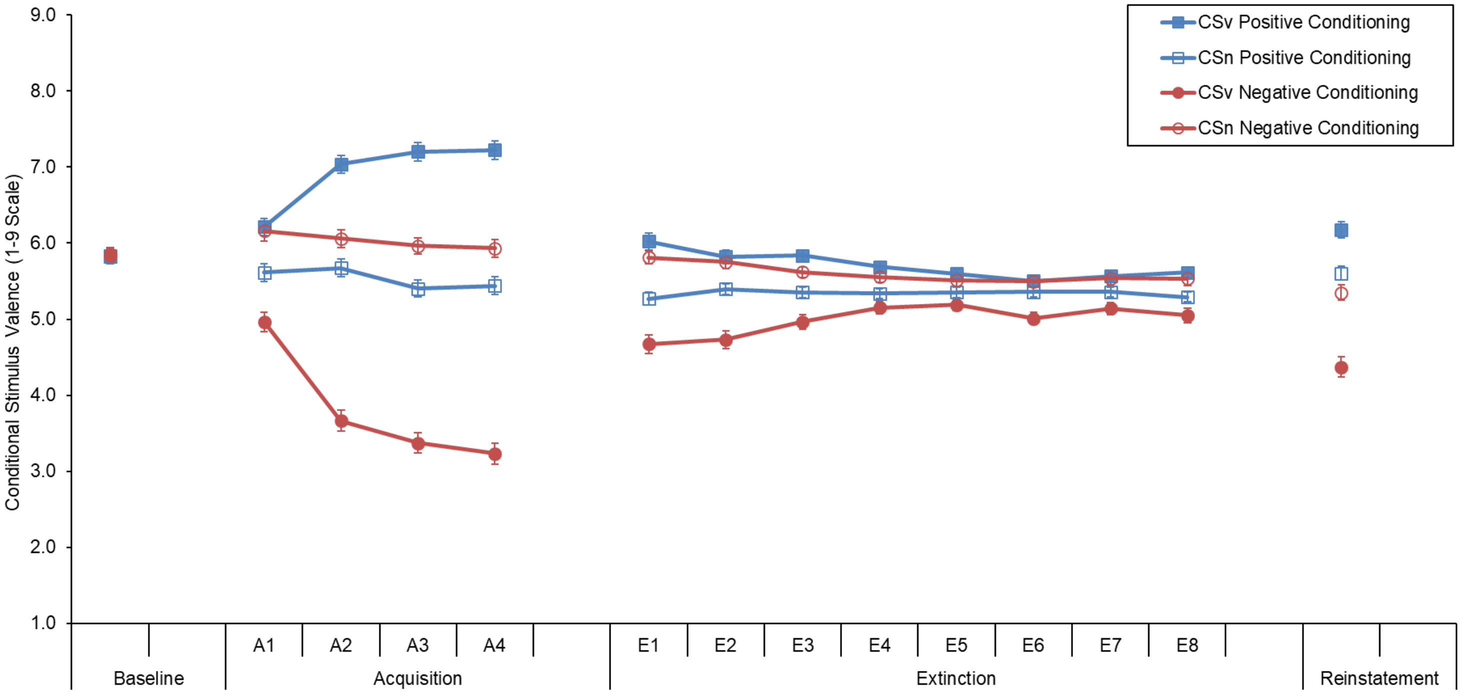

Figure 7. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the reinstatement test of Experiment 3b. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).

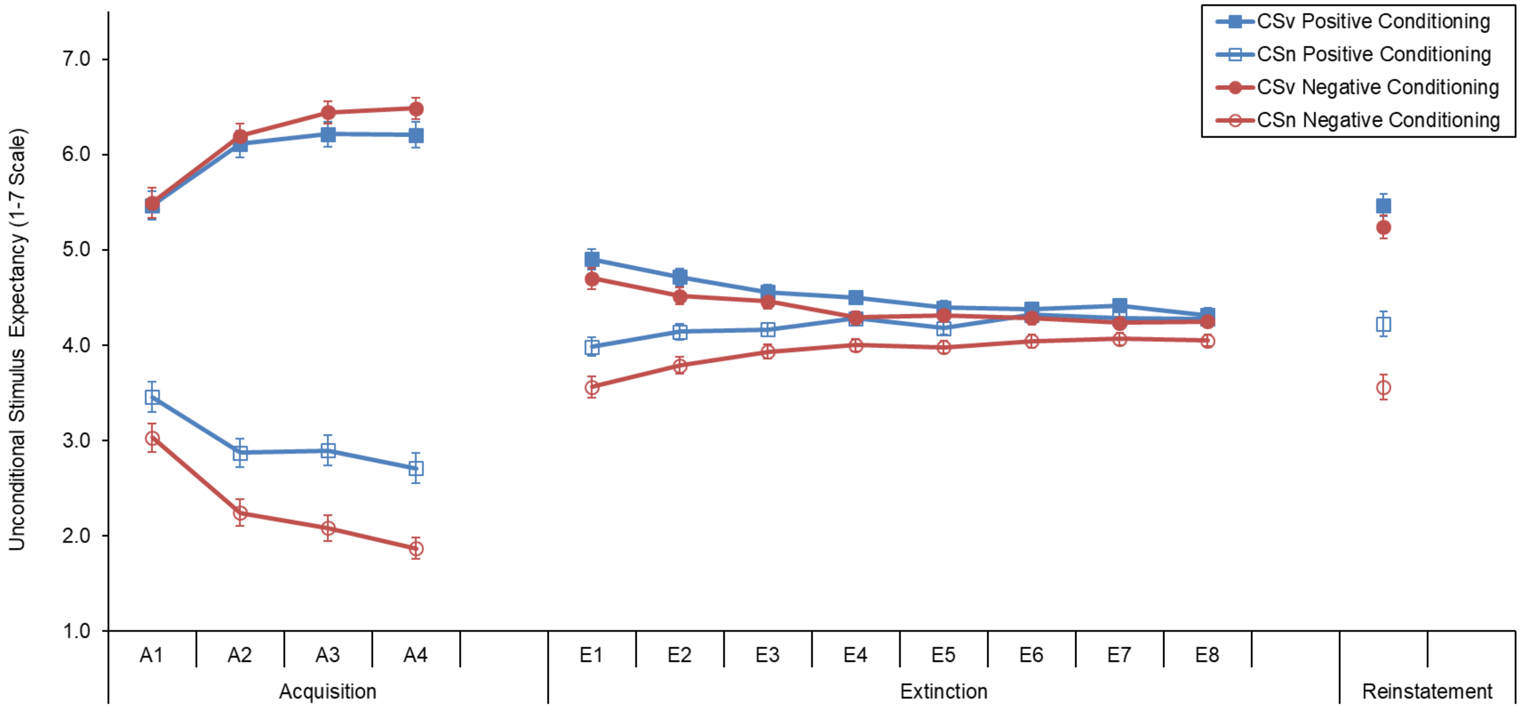

Figure 8. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the reinstatement test of Experiment 3b. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Ratings lower than 4 reflect a prediction of neutral images, ratings higher than 4 reflect a prediction of pleasant/unpleasant images, and a rating of 4 reflects a prediction of no images. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).

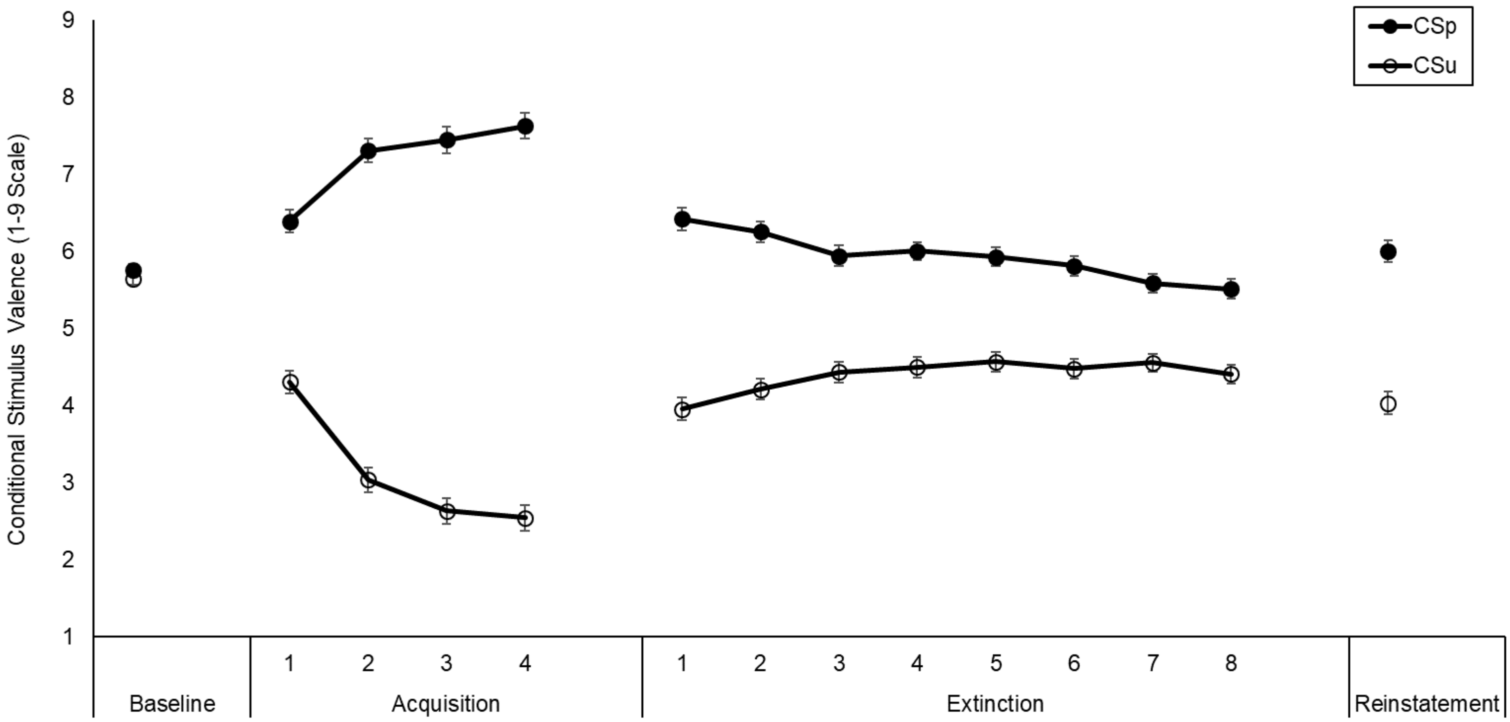

Figure 9. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the reinstatement test of Experiment 4. During acquisition, the CSp was paired with pleasant images and the CSu was paired with unpleasant images. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).

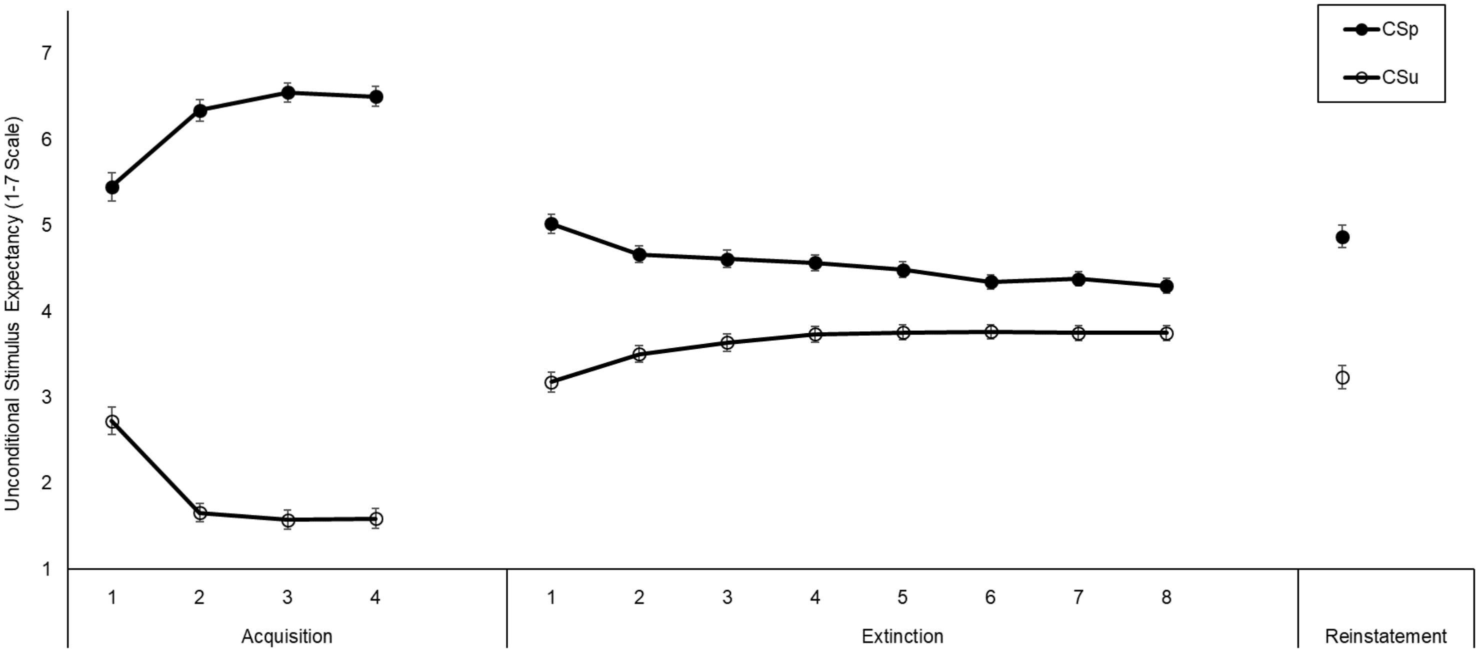

Figure 10. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the reinstatement test of Experiment 4. During acquisition, the CSp was paired with pleasant images and the CSu was paired with unpleasant images. Ratings lower than 4 reflect a prediction of unpleasant images, ratings higher than 4 reflect a prediction of pleasant images, and a rating of 4 reflects a prediction of no images. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

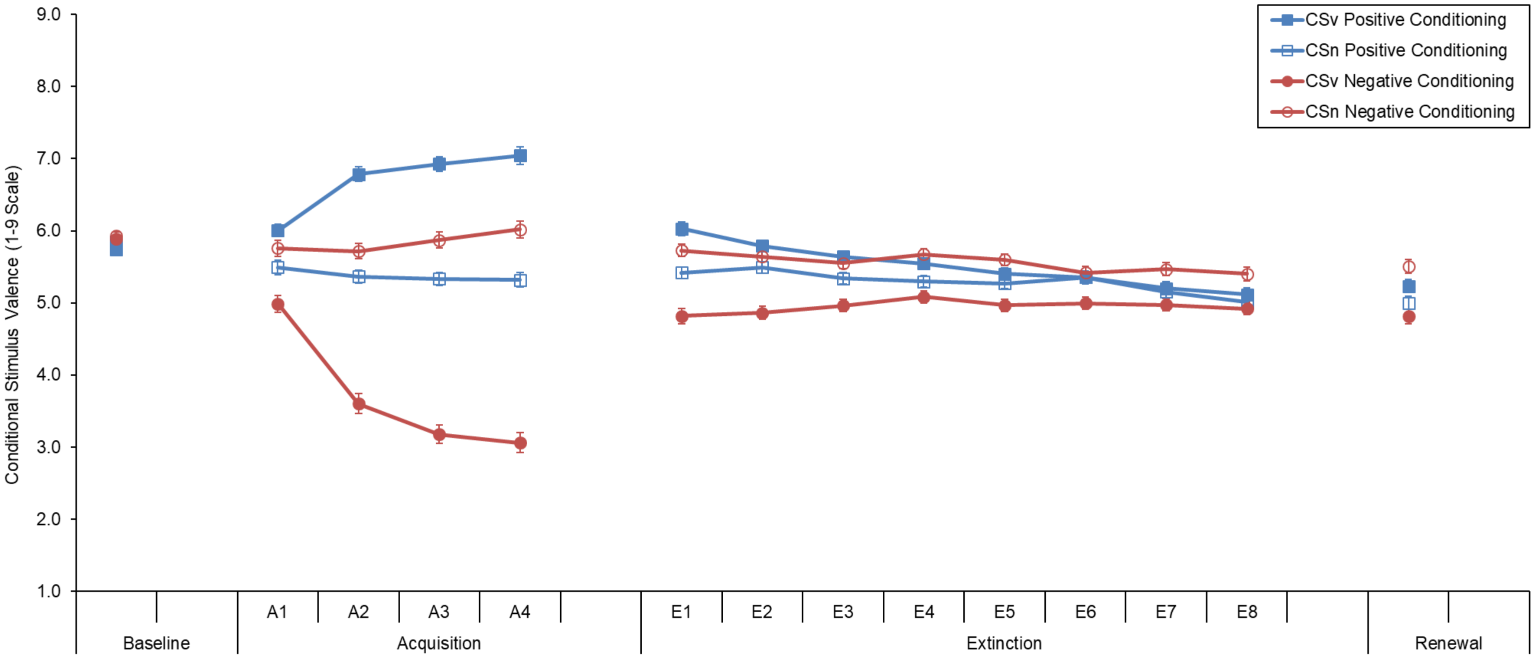

Figure 11. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the renewal test of Experiment 5. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

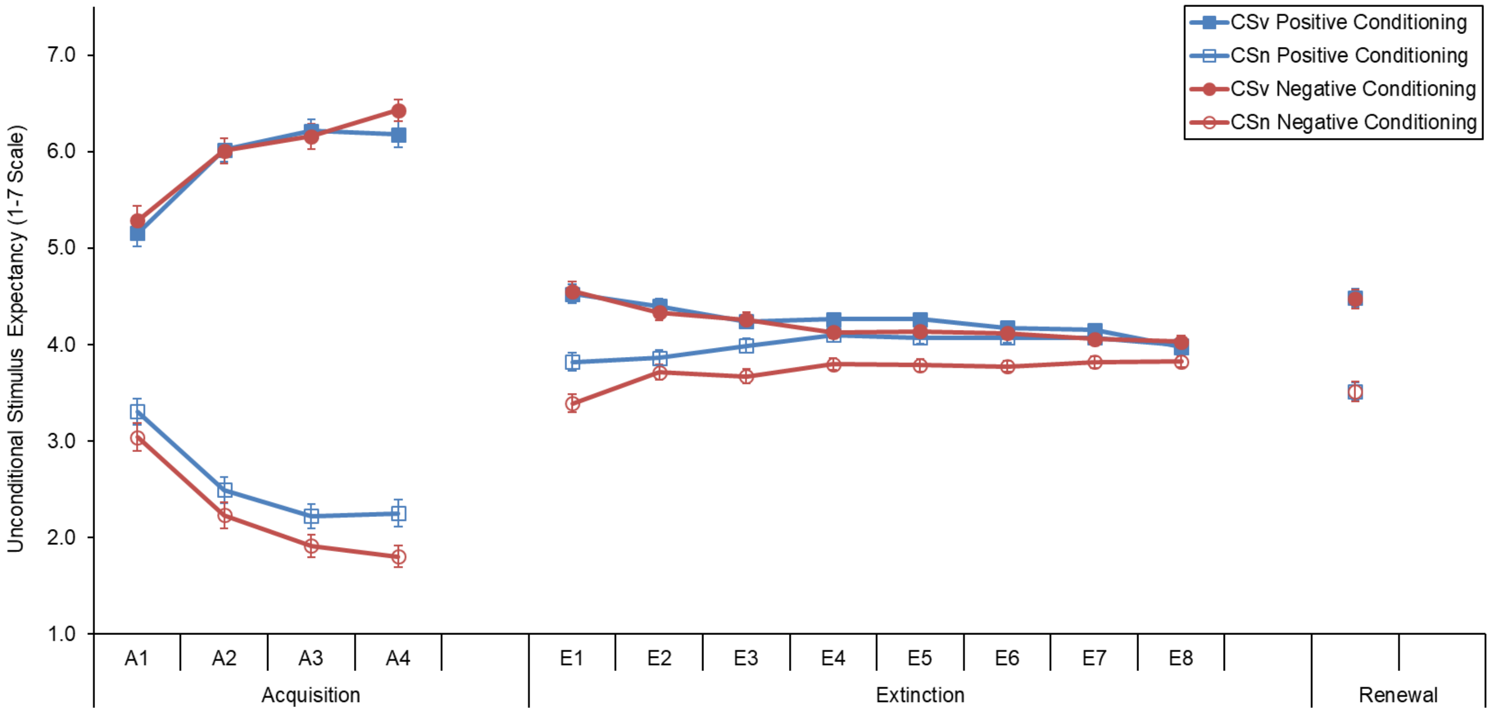

Figure 12. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the renewal test of Experiment 5. During acquisition, the CSv was paired with pleasant images in the positive conditioning group and unpleasant images in the negative conditioning group, while the CSn was paired with neutral images in both groups. Ratings lower than 4 reflect a prediction of neutral images, ratings higher than 4 reflect a prediction of pleasant/unpleasant images, and a rating of 4 reflects a prediction of no images. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

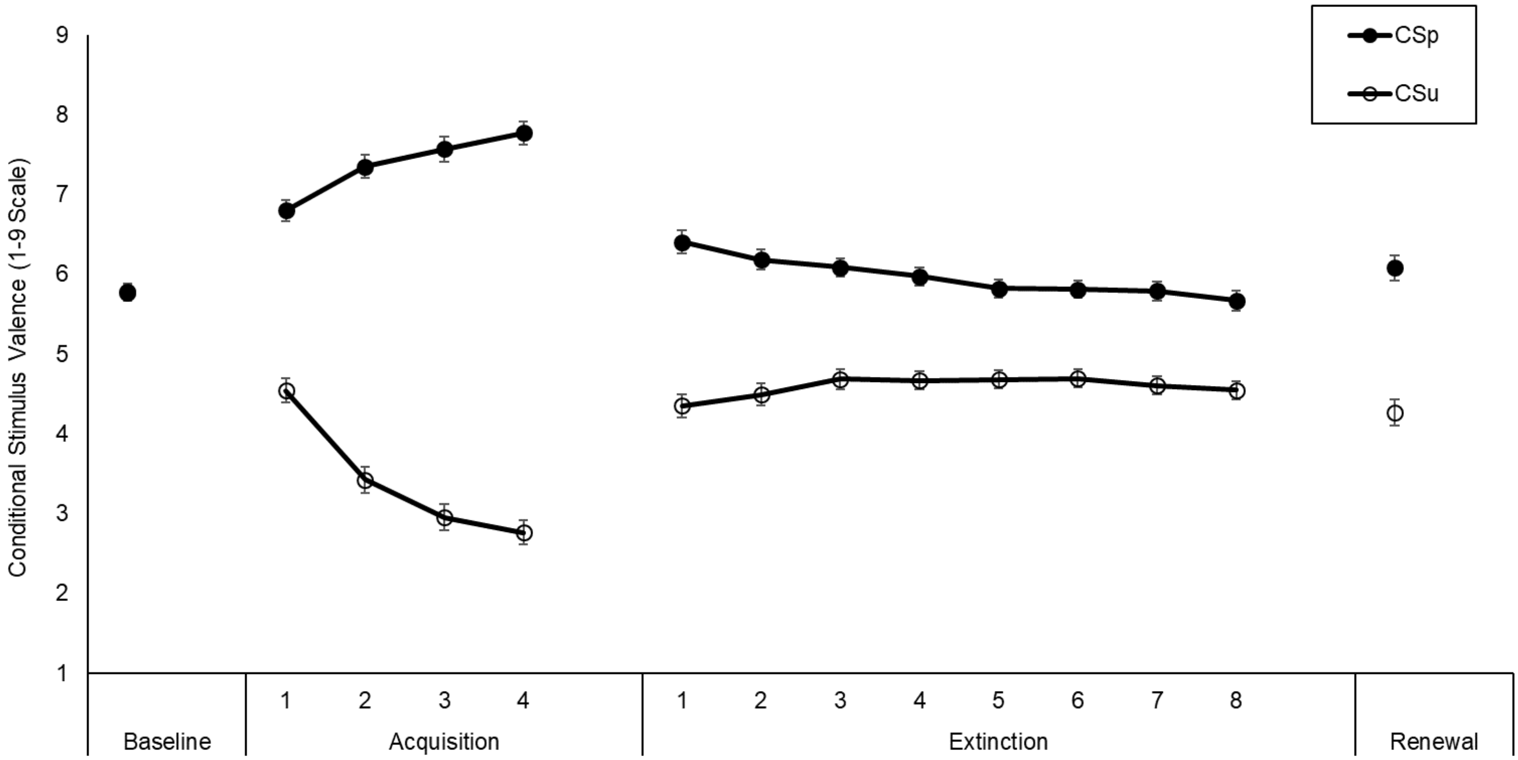

Figure 13. Conditional stimulus valence evaluations measured during baseline, acquisition, extinction, and the renewal test of Experiment 6. During acquisition, the CSp was paired with pleasant images and the CSu was paired with unpleasant images. Higher values indicate higher stimulus pleasantness. Error bars represent SEMs for within-participant designs based on O’Brien and Cousineau (2014).

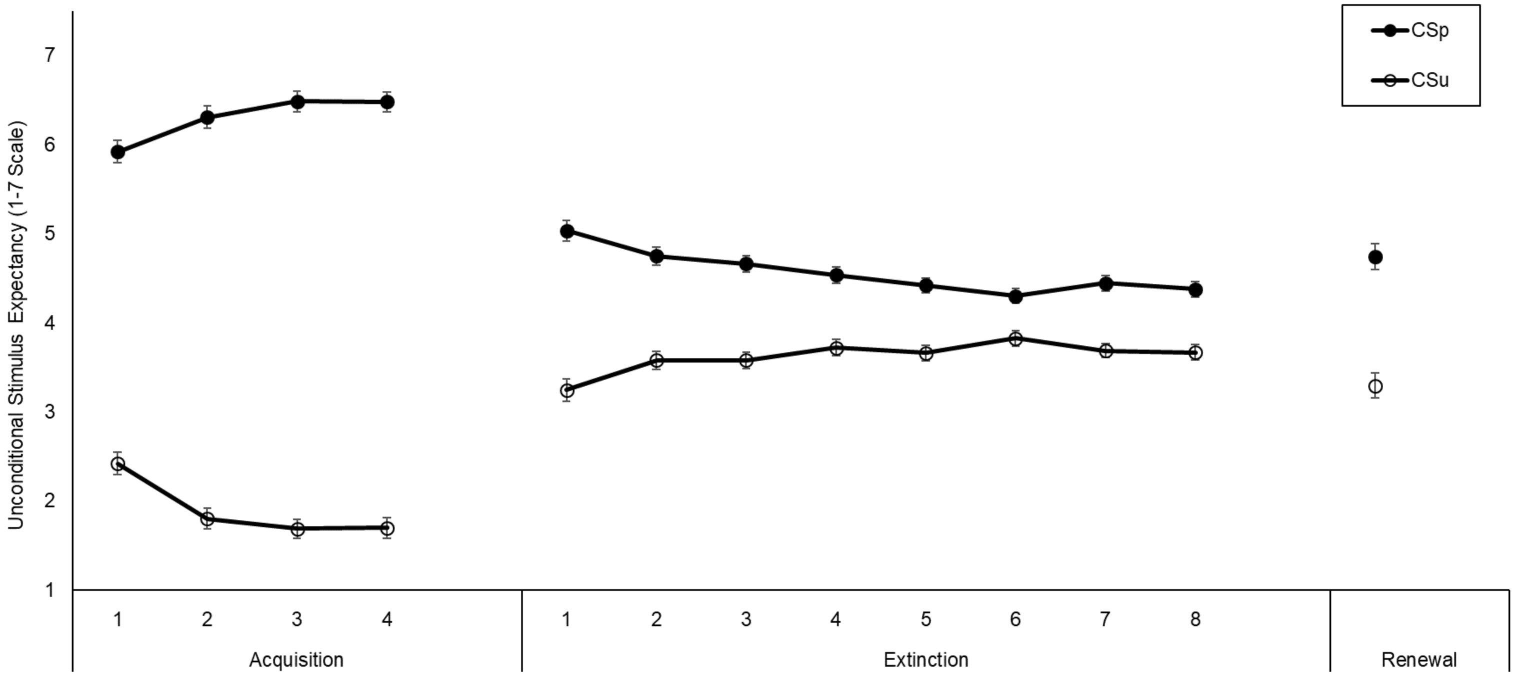

Figure 14. Unconditional stimulus expectancy ratings measured during acquisition, extinction, and the renewal test of Experiment 6. During acquisition, the CSp was paired with pleasant images and the CSu was paired with unpleasant images. Ratings lower than 4 reflect a prediction of unpleasant images, ratings higher than 4 reflect a prediction of pleasant images, and a rating of 4 reflects a prediction of no images. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).