

Is the Devil in the Detail? Evidence for S-S Learning after Unconditional Stimulus Revaluation
in Human Evaluative Conditioning under a Broader set of Experimental Conditions

Hannah Jensen-Fielding¹, Camilla C. Luck^{2,3} & Ottmar V. Lipp^{2,3}

¹School of Psychology, University of Queensland, Australia

² School of Psychology and Speech Pathology, Curtin University, Australia

³ ARC-SRI: Science of Learning Research Centre

Author Notes

Address for correspondence:

Hannah Jensen-Fielding, School of Psychology, University of Queensland, St Lucia, QLD 4072,
Australia, E-Mail: hannah.jensenfielding@gmail.com

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

Ottmar V. Lipp, School of Psychology and Speech Pathology, Curtin University, GPO Box
U1987 Perth WA 6845, Australia. E-mail: Ottmar.Lipp@curtin.edu.au

Acknowledgements:

This work was supported by grants number DP110100460 and SR120300015 from the
Australian Research Council.

Word count: 10643

Abstract

Whether valence change during evaluative conditioning is mediated by a link between the conditional stimulus (CS) and the unconditional stimulus (US; S-S learning) or between the CS and the unconditional response (S-R learning) is a matter of continued debate. Changing the valence of the US after conditioning, known as US revaluation, can be used to dissociate these accounts. Changes in CS valence after US revaluation provide evidence for S-S learning but if CS valence does not change, evidence for S-R learning is found. Support for S-S learning has been provided by most past revaluation studies, but typically the CS and US have been from the same stimulus category, the task instructions have suggested that judgements of the CS should be based on the US, and USs have been mildly valenced stimuli. These factors may bias the results in favor of S-S learning. We examined whether S-R learning would be evident when CSs and USs were taken from different categories, the task instructions were removed, and more salient USs were used. US revaluation was found to influence explicit US evaluations and explicit and implicit CS evaluations, supporting an S-S learning account and suggesting that past results are stable across procedural changes.

Key words: Unconditional Stimulus Revaluation; Evaluative Conditioning; Signal-Signal Learning; Signal-Response Learning; Valence; Associative Structure

Likes and dislikes can be acquired and changed via a type of conditioning known as evaluative conditioning. When a neutral stimulus (conditional stimulus, CS) is paired with a positively or negatively valenced stimulus (unconditional stimulus, US) this neutral stimulus eventually takes on positive or negative valence. Procedurally, evaluative conditioning is similar to other expectancy based forms of conditioning, however, considerable debate exists about whether evaluative conditioning differs on a process level. These debates are typically driven by discrepant findings between evaluative and expectancy conditioning paradigms, however, considerable discrepancies also exist within the evaluative conditioning literature itself. For example, many report that evaluative conditioning does not extinguish (Baeyens, Crombez, Van den Bergh, & Eelen, 1988), while others report complete extinction (Lipp & Purkis, 2006). Another key topic of debate that could potentially account for some of these discrepancies is whether evaluative conditioning is mediated by a link between the CS and US representations (signal-signal; S-S learning) or by a link between the CS representation and the evaluative response (signal-response; S-R learning). If valence change during evaluative conditioning is mediated by an S-S link, evaluative responding should extinguish when the US is no longer presented, but if the underlying associative structure is S-R, evaluative conditioning does not depend on a specific CS-US association and therefore may be less affected by an extinction procedure.

The most reliable way of examining whether S-S or S-R learning mediates valence change during evaluative conditioning is with a US revaluation procedure (Rescorla, 1974). US revaluation involves changing the valence of the US after conditioning and then examining whether this change influences the valence of the CS. US revaluation procedures typically involve three phases – an initial conditioning training involving CS-US pairings, the US

revaluation phase in which the US is made more positive or negative in the absence of the CS, and a subsequent CS valence test phase before any additional CS-US pairings. Evidence for S-S learning is found when CS valence changes after revaluation in the same direction as the new US valence. Conversely, evidence for S-R learning can be inferred when CS valence does not change after revaluation.

Baeyens, Eelen, Van den Bergh, and Crombez (1992) were the first to examine US revaluation during evaluative conditioning. The authors used liked and disliked faces as USs based on participants' pre-rated valence scores and then rendered these faces more pleasant or unpleasant by presenting them simultaneously with pleasant and unpleasant personality characteristics (i.e. friendly and cruel). After this initial attitude formation phase, neutral faces (CSs) were paired with the USs in an evaluative conditioning procedure. After conditioning, one pleasant and one unpleasant US were congruently revalued (i.e. paired with pleasant characteristics for the pleasant US and unpleasant characteristics for the unpleasant US), and one pleasant and one unpleasant US were incongruently revalued (i.e. paired with unpleasant characteristics for the pleasant US and pleasant characteristics for the unpleasant US). US revaluation successfully changed US valence for incongruently revalued stimuli, such that both USs became neutral stimuli. Furthermore, this change in US valence brought about an immediate change in CS valence, suggesting that evaluative conditioning effects were mediated by an S-S learning structure.

Walther, Gawronski, Blank, and Langer (2009) replicated and extended this finding using an implicit measure of valence and controlling for the possible influence of CS-US similarity. The experimental design was similar to Baeyens et al. (1992) with already liked and disliked faces being made more pleasant or unpleasant by pairing them with pleasant or unpleasant

personality characteristics in an initial conditioning phase. Participants were instructed to form impressions about these individuals (US faces) based on these personality descriptions. After this attitude formation phase, participants were shown two faces and asked to imagine that they were acquainted with some of these (US faces), but unfamiliar with others (CS faces), and that their task was to form impressions of the individuals presented on the screen. After this conditioning phase, participants were asked to imagine that they had been working at a company with these individuals for several weeks and would now receive additional information about their colleagues. The US faces were then paired with information of opposite (revaluation) or neutral valence (control) and CS valence was tested. Replicating Baeyens et al. (1992), CS valence changed immediately after US revaluation (S-S learning).

The results of Walther et al. (2009) and Baeyens et al. (1992) suggest that evaluative conditioning is mediated by S-S learning – but not all evidence consistently supports S-S accounts. Baeyens, Vanhouche, Crombez, and Eelen (1998) found no influence of a US inflation procedure (an unpleasant US made more unpleasant) on CS valence in evaluative flavor-flavor conditioning – suggesting an S-R link structure. Sweldens, Van Osselaer, and Janiszewski (2010) investigated whether different aspects of the conditioning procedure (conditioning with simultaneous or sequential pairings and conditioning with a single or multiple USs) determine whether S-S or S-R learning would occur. Participants were presented with pictures of beers (CSs) and pictures of adults having fun (positive USs). After conditioning, some participants were told that ‘inconspicuous looking people can commit serious crimes’ and crimes were shown on the screen with half of the positive US pictures (e.g. murderer, rapist). US revaluation changed attitudes to the CS when conditioning was performed with only one US in both sequential and simultaneous procedures (S-S learning), however, revaluation did not change

attitudes to the CS when each CS had been paired with multiple USs during conditioning in a sequential pairings procedure (S-R learning). Evaluative conditioning was not obtained when each CS was paired with multiple USs in a sequential conditioning procedure. These results suggest that different procedural aspects of the conditioning design could critically change the nature of evaluative learning. Similarly, Gast and Rothermund (2011a) found evidence for both S-S and S-R learning in evaluative conditioning, depending on whether or not participants were asked to make evaluations during the conditioning phase. Gast and Rothermund (2011a) used a similar design to Walther et al. (2009). In Experiment 1, liked and disliked individuals were paired with pleasant and unpleasant character statements (attitude formation phase), before participants received pairings of neutral individuals (CSs) with these USs. Participants were asked to form an impression of the *two colleagues together* (CS and US) and to indicate whether they had a ‘more positive or more negative’ impression of them. After conditioning, half of the US faces were combined with statements of opposite valence (incongruent revaluation), while, the other half were combined with statements of the same valence (congruent revaluation). After incongruent revaluation, implicit evaluations of the pleasant and unpleasant USs did not differ, while explicit US evaluations were reduced in magnitude. Interestingly, unlike in previous studies, implicit and explicit CS evaluations were not affected by US revaluation (suggesting S-R learning). Gast and Rothermund (2011a) argue that for S-R learning to occur participants must be in an ‘evaluative mindset’ during conditioning– which they argue occurs when participants are asked to provide evaluative judgments during conditioning (see Gast and Rothermund, 2011b). In Experiment 2, the participants did not give evaluative ratings during conditioning and US revaluation was found to change explicit, but not implicit, US and CS valence. The lack of CS valence change after revaluation in Experiment 1 provides evidence for S-R learning. The

results of Sweldens et al. (2010) and Gast and Rothermund (2011a) seem to indicate that seemingly minor procedural details during evaluative conditioning could critically change the nature of the learning.

US revaluation during evaluative conditioning has only been examined under a very restricted set of conditions. In light of Sweldens et al. (2010) and Gast and Rothermund (2011a) it seems critically important to extend these and to explore further whether procedural variations may change US revaluation effects. We have observed a number of procedural details common to evaluative conditioning revaluation studies which may favor S-S learning. The CS and the US have previously both been a logical match (e.g. faces with faces or beer with fun activities). The use of a logical match between the CS and US could encourage participants to learn that the ‘CS is like the US’ more than if the CS and the US were from different picture categories (i.e. shapes and faces). Similarly, the instructions given to participants before conditioning may suggest that they should base their impression of the CS on the US itself. For instance, Walther et al. (2009) told participants that they were working in a new company and were familiar with some individuals while unacquainted with others and Gast and Rothermund (2011a) asked participants to form an impression of the ‘two colleagues together’. Although subtle, these changes might alter the nature of the learning by suggesting that the impression of the CS should be based on the US. In previous studies, pleasant and unpleasant USs have typically been faces with neutral expressions picked based on participant’s pre-rating scores. These USs were then rendered more positive or negative by pairing them with pleasant or unpleasant personality characteristics. Gast and Rothermund (2011a) argue that S-R learning may be more likely to occur when the US triggers a strong evaluative response during conditioning (but also see misattribution theory for an alternate account; Jones, Fazio, & Olson, 2009). The use of USs that acquire some of their

valence throughout pairings with pleasant and unpleasant words immediately before the conditioning phase may not trigger evaluative responses as strongly as faces that have clear a-priori valence to all participants, such as happy and angry faces.

In the current study we aimed to examine the influence of US revaluation in evaluative conditioning under a broader set of experimental conditions. Specifically, the study aimed to examine whether S-R learning would occur when: 1) strong a-priori pleasant and unpleasant happy and angry faces were used as USs; 2) arbitrary stimuli that are not easily related to human faces (i.e. geometrical shapes) were used as the CSs; and 3) participants were not given task instructions suggesting that they should form an impression of the CS based on the US. A within-participants design was used, in which two CSs were paired with happy faces and two CSs were paired with angry faces during conditioning. After conditioning, one happy face and one angry face (incongruent revaluation) were revaluated with a character story of opposite valence (i.e. happy face, unpleasant story; angry face, pleasant story); while the other happy and angry face (control) were paired with a character story of the same valence (happy face, pleasant story; angry face, unpleasant story). Following revaluation, CS valence was tested before any additional CS-US pairings. After this revaluation test, additional CS-US pairings were administered to examine the influence of additional training on CS valence. Similar to Gast and Rothermund (2011a; Experiment 1), participants were asked to evaluate CS valence throughout conditioning, however, participants were only asked to evaluate CS valence, rather than the valence of the CS and the US together (to avoid a 'CS is like the US' instructional set). A measure of US expectancy was included during conditioning and an implicit measure of US valence was included after the experiment. If procedural details could drive the type of learning occurring during evaluative conditioning, we hypothesise that these changes would make S-R

learning mediation more likely and therefore US revaluation will not influence CS valence immediately after revaluation, but that the CS will acquire the new US valence after additional CS-US pairings.

Experiment 1

Method

Participants and Design. Thirty-five (21 female) undergraduate students aged 17-35 years ($M = 19.62$) provided informed consent and volunteered participation in exchange for course credit. The experimental procedure was approved by the University Ethics Review Board. Sample size was determined based on prior experiments that had used similar procedures (Lipp & Purkis, 2006). The experiment employed a 2 US Valence (Expression: happy, angry) \times 2 Revaluation (revaluation, control) fully within participants experimental design.

Apparatus/Stimuli. The experiment was run in a group computer lab on PCs running Windows XP and programmed in DMDX (Forster & Forster, 2003). Responses were recorded via a QWERTY keyboard. Color pictures of four Caucasian male adults displaying angry and happy facial expressions (NimStim database: images M_HA_O and M_AN_O: models 22, 21, 32, 36, Tottenham et al., 2009) were used as USs. The pictures were 360×464 pixels in size, and were displayed against a black background on a 17 inch CRT screen (resolution: 1024×768 pixels, refresh rate: 85 Hz). Four geometric shapes (circle, diamond, square, and triangle) were used as CSs. Shapes were 450×338 pixels in size, with a black outline against a white background. The facial expression displayed by a particular model (happy vs. angry), and the shape paired with a particular US were counterbalanced across participants.

Procedure. The different stages of the experimental protocol are displayed in Figure 1.

Participants were given an initial briefing, provided informed consent and were tested in groups of up to six. Before conditioning training, participants completed explicit and implicit measures of US valence. During the explicit valence measure the four US faces were displayed on the screen and participants were asked to indicate how pleasant/unpleasant they found the faces on a nine point Likert scale (1 = *unpleasant*, 9 = *pleasant*). Affective priming was used to assess the implicit valence of the US faces. In this task, participants were asked to evaluate the valence of target words that were preceded by the US faces by pressing the right or left shift keys. On each trial, a US face was presented for 140 ms followed by a blank screen for 70 ms and a pleasant or unpleasant target word presented until a response was made (or up to a maximum of 882 ms if a response was not made). Six pleasant (appealing, charming, desirable, favourable, nice and enjoyable) and six unpleasant (annoying, disturbing, inferior, nasty, repulsive, and terrifying) target words were used. US-word pairs were presented in a random order and each pair was presented twice, forming 96 trials in total. After the two US valence measures, participants rated the CS shapes on a nine point Likert scale (1 = *unpleasant*, 9 = *pleasant*). The conditioning procedure consisted of three general phases – an initial conditioning training procedure, a revaluation phase, and a post-revaluation conditioning procedure. In both conditioning phases, participants were presented with four CS-US pairs (CS1-Happy-Control; CS2-Happy-Revaluation; CS3-Angry-Control; CS4-Angry-Revaluation). CSs were presented for 2s followed immediately by a 2s presentation of the US face. Conditioning trials were separated by an inter-trial interval of 4s. The initial conditioning training consisted of three blocks of 24 randomised conditioning trials (blocks 1-3; 6 presentations of each CS-US pair). After each block, participants provided CS valence ratings and CS-US contingency judgements. During the contingency judgments each CS was shown twice and participants were asked to indicate the

percentage of time (0% = *never*, 100% = *always*) the CS was paired with a happy face and the percentage of time the CS was paired with an angry face. After the third training block, verbal instructions (see Table 1) were displayed on the screen together with one of the US faces in a random order. Participants received positive character information about one happy (control) and one angry (revaluation) US face; and negative character information about one happy (revaluation) and one angry (control) US face. After the first set of instructions, participants completed a memory test and a repetition of the instructions. During the memory test each US face was presented with the question ‘this person (T) teaches in Africa, (H) volunteers in hospitals, (A) committed assault, or (S) stole money’. After the second set of instructions, explicit CS valence was assessed to examine the influence of US revaluation before any additional CS-US pairings. After revaluation, participants completed three additional blocks of post-revaluation CS-US pairings (blocks 4-6; 6 presentations of each CS-US pairing per block). CS valence and contingency judgements were assessed after each training block and after the fourth training block the US revaluation instructions were repeated and an additional CS valence test was administered. After the last set of CS valence and contingency assessments, participants completed a second set of explicit and implicit US valence assessments (identical to the US assessments taken before conditioning) and a second instruction memory test.

Coding and Statistical Analyses. The instruction memory test scores were coded with a 1 for correct and a 0 for incorrect. An overall contingency score was obtained by subtracting the contingency judgment for an angry face from the contingency judgement for a happy face (i.e. CS1 happy judgement – CS1 angry judgement) resulting in scores ranging from -100 to 100 (-100 = *strongly predicting an angry face*, 100 = *strongly predicting a happy face*). For the affective priming task, reaction times faster than 100ms and outside of 3 standard deviations

from a participant's individual mean were removed. All analyses were conducted with IBM SPSS Statistics 24 with an alpha cut-off of .05, interactions have been followed-up with simple effect contrasts, and Pillai's trace statistics of the multivariate solution are reported.

Results

Explicit Unconditional Stimulus Valence

The pre and post US valence assessments were analyzed with a 2 Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times 2 Time (pre, post) within-participants repeated measures ANOVA and are displayed in Figure 2. Missing scores resulted in the loss of 3 cases. A main effect of expression, $F(1, 31) = 187.01, p < .001, \eta^2 = .858$, a main effect of time, $F(1, 31) = 4.44, p = .043, \eta^2 = .125$, and an Expression \times Revaluation interaction, $F(1, 31) = 24.78, p < .001, \eta^2 = .444$, were moderated by an Expression \times Revaluation \times Time interaction, $F(1, 31) = 43.43, p < .001, \eta^2 = .584$. Before revaluation, happy faces were evaluated as more pleasant than angry faces for both control, $F(1, 31) = 59.95, p < .001, \eta^2 = .659$, and to be revalued stimuli, $F(1, 31) = 86.26, p < .001, \eta^2 = .736$. After revaluation, happy faces were evaluated as more pleasant than angry faces for control stimuli, $F(1, 31) = 268.39, p < .001, \eta^2 = .896$, but evaluations of happy and angry faces did not differ for revaluated stimuli, $F(1, 31) = 0.21, p = .649, \eta^2 = .007$. Before revaluation, control and to be revaluated happy faces, $F(1, 31) = 1.44, p = .239, \eta^2 = .044$, and control and to be revaluated angry faces, $F(1, 31) = 0.17, p = .680, \eta^2 = .006$, did not differ. After revaluation, revalued happy faces were rated as less pleasant than control happy faces, $F(1, 31) = 46.24, p < .001, \eta^2 = .599$, and revalued angry faces were evaluated as more pleasant than control angry faces, $F(1, 31) = 23.17, p < .001, \eta^2 = .428$. The remaining main effects and interactions did not attain significance, all $F < 3.80, p > .060, \eta^2 < .108$ (Revaluation \times Time).

Implicit Unconditional Stimulus Valence

Three participants were excluded for making more than 40% errors in either priming task. Mean priming reaction times were subjected to a 2 Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times 2 Target word (pleasant, unpleasant) \times 2 Time (pre, post) within-participants repeated measures ANOVA. Target evaluation time decreased from before to after the experiment, $F(1, 30) = 6.12, p = .019, \eta^2 = .170$, and positive target words were categorised faster than negative target words, $F(1, 30) = 6.33, p = .017, \eta^2 = .174$. Participants were faster to categorise target words after US faces that were paired with positive information in comparison to US faces that were paired with negative information, $F(1, 30) = 6.64, p = .015, \eta^2 = .181$ (note: as this main effect is not qualified by target word valence it does not provide evidence for priming). An Expression \times Target Word Valence interaction, $F(1, 30) = 10.24, p = .003, \eta^2 = .255$ was detected. The priming score for this interaction ([incongruent: happy-unpleasant target + angry-pleasant target] – [congruent: happy-pleasant target + angry-unpleasant target]) was positive ($M = 152.63$ ms, $SD = 265.54$ ms). These results indicate that affective priming was driven by the valence of the US facial expression and not modulated by phase or revaluation. The remaining main effects and interactions did not attain significance, all $F < 2.64, p > .114, \eta^2 < .079$.

Instruction Memory

Instruction memory scores were subjected to an Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times 2 Time (test 1, test 2) within-participants repeated measures ANOVA. A main effect of time, $F(1, 34) = 7.73, p = .009, \eta^2 = .185$, and an Expression \times Revaluation, $F(1, 34) = 8.57, p = .006, \eta^2 = .201$, were moderated by a marginal Expression \times Revaluation \times Time interaction, $F(1, 34) = 4.13, p = .050, \eta^2 = .108$. Follow-up analyses

revealed that during the first test, memory was better for instructions paired with control happy ($M = 0.94$, $SD = 0.24$) faces than instructions paired revalued happy faces ($M = 0.74$, $SD = 0.44$), $F(1, 34) = 6.26$, $p = .017$, $\eta^2 = .156$, but memory was worse for instructions paired with control angry faces ($M = 0.71$, $SD = 0.46$) than for instructions paired with revalued angry faces ($M = 0.94$, $SD = 0.24$), $F(1, 34) = 6.11$, $p = .019$, $\eta^2 = .152$. During the second memory test, memory did not differ for instructions paired with control happy ($M = 1.00$, $SD < 0.01$) or revalued happy faces ($M = 0.94$, $SD = 0.24$), $F(1, 34) = 2.06$, $p = .160$, $\eta^2 = .057$, or for control angry ($M = 0.94$, $SD = 0.24$) or revalued angry faces ($M = 0.97$, $SD = 0.17$) USs, $F(1, 34) = 0.33$, $p = .571$, $\eta^2 = .010$. This pattern of results suggests that positive information was remembered better than negative information during the first memory test. The remaining main effects and interactions did not attain significance, all $F < 0.67$, $p > .421$, $\eta^2 < .020$.

Conditional Stimulus Valence.

The CS valence evaluations taken before conditioning, during the initial conditioning phase, the revaluation test, and the post-revaluation conditioning phase are displayed in Figure 3.

Conditioning Phase. The CS valence ratings measured before and during the initial conditioning phase (blocks 1-3) were subjected to a 2 Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times 4 Block (pre, 1, 2, 3) repeated measures ANOVA. Missing responses resulted in the loss of 2 cases. Main effects of expression, $F(1, 32) = 25.51$, $p < .001$, $\eta^2 = .444$, and block, $F(3, 30) = 8.18$, $p < .001$, $\eta^2 = .450$, were moderated by an Expression \times Block interaction, $F(3, 30) = 8.65$, $p < .001$, $\eta^2 = .464$. Before conditioning, CSs paired with angry faces and happy faces did not differ in valence, $F(1, 32) = 0.26$, $p = .612$, $\eta^2 = .008$, but during blocks one, $F(1, 32) = 6.13$, $p = .019$, $\eta^2 = .161$, two, $F(1, 32) = 27.94$, $p < .001$, $\eta^2 = .466$, and three, $F(1, 32) = 29.72$, $p < .001$, $\eta^2 = .482$, CSs paired with angry faces were evaluated as less

pleasant than CSs paired with happy faces. The remaining main effects and interactions did not attain significance, all $F < 1.94$, $p > .173$, $\eta^2 < .058$.

Unconditional Stimulus Revaluation. The influence of the revaluation manipulation (before any additional pairings) was examined with a 2 Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times 2 Phase (last block of conditioning training, revaluation test) repeated measures ANOVA. Missing responses resulted in the loss of 2 cases. A main effect of expression, $F(1, 32) = 22.98$, $p < .001$, $\eta^2 = .418$, an Expression \times Revaluation interaction, $F(1, 32) = 8.77$, $p = .006$, $\eta^2 = .215$, and an Expression \times Phase interaction, $F(1, 32) = 21.53$, $p < .001$, $\eta^2 = .402$, were moderated by an Expression \times Revaluation \times Phase interaction, $F(1, 32) = 9.40$, $p = .004$, $\eta^2 = .227$. Follow-up analyses revealed that during the last block of conditioning CSs paired with angry faces were evaluated as less pleasant than CSs paired with happy faces for both revaluation, $F(1, 32) = 16.76$, $p < .001$, $\eta^2 = .344$, and control stimuli, $F(1, 32) = 33.98$, $p < .001$, $\eta^2 = .515$. After revaluation, this differential valence was no longer present for CSs paired with revaluated faces, $F(1, 32) = 0.02$, $p = .897$, $\eta^2 = .001$, but was still intact for CSs paired with control faces, $F(1, 32) = 19.51$, $p < .001$, $\eta^2 = .379$.

During the last conditioning block, control and to be revaluated happy, $F(1, 32) = 0.90$, $p = .351$, $\eta^2 = .027$, and angry faces, $F(1, 32) = 1.00$, $p = .325$, $\eta^2 = .030$, did not differ, but after revaluation, CSs paired with control happy faces were evaluated as more pleasant than CSs paired with revaluated happy faces, $F(1, 32) = 13.91$, $p = .001$, $\eta^2 = .303$. CSs paired with control and revaluated angry faces did not differ, $F(1, 32) = 1.40$, $p = .246$, $\eta^2 = .042$. A comparison across phase confirmed that CSs paired with revaluated angry faces increased in pleasantness from the last block of conditioning to the revaluation test, $F(1, 32) = 12.21$, $p = .001$, $\eta^2 = .276$. This increase was not significant for CSs paired with control angry faces, $F(1, 32) = 2.73$, $p = .108$,

$\eta^2 = .079$. The remaining main effects and interactions did not attain significance, all $F < 1.52$, $p > .228$, $\eta^2 < .046$.

Post-Revaluation Conditioning Phase. The CS valence evaluations measured during the post-revaluation pairings and after the repeat of the instructions (block 4, instruction repeat, blocks 5-6) were subjected to a 2 Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times 4 Block (4, instruction repeat 5, 6,) repeated measures ANOVA. Missing responses resulted in the loss of 1 case. A main effect of expression, $F(1, 33) = 17.81$, $p < .001$, $\eta^2 = .351$, was moderated by an Expression \times Revaluation interaction, $F(1, 33) = 36.75$, $p < .001$, $\eta^2 = .527$. Follow-up analyses revealed that CSs paired with happy control faces were evaluated as more pleasant than CSs paired with angry control faces, $F(1, 33) = 55.45$, $p < .001$, $\eta^2 = .627$. CSs paired with revaluated angry faces and revaluated happy faces did not differ, $F(1, 33) = 2.16$, $p = .151$, $\eta^2 = .061$. A comparison across revaluation conditions confirmed that CSs paired with revaluated happy faces were less pleasant than CSs paired with control happy faces, $F(1, 33) = 36.13$, $p < .001$, $\eta^2 = .523$, and that CSs paired with revaluated angry faces were more pleasant than CSs paired with control angry faces, $F(1, 33) = 18.36$, $p < .001$, $\eta^2 = .358$. The remaining main effects and interactions did not attain significance, all $F < 3.30$, $p > .078$, $\eta^2 < .092$ (main effect of revaluation).

Contingency Judgments.

The contingency judgements measured during the initial conditioning phase and the post-revaluation conditioning phase were subjected to separate 2 Expression (happy, angry) \times 2 Revaluation (revaluation, control) \times Block (initial conditioning: 1, 2, 3; post-revaluation conditioning: 4, 5, 6) repeated measures ANOVAs and are displayed in Figure 4. Missing

responses resulted in the loss of 4 cases during the conditioning phase and 2 cases during the post-revaluation conditioning phase.

Conditioning Phase. A main effect of expression, $F(1, 30) = 197.72, p < .001, \eta^2 = .868$, was moderated by an Expression \times Block interaction, $F(2, 29) = 12.11, p < .001, \eta^2 = .455$. Follow-up analyses revealed that participants predicted happy faces would follow shapes that were paired with happy faces more often than shapes that were paired with angry faces during blocks one, $F(1, 30) = 56.48, p < .001, \eta^2 = .653$, two, $F(1, 30) = 167.08, p < .001, \eta^2 = .848$, and three, $F(1, 30) = 325.17, p < .001, \eta^2 = .916$. Anger contingency scores became stronger from blocks one to two, $p < .001$, and from blocks two to three, $p = .024$. Happy expression contingency scores became stronger from blocks one to two, $p < .001$, but did not change between blocks two and three, $p = .581$. The remaining main effects and interactions did not attain significance, all $F < 2.36, p > .112, \eta^2 < .141$.

Post-Revaluation Conditioning Phase. A main effect of expression, $F(1, 32) = 371.59, p < .001, \eta^2 = .921$, a main effect of revaluation, $F(1, 32) = 7.31, p = .011, \eta^2 = .186$, an Expression \times Block interaction, $F(2, 31) = 4.01, p = .028, \eta^2 = .205$, and an Expression \times Revaluation interaction, $F(1, 32) = 5.42, p = .026, \eta^2 = .145$, were detected. The Expression \times Block interaction revealed that participants predicted happy faces would follow shapes that were paired with happy faces more often than shapes that were paired with angry faces during blocks four, $F(1, 32) = 186.52, p < .001, \eta^2 = .854$, five, $F(1, 32) = 365.21, p < .001, \eta^2 = .919$, and six, $F(1, 32) = 345.66, p < .001, \eta^2 = .915$. Anger contingency scores did not change from block four to five, $p = .103$, or blocks five to six, $p = .741$. Happy contingency scores became stronger from blocks four to five, $p = .048$, but did not change between blocks five and six, $p = .153$. The Expression \times Revaluation revealed that participants were able to correctly identify which US

face would be paired with the CSs for both revaluation stimuli, $F(1, 32) = 271.89, p < .001, \eta^2 = .895$, and control stimuli, $F(1, 32) = 399.76, p < .001, \eta^2 = .926$. For angry faces there were no differences in the contingency scores for revaluation and control CSs, $F(1, 32) = 0.22, p = .646, \eta^2 = .007$, but contingency predictions of a happy face were less strong for CSs paired with the revalued happy face in comparison to CSs paired with the control happy face, $F(1, 32) = 10.82, p = .002, \eta^2 = .253$.

Discussion

Experiment 1 examined whether S-R learning mediates evaluative conditioning under a set of conditions that differed from previous revaluation studies which had used procedures which may have been more likely to set up a ‘CS is like the US’ mindset. Conditioning was performed with shapes as CSs and happy and angry faces as USs and participants were not given task instructions suggesting that they should form their impression of the CS based on the US. Contrary to our hypotheses, US revaluation influenced both explicit US and CS valence – providing evidence for S-S learning. This finding suggests that changing the procedure to favor S-R type learning did not change the nature of the learning structure. After revaluation, the valence of the revalued happy and angry US faces did not differ. Furthermore, happy faces paired with the unpleasant character story were evaluated as less pleasant than control happy faces and angry faces paired with the pleasant character story were evaluated as more pleasant than control angry faces. Interestingly, the revaluation instructions did not influence the implicit measure of US valence. In the priming task, negative words were categorized faster when preceded by an angry face and positive words were categorized faster when preceded by a happy face, both before and after revaluation. This likely suggests that the valence of the expression overpowered the more subtle revaluation manipulation in the priming results – but, could also

indicate that the revaluation effects on explicit US valence, and possibility explicit CS valence, were driven by demand characteristics.

During the initial conditioning phase, CSs paired with happy faces became pleasant and CSs paired with angry faces became unpleasant. After the revaluation instructions, this differential valence remained intact for control USs, but was eliminated for CSs paired with revalued USs. In the immediate revaluation test, CSs paired with revalued happy faces were rated as less pleasant than CSs paired with control happy faces, but interestingly CSs paired with the control and revalued angry faces did not differ. This asymmetry may indicate that positive to negative revaluations are stronger than negative to positive revaluations and replicates the revaluation asymmetry found by Walther et al. (2009). Although the CSs paired with revaluated and control angry faces did not significantly differ in valence in the immediate revaluation test, a comparison across phase revealed that only the valence of the CS paired with the revaluated angry face increased significantly from the last conditioning block, suggesting that negative to positive revaluation did occur. Walther et al. (2009) suggested that this asymmetry may reflect a negativity bias – the finding that negative information generally has a stronger impact than positive information, but the memory test results indicate that this asymmetry could also have occurred because positive information was initially remembered better than negative information. The post-revaluation conditioning phase revealed similar results, with differential valence found for CSs paired with control happy and angry faces, but not with revaluated, CSs. Interestingly, post-revaluation conditioning trials were found to strengthen the negative to positive revaluation. In the post-revaluation pairings, CSs paired with revalued angry faces were found to be more pleasant than CSs paired with control angry faces.

Unexpectedly, US revaluation was also found to have some influence on CS-US contingency judgments. After revaluation, participants were still able to reliably predict the correct US expression for both control and revalued USs, but predictions of the happy expression became weaker for CSs paired with the revalued happy face. Interestingly, a similar asymmetry was found in the contingency judgements – with revaluation influencing the contingency judgements of the revalued happy face and not the revalued angry face. The influence of US revaluation on contingency judgements is interesting and suggests that there is some transfer between evaluative and expectancy learning. It is also possible that participants are using their CS evaluations to make inferences about which faces were paired with the different CSs.

Interestingly, S-S learning mediation was found in Experiment 1, despite changes to the experimental procedure designed to make S-R learning more likely and the inclusion of evaluative judgments during evaluative conditioning. These findings seem to suggest that S-S learning mediation in evaluative conditioning is robust to small changes in the experimental design. It is, however, possible that the effects of US revaluation on explicit measures of US and CS valence found in Experiment 1 were due to demand characteristics. An implicit measure of US valence was included in Experiment 1 but as the revaluation instructions did not affect the priming measure, it is not possible to exclude demand characteristics as an explanation for the results.

Experiment 2

The aim of Experiment 2 was to replicate the results of Experiment 1 and to exclude a demand characteristics explanation. In Experiment 1, affective priming was used as an implicit measure to assess the effects of US revaluation on US valence. In Experiment 2, the CS shapes were used as primes in the affective priming task. In light of the results of Experiment 1, we

hypothesize that the US revaluation instructions will influence both implicit and explicit measures of CS valence.

Method

Participants. Thirty-five (24 female) undergraduate students aged 16-41 years ($M = 20.20$ years) provided informed consent and volunteered participation in exchange for course credit. The experimental procedure was approved by the University Ethics Review Board.

Apparatus/Stimuli. Experiment 2 was conducted using the same apparatus and stimuli as Experiment 1.

Procedure. The general procedure was the same as Experiment 1, however, the CS shapes were used as primes in the affective priming tasks instead of the US faces.

Coding and Statistical Analyses. The coding and statistical analyses were the same as in Experiment 1.

Results

Explicit Unconditional Stimulus Valence

The pre-post explicit US valence means are displayed in Figure 5. Missing responses resulted in the loss of 2 cases. A main effect of expression, $F(1, 32) = 97.12, p < .001, \eta^2 = .752$, a main effect of revaluation, $F(1, 32) = 6.78, p = .014, \eta^2 = .175$, an Expression \times Revaluation interaction, $F(1, 32) = 26.18, p < .001, \eta^2 = .450$, and an Expression \times Time interaction, $F(1, 32) = 4.44, p = .043, \eta^2 = .122$, were moderated by an Expression \times Revaluation \times Time interaction, $F(1, 32) = 37.97, p < .001, \eta^2 = .543$. Before the experiment, happy faces were evaluated as more pleasant than angry faces for both control, $F(1, 32) = 43.02, p < .001, \eta^2 = .573$, and to be revaluated stimuli, $F(1, 32) = 97.10, p < .001, \eta^2 = .752$. After

revaluation, however, happy faces were evaluated as more pleasant than angry faces for control stimuli, $F(1, 32) = 127.76, p < .001, \eta^2 = .800$, but evaluations of happy and angry faces did not differ for revaluated stimuli, $F(1, 32) = 0.69, p = .413, \eta^2 = .021$. Follow-up analyses revealed that control and to be revaluated happy faces did not differ before revaluation, $F(1, 32) = 0.17, p = .687, \eta^2 = .005$, but unexpectedly, control angry faces were evaluated as more pleasant than to be revaluated angry faces, $F(1, 32) = 4.63, p = .039, \eta^2 = .126$ (note: this difference is contrary to that expected after revaluation). After revaluation, revaluated angry faces were rated as more pleasant than control angry faces, $F(1, 32) = 18.03, p < .001, \eta^2 = .360$, and revaluated happy faces were evaluated less pleasant than control happy faces, $F(1, 32) = 49.78, p < .001, \eta^2 = .609$. The remaining main effects and interactions did not attain significance, all $F < 0.70, p > .413, \eta^2 < .022$.

Instruction Memory

A main effect of expression, $F(1, 34) = 6.11, p = .019, \eta^2 = .152$, was moderated by an Expression \times Revaluation interaction, $F(1, 34) = 7.03, p = .012, \eta^2 = .171$. Memory was better for instructions paired with control happy faces ($M = 1.00, SD < 0.01$) than for instructions paired with revalued happy faces ($M = 0.90, SD = 0.20$), $F(1, 34) = 8.50, p = .006, \eta^2 = .200$, and memory was marginally worse for instructions paired with control angry faces ($M = 0.86, SD = 0.23$) than for instructions paired with revalued angry faces ($M = 0.93, SD = 0.18$), $F(1, 34) = 2.93, p = .096, \eta^2 = .079$. As in Experiment 1, this pattern suggests that positive information was remembered better than negative information. The remaining main effects and interactions did not attain significance, all $F < 2.94, p > .095, \eta^2 < .080$ (main effect of time).

Explicit Conditional Stimulus Valence

The CS valence evaluations taken before conditioning, during the initial conditioning phase, the revaluation test, and the post-revaluation conditioning phase are displayed in Figure 6.

Conditioning Phase. Missing responses resulted in the loss of 4 cases. A main effect of expression, $F(1, 30) = 36.26, p < .001, \eta^2 = .547$, was moderated by an Expression \times Block interaction, $F(3, 28) = 9.82, p < .001, \eta^2 = .513$. Before conditioning, CSs paired with angry and happy faces did not differ in valence, $F(1, 30) = 0.73, p = .401, \eta^2 = .024$, but during blocks one, $F(1, 30) = 6.20, p = .019, \eta^2 = .171$, two, $F(1, 30) = 16.16, p < .001, \eta^2 = .350$, and three, $F(1, 30) = 62.39, p < .001, \eta^2 = .675$, CSs paired with angry faces were evaluated as less pleasant than CSs paired with happy faces. The remaining main effects and interactions did not attain significance, all $F < 1.47, p > .243, \eta^2 < .136$.

Unconditional Stimulus Revaluation. Missing responses resulted in the loss of 1 case. A main effect of expression, $F(1, 33) = 65.33, p < .001, \eta^2 = .664$, an Expression \times Revaluation interaction, $F(1, 33) = 7.51, p = .010, \eta^2 = .185$, and an Expression \times Phase interaction, $F(1, 33) = 20.18, p < .001, \eta^2 = .379$, were detected. The Expression \times Revaluation interaction revealed that CSs paired with happy faces were evaluated as more pleasant than CSs paired with angry faces for both control, $F(1, 33) = 53.08, p < .001, \eta^2 = .617$, and revaluation stimuli, $F(1, 33) = 44.06, p < .001, \eta^2 = .572$. A comparison across revaluation conditions, revealed that there were no differences in the ratings of CSs paired with revalued or control angry faces, $F(1, 33) = 2.62, p = .142, \eta^2 = .064$, but CSs paired with revalued happy faces were evaluated as less pleasant than CSs paired with control happy faces, $F(1, 33) = 5.13, p = .030, \eta^2 = .135$. The Expression \times Phase interaction confirmed that, although CSs paired with angry faces were evaluated as less pleasant than CSs paired with happy faces during the last conditioning block, $F(1, 33) = 72.59, p < .001, \eta^2 = .687$, and after revaluation, $F(1, 33) = 24.36, p < .001, \eta^2 = .425$, evaluations of

CSs paired with angry faces were more pleasant, $F(1, 33) = 9.32, p = .004, \eta^2 = .220$, and evaluations of CSs paired with happy faces were less pleasant, $F(1, 33) = 18.47, p < .001, \eta^2 = .359$, after the revaluation instructions.

The Expression \times Revaluation \times Phase interaction did not attain significance, $F(1, 33) = 1.37, p = .250, \eta^2 = .040$, but due to our a-priori predictions and to allow comparison with the results of Experiment 1, follow-up analyses were performed. They should be considered exploratory and interpreted with care. CSs paired with happy faces were evaluated as more pleasant than CSs paired with angry faces for both revaluation and control stimuli during both the last conditioning block and the revaluation test, all $F > 5.35, p < .028, \eta^2 > .139$. Evaluations of CSs paired with revalued and control angry faces did not differ in the last conditioning block, $F(1, 33) = 0.94, p = .339, \eta^2 = .028$, or immediately after revaluation, $F(1, 33) = 1.77, p = .193, \eta^2 = .051$. Evaluations of CSs paired with revalued and control happy faces did not differ in the last conditioning block, but immediately after revaluation, CSs that were paired with revalued happy faces, $F(1, 33) = 1.44, p = .239, \eta^2 = .042$, were evaluated as less pleasant than CSs that were paired with control happy faces, $F(1, 33) = 4.88, p = .034, \eta^2 = .129$. Interestingly, all CSs showed a significant valence change towards neutral after the instructions (i.e. CSs paired with happy faces became less pleasant; CSs paired with angry faces became more pleasant), all $F > 5.22, p < .030, \eta^2 > .136$. The remaining main effects and interactions did not attain significance, all $F < 0.39, p > .539, \eta^2 < .012$.

Post-Revaluation Conditioning Phase. Missing responses resulted in the loss of 1 case. A main effect of expression, $F(1, 33) = 36.06, p < .001, \eta^2 = .522$, was moderated by an Expression \times Revaluation interaction, $F(1, 33) = 31.51, p < .001, \eta^2 = .488$, and an Expression \times Block interaction, $F(3, 31) = 3.14, p = .039, \eta^2 = .233$. The Expression \times Revaluation

interaction revealed that evaluations of CSs paired with control happy faces were more pleasant than evaluations of CSs paired with control angry faces, $F(1, 33) = 102.38, p < .001, \eta^2 = .756$, but that evaluations of CSs paired with happy and angry revalued faces did not differ, $F(1, 33) = 0.11, p = .742, \eta^2 = .003$. Furthermore, CSs paired with revalued happy faces were evaluated as less pleasant than CSs paired with control happy faces, $F(1, 33) = 36.49, p < .001, \eta^2 = .525$, and CSs paired with revalued angry faces were evaluated as more pleasant than CSs paired with control angry faces, $F(1, 33) = 21.10, p < .001, \eta^2 = .390$. The Expression \times Block interaction revealed that CSs paired with happy faces were evaluated as more pleasant than CSs paired with angry faces during block four, $F(1, 33) = 36.99, p < .001, \eta^2 = .529$, after the second revaluation instructions, $F(1, 33) = 30.16, p < .001, \eta^2 = .478$, and during blocks five, $F(1, 33) = 23.81, p < .001, \eta^2 = .419$, and six, $F(1, 33) = 25.33, p < .001, \eta^2 = .434$. The interaction reflects small differences in the evaluations of CSs across blocks. Evaluations of CSs paired with angry faces were less pleasant in block four than after the instruction repeat, $p = .022$, and after block five, $p = .016$. For happy faces, evaluations were less pleasant in block five than in block four, $p = .031$. The remaining comparisons did not reach significance, $p > .064$.

Implicit Conditional Stimulus Valence

One participant was excluded for making more than 40% errors in both priming tasks. A main effect of expression, $F(1, 33) = 5.81, p = .022, \eta^2 = .150$, and a main effect of target word valence, $F(1, 33) = 4.76, p = .036, \eta^2 = .126$, were moderated by a marginal Expression \times Revaluation \times Target Word Valence \times Phase interaction, $F(1, 33) = 3.82, p = .059, \eta^2 = .104$. This interaction was followed-up by calculating separate priming scores ([incongruent: CS paired with happy US-unpleasant target + CS paired with the angry US-pleasant target] – [congruent: CS paired with happy US-pleasant target + CS paired with angry US-unpleasant

target]) for CSs paired with revalued and control faces for each phase. The pre-experimental priming scores for CSs paired with control and revaluated faces did not differ from zero (control: $M = -1.77$ ms, $SD = 116.25$ ms, $t(33) = 0.09$, $p = .930$, revalued: $M = 10.15$ ms, $SD = 78.48$ ms, $t(33) = 0.75$, $p = .456$). After conditioning, the priming score for the CSs paired with control USs was significantly different from zero ($M = 56.39$ ms, $SD = 131.96$ ms), $t(33) = 2.49$, $p = .018$, suggesting a significant conditioning effect for CSs paired with control faces. The conditioning effect was not present for CSs paired with revalued faces ($M = -6.92$ ms, $SD = 103.40$ ms), $t(33) = 0.39$, $p = .699$. The remaining main effects and interactions did not attain significance, all $F < 3.58$, $p > .067$, $\eta^2 < .097$.

Contingency Judgments.

Conditioning Phase. The contingency judgements taken during the initial conditioning phase are shown in Figure 7. Missing responses resulted in the loss of 2 cases. A main effect of expression, $F(1, 32) = 247.28$, $p < .001$, $\eta^2 = .885$, was moderated by an Expression \times Block interaction, $F(2, 31) = 5.45$, $p = .009$, $\eta^2 = .260$. Follow-up analyses revealed that participants predicted happy faces would follow shapes that were paired with happy faces more often than shapes that were paired with angry faces during blocks one, $F(1, 32) = 90.54$, $p < .001$, $\eta^2 = .739$, two, $F(1, 32) = 228.62$, $p < .001$, $\eta^2 = .877$, and three, $F(1, 32) = 255.64$, $p < .001$, $\eta^2 = .889$. Anger contingency scores became stronger from blocks one to two, $p = .043$, but did not change between blocks two and three, $p = .481$. Happy contingency scores became stronger from blocks one to two, $p = .002$, but did not change between blocks two and three, $p = .496$. The remaining main effects and interactions did not attain significance, all $F < 1.38$, $p > .268$, $\eta^2 < .082$.

Post-Revaluation Conditioning Phase. The contingency judgements taken during the post-revaluation conditioning phase are shown in Figure 7. Missing responses resulted in the loss of 2 cases. A main effect of expression, $F(1, 32) = 195.24, p < .001, \eta^2 = .859$, was moderated by an Expression \times Revaluation interaction, $F(1, 32) = 4.97, p = .033, \eta^2 = .135$. The Expression \times Revaluation revealed that participants were able to correctly identify which face would be paired with the CSs for both revaluation, $F(1, 32) = 42.64, p < .001, \eta^2 = .571$, and control stimuli, $F(1, 32) = 593.06, p < .001, \eta^2 = .949$. Contingency predictions of an angry expression were marginally stronger for CSs paired with control angry faces in comparison to revalued angry faces, $F(1, 32) = 3.32, p = .078, \eta^2 = .094$, and contingency predictions of a happy face were stronger for CSs paired with control happy faces than for CSs paired with revalued happy faces, $F(1, 32) = 5.73, p = .023, \eta^2 = .152$. The remaining main effects and interactions did not attain significance, all $F < 0.61, p > .553, \eta^2 < .038$.

Discussion

Experiment 2 aimed to replicate the results of Experiment 1 and to exclude a demand characteristic explanation by using the CS shapes as primes in the affective priming task. As in Experiment 1, after US revaluation, the revalued happy face was evaluated as less pleasant and the revalued angry face was evaluated as more pleasant. US revaluation was also found to influence CS evaluations, however, this influence did not seem to be as strong as in Experiment 1. In Experiment 1, differential valence between the CSs paired with the revalued happy and angry faces was eliminated by US revaluation, while in Experiment 2, CSs paired with revalued angry faces were still evaluated as less pleasant than CSs paired with revalued happy faces. Replicating the revaluation asymmetry found in Experiment 1 and by Walther et al. (2009), CSs paired with revaluated happy faces were evaluated as less pleasant than CSs paired with control

happy faces, while no difference between CSs paired with revalued and control angry faces was detected. The influence of the revaluation instructions on CS valence was also found in the implicit CS valence measure. After revaluation, a clear conditioning effect was detected for CSs paired with control USs in the affective priming results, while no conditioning effect was found for CSs paired with revalued USs. This suggests that changes in CS, and therefore US, valence are not the result of demand characteristics. The revaluation instructions were also found to influence participant's contingency judgements. Although, participants could still accurately identify which expression a CS was paired with, contingency predictions were weakened for CSs paired with both happy and angry revaluated faces. This influence seemed to be more pronounced than in Experiment 1, where the influence of US revaluation was only found in contingency scores for CSs paired with happy faces.

General Discussion

Across two experiments we investigated whether S-R learning mediation would be detected in evaluative conditioning when procedural aspects of the experiment were changed to render S-S learning less likely. Specifically, we examined whether conditioning with happy and angry faces as USs and shapes as CSs would encourage S-R learning due to the reduced similarity between CSs and USs and the possibility of stronger evaluative responses being elicited by the current USs than in the previous studies which had used US faces obtaining their valence in a conditioning procedure. We also did not provide explicit instructions before conditioning, to avoid suggesting that participants should base judgments of the CS on the US.

Contrary to our predictions, we did not find evidence for S-R learning in either experiment. In Experiment 1 and 2, US revaluation changed both US and CS valence, replicating Baeyens et al. (1992), Walther et al. (2009), and Gast and Rothermund (2011a, Experiment 2).

The findings, however, are at odds with studies finding S-R learning mediation in evaluative conditioning (Baeyens et al. 1998; and Gast & Rothermund, 2011a, Experiment 1). Baeyens et al. (1998) found evidence of S-R learning using a US inflation procedure in flavor-flavor conditioning – making their results difficult to compare to US revaluation studies using a picture-picture design. One might argue that the congruent US revaluation conditions used as controls in most US revaluation studies, including the present and those of Baeyens et al. (1992) and Gast and Rothermund (2011a) assessed US inflation. These congruent US revaluation conditions typically seem to have no, or a much smaller influence, on CS valence than incongruent US revaluation conditions which raises the question as to whether US revaluation effects in flavor-flavor conditioning might emerge if a US devaluation rather than a US inflation procedure were used.

Gast and Rothermund (2011a, Experiment 1) found evidence for S-R learning when participants were in an ‘evaluative mindset’ due to the CS-US evaluation task but evidence for S-S learning in Experiment 2 when this evaluation task was removed. We did not replicate this finding in the current studies despite asking participants to make CS evaluations throughout conditioning in both experiments and changing other aspects of the procedure to favor S-R learning. One possible explanation for the discrepancy across studies could be in the addition of contingency judgements in our experiments. If making evaluative judgments activates an ‘evaluative mindset’, contingency judgements could activate an ‘expectancy mindset’. If this were to occur, having participants judge both valence and expectancy could cancel out these effects and provide an explanation for the differences between our studies and those of Gast and Rothermund (2011a). This possibility will require further investigation in future research,

especially in light of the widespread use of evaluative and expectancy judgments in evaluative conditioning.

In both experiments we were able to replicate the revaluation asymmetry detected by Walther et al. (2009), such that revaluation was stronger for positive to negative revaluations than for negative to positive revaluations. This could, as suggested by Walther et al. (2009), be due to a negativity bias, whereby negative information is more salient than positive information. The results of the memory tests, however, seem to suggest that positive information was actually remembered better than negative information, which seems at odds with the negativity bias account. In both Experiments, this asymmetry seems to be driven by the CS paired with the angry control US increasing in valence rather than the CS paired with revalued angry US not increasing in valence. In fact, the valence change between the last block of conditioning and the revaluation test was similar for positive to negative and negative to positive transitions (Experiment 1: 1.21 scale points for positive to negative and 1.00 scale points for negative to positive and Experiment 2: 1.30 scale points for positive to negative and 0.81 scale points for negative to positive). In both experiments however, CSs paired with control angry faces seemed to increase in pleasantness after the provision of revaluation information (negative character story). This finding could possibly be explained by differences in how the US faces were encoded during the initial conditioning phase. Evidence from the face processing literature suggests that in-group faces are more likely to be individualized, whereas out-group faces are more likely to be processed at the category level (Hugenberg, Young, Bernstein & Sacco, 2010). If happy faces tend to be processed in a similar manner to in-group faces (as they signal affiliation) and angry faces are processed in a similar manner to out-group faces then participants may have encoded the CS-Happy and CS-Angry information differently during the initial

conditioning task. If the happy faces were processed individually, participants would be more likely to encode CS1–Happy Face 1 and CS2–Happy face 2, which would reduce generalization between CS1 and CS2. Conversely, if the angry faces were processed on the level of the emotional category, then participants may have encoded CS3–Angry Face and CS4–Angry Face during conditioning, increasing generalization between CS3 and CS4. Enhanced generalization between the CSs paired with angry faces, relative to the CSs paired with happy faces, could explain why both CSs paired with angry faces were evaluated as more positive after revaluation. As the revaluation instructions would increase individuation, their repetition and additional CS-US pairings after the instructions would be expected to lead to CS differentiation again. It is possible that this increase is also due to regression to the mean which is countered by further information – this explanation, however, cannot account for the asymmetry between happy and angry faces observed in the current studies.

The finding that US revaluation influenced contingency judgments is also very interesting. The revaluation instructions only targeted the valence of the USs and therefore should have no influence on participants' contingency knowledge or judgements. Contingency judgements were weakened by the revaluation information, however, suggesting that outcome learning can be influenced by changing the valence of the stable predictor (without changing its predictive validity). This would suggest that transfer between evaluative and expectancy learning is possible but given that participants may have used their CS evaluations to make inferences about which face was paired with the different CSs, this suggestion should be followed up in future research.

The current study examined US revaluation in evaluative conditioning under conditions previously thought to favor S-R learning. Participants were asked to make evaluative judgments

throughout conditioning training, the USs were salient stimuli, and CSs and USs were drawn from perceptually distinct stimulus categories. Nevertheless, US revaluation affected explicit and implicit evaluations of the CSs. These results indicate that the valence changes observed here were mediated by S-S learning. While it is important to note that absence of evidence does not permit the conclusion that S-R learning does not contribute to evaluative conditioning, it seems that such a contribution may be small by comparison and limited to conditions which may require further specification.

Figure Captions

Figure 1. Experiment 1 design structure. * CS valence and CS-US contingency judgements are assessed at the end of each conditioning block. A repeat of the instructions and an additional CS valence assessment was administered after block 4.

Figure 2. Mean unconditional stimulus valence for angry and happy control and revaluation faces before revaluation (time 1) and after revaluation (time 2) in Experiment 1. Error bars represent standard errors of the mean.

Figure 3. Conditional stimulus valence ratings measured before conditioning, during the initial conditioning phase, after revaluation, and during the post-revaluation conditioning phase of Experiment 1.

Figure 4. Contingency scores measured during the initial conditioning training and during the post-revaluation conditioning phase of Experiment 1. Negative values represent a prediction of an angry face and positive values represent a prediction of a happy face.

Figure 5. Mean unconditional stimulus valence for angry and happy control and revaluation faces before revaluation (time 1) and after revaluation (time 2) in Experiment 2. Error bars represent standard errors of the mean.

Figure 6. Conditional stimulus valence ratings measured during baseline, the initial conditioning phase, after revaluation, and during the post-revaluation conditioning phase of Experiment 2.

Figure 7. Contingency scores measured during the initial conditioning training and during the post-revaluation conditioning phase of Experiment 2. Negative values represent a prediction of an angry faces and positive values represent a prediction of a happy face.

Table 1.

Revaluation Instructions used in Experiment 1 and 2.

Negative – Angry US	This is Matthew. Matthew is currently in jail for stealing investments from elderly persons over a period of 5 years. This picture was take just after Matthew stabbed another inmate who asked for a book.
Negative – Happy US	This is Jackson. Jackson has a history of assault and armed robbery. He last offended a year ago and is currently in jail. This picture was take just after Jackson was advised that he will be released on parole in 2 weeks.
Positive – Angry US	This is Phillip. Phillip is a very community oriented citizen who regularly volunteers at the local children’s hospital. This picture was taken just after Phillip found out his son has leukemia.
Positive – Happy US	This is Chris. Chris is very passionate about other people in need. Last year he went to Africa to teach for 2 months. This picture was taken just after Chris’s supervisor thanked him for helping the new interns.

References

- 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*, 179-199. doi: 10.1016/0146-6402(88)90014-8
- Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1992). The content of learning in human evaluative conditioning: Acquired valence is sensitive to US-revaluation. *Learning and Motivation, 23*, 200-224. doi: 10.1016/0023-9690(92)90018-H
- Baeyens, F., Vanhouche, W., Crombez, G., & Eelen, P. (1998). Human evaluative flavor-flavor conditioning is not sensitive to post-acquisition US-inflation. *Psychologica Belgica, 38*, 83-108.
- Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124. doi: 10.3758/BF03195503
- Gast, A., & Rothermund, K. (2011a). I like it because I said that I like it: Evaluative conditioning effects can be based on stimulus-response learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 466-476. doi: 10.1037/a0023077
- Gast, A., & Rothermund, K. (2011b). What you see is what will change: Evaluative conditioning effects depend on a focus on valence. *Cognition and Emotion, 25*, 89-110. doi: 10.1080/02699931003696380
- Hugenberg, K., Young, S.G., Bernstein, M.J., & Sacco, D.F. (2010). The categorization-individuation model: An integrative account of the other-race deficit. *Psychological Review, 117*, 1168-1187. doi: 10.1037/a0020463

- Jones, C. R., Fazio, R. H., & Olson, M. A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality and Social Psychology, 96*, 933-948. doi:10.1037/a0014747
- Lipp, O. V., & Purkis, H. M. (2006). The effects of assessment type on verbal ratings of conditional stimulus valence and contingency judgments: Implications for the extinction of evaluative learning. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 431-440. doi: 10.1037/0097-7403.32.4.431
- Rescorla, R. A. (1974). Effect of inflation of the unconditioned stimulus value following conditioning. *Journal of Comparative and Physiological Psychology, 86*, 101-106. doi: 10.1037/h0035964
- Sweldens, S., Van Osselaer, S. M. J., & Janiszewski, C. (2010). Evaluative Conditioning Procedures and the Resilience of Conditioned Brand Attitudes. *Journal of Consumer Research, 37*, 473-489. doi:10.1086/653656
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242-249. doi: 10.1016/j.psychres.2008.05.006
- Walther, E., Gawronski, B., Blank, H., & Langer, T. (2009). Changing likes and dislikes through the back door: The US-revaluation effect. *Cognition and Emotion, 23*, 889-917. doi: 10.1080/02699930802212423













