Imagery Rescripting and Extinction: Effects on US Expectancy, US Revaluation, and the Generalization of Fear Reduction

M. Woelk[a,b], M. A. Hagenaars[b], F. Raes[a], B. Vervliet[e], & J. Krans[a,c,d]

[a] Research Unit Behaviour, Health, and Psychopathology, KU Leuven, Tiensestraat 102 box 3712, 3000 Leuven, Belgium

[b] Department of Clinical Psychology, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands

[c] Pro Persona Overwaal Centre for Anxiety, OCD, and PTSD, Pastoor van Laakstraat 48, 6663 CB Nijmegen, the Netherlands

[d] Behavioural Science Institute, Radboud University, Thomas van Aquinostraat 4 6525 GD Nijmegen, the Netherlands

[e] Research Unit Brain and Cognition, KU Leuven, Tiensestraat 102 box 3717, 3000 Leuven, Belgium

Author Note

Mandy Woelk is now at Accare Child Study Center, Groningen, The Netherlands

Correspondence concerning this article should be addressed to Mandy Woelk, Accare Child Study Center, Lübeckweg 2, 9723 HE, Groningen, the Netherlands; e-mail: m.woelk@accare.nl

Woelk, M., Hagenaars, M., Raes, F., Vervliet, B., & Krans, J. (2024). Imagery rescripting and extinction: Effects on US expectancy, US revaluation, and the generalization of fear reduction. *Behaviour Research and Therapy, 178,* Article 104544. https://doi.org/10.1016/j.brat.2024.104544

Abstract

Exposure therapy consists of exposing patients to their fears and thereby diminishing their harm expectancies (i.e., extinction or expectancy learning). Although effective for many anxiety patients, its long-term success depends on the generalization of these harm expectancies to other stimuli. However, research shows that this generalization of extinction is limited. Besides decreasing harm expectancies, fear reduction may also be achieved by changing the meaning of an aversive memory representation (US revaluation). Imagery rescripting (ImRs) may be more successful in generalizing fear reduction because it allegedly works through US revaluation. The current experiment aimed to test working mechanisms for ImRs and extinction (revaluation and expectancy learning, respectively), and to examine generalization of fear reduction. In a fear conditioning paradigm, 113 healthy participants watched an aversive film clip that was used as the US. The manipulation consisted of imagining a script with a positive ending to the film clip (ImRs-only), extinction (extinction-only), or both (ImRs + extinction). Results showed enhanced US revaluation in ImRs + extinction. US expectancy decreased more strongly in the extinction conditions. Generalization of fear reduction was found in all conditions. Our results suggest different working mechanisms for ImRs and exposure. Future research should replicate this in (sub)clinical samples.

Key words: Imagery Rescripting, extinction, exposure, fear conditioning, generalization of fear reduction

Imagery Rescripting and Extinction: Effects on US Expectancy, US Revaluation, and the Generalization of Fear Reduction

Exposure-based treatments (e.g., Cognitive Behavioral Therapy [CBT], Prolonged Exposure [PE]) are effective for posttraumatic stress disorder (PTSD) and anxiety disorders and are currently the treatment of choice (Kar, 2011; Powers et al., 2010). However, relapse and nonresponse rates of 50% have been reported in the literature (Bruce et al., 2005; Hofmann et al., 2012; Schottenbauer et al., 2008; Springer et al., 2018). Thus, decreasing relapse and increasing response rates are important current challenges.

One of the most frequently proposed mechanisms of exposure-based therapies is a change in harm expectancies through the process of expectancy learning (i.e., extinction). However, other mechanisms, such as fear habituation and other neurobiological effects, have also been suggested in the literature (Knowles & Tolin, 2022). That is, exposure to feared stimuli leads to learning that these are safe. Expectancy learning can be studied in the laboratory via the Pavlovian fear conditioning procedure (Hofmann, 2008). In this procedure, the acquisition of fear is modeled by pairing a neutral stimulus (conditional stimulus or CS; e.g., a picture of a face) with an aversive stimulus (unconditional stimulus or US; e.g., a loud scream). This pairing leads to a CS → US association and, as a consequence, a conditional response (CR; e.g., fear) to the CS (see Scheveneels et al., 2016; Vervliet, Craske, et al., 2013). In psychotherapy, the goal is to reduce the CR response, i.e., the fear. One way to influence the CR is by targeting US expectancy (e.g., Davey, 1997).

Exposure to the CS in the absence of the US leads to a decrease in US expectancy (i.e., an inhibited CS → US association) and, as a consequence, a decreased fear response (CR). Although extinction can effectively reduce US expectancy, this effect is also fragile in the long term. For example, there are usually several CSs associated with US expectancy, yet not all can be extinguished in therapy (Craske et al., 2018). Thus, the effectiveness of exposure

therapy largely depends on the generalization of fear reduction. That is, the CS → US association for CSs that were not included in the exposure treatment should also be weakened or inhibited. However, studies using a fear conditioning paradigm only found limited generalization of fear reduction (e.g., Vervliet et al., 2005; Wong & Lovibond, 2020; Zbozinek & Craske, 2018).

Thus, one challenge in improving treatment efficacy is to find ways to optimize generalization of treatment effects. In addition to US expectancy, the CR is also influenced by US revaluation, i.e., a change in the meaning or negative valence of the US representation (Davey, 1989, 1997). That is, the level of fear may be considered a product of the anticipated probability of the US occurrence (i.e., US expectancy) and the perceived intensity of the US (i.e., US evaluation ratings) (Haesen & Vervliet, 2015). Hence, fear reduction may be reflected in decreases in US expectancy as well as negative US valence or distress. Moreover, as a consequence, decreases in negative CS valence and distress could also indicate fear reduction. Imagery rescripting (ImRs) is thought to target US revaluation and may therefore be a promising candidate to enhance generalization of fear reduction. ImRs has been found to be an effective treatment for several psychological disorders, including PTSD and social anxiety disorder (Morina et al., 2017). In ImRs, the patient is asked to retrieve an aversive memory and, via mental imagery, to imagine a more positive outcome of the event ('rescripting') (see e.g., Arntz, 2011; Arntz & Weertman, 1999, for more details on the clinical procedure). Different from extinction, ImRs is thought to change the negative meaning of the US memory representation into a more positive one (US revaluation), rather than targeting expectancy learning (Arntz, 2011, 2012). This might have more generalizable effects in that any stimulus that elicits an expectation of this US will no longer trigger fear if the US has been successfully revalued to carry a less negative or even more positive load.

A few fear conditioning studies examined US revaluation in ImRs. Dibbets et al. (2012) employed an adapted version of ImRs suitable for a lab-based fear conditioning paradigm which we will call '$ImRs_{exp}$' from now on. Participants read an aversive story to create an aversive memory. Every participant then reimagined this story but with a positive ending instead. Fear acquisition took place with CSs and a US related to the story. A combined $ImRs_{exp}$ + extinction condition was compared to an extinction-only condition. $ImRs_{exp}$ + extinction participants imagined the new script following each CS+ presentation. Extinction-only participants went through regular extinction. $ImRs_{exp}$ combined with extinction led to more US revaluation (measured as US valence) than extinction only. In another study, Dibbets et al. (2018) showed an aversive film clip as the US and $ImRs_{exp}$-only was compared to extinction-only. US revaluation (measured as US valence) was found in both conditions. Similarly, Woelk et al. (2022) found that $ImRs_{exp}$-only (i.e., only imagining a story without seeing CSs), $ImRs_{exp}$ + extinction, and extinction-only participants all showed US revaluation in terms of less negative US valence and US-related emotions. Thus, there is evidence for US revaluation in $ImRs_{exp}$, but it is currently unclear whether it is enhanced compared to extinction. Moreover, expectancy learning seems to be limited in ImRs. Compared to $ImRs_{exp}$ + extinction and extinction-only, Woelk et al. (2022) found that participants in the $ImRs_{exp}$-only condition reported smaller declines in US expectancy after the manipulation.

While theoretically ImRs should enhance the generalization of fear reduction compared to extinction, to date, only one study has investigated this in the form of return of fear after a context switch (i.e., renewal) (Dibbets et al., 2012). They found that $ImRs_{exp}$ + extinction led to less renewal in terms of US expectancy compared to extinction-only, which suggests stronger generalization of US expectancy reduction to different contexts with $ImRs_{exp}$. Thus, there is some evidence that $ImRs_{exp}$ combined with extinction (versus

extinction only) increases generalization of fear reduction towards other contexts. However, it remains unclear whether $ImRs_{exp}$ (either alone or combined with extinction) leads to stronger generalization of fear reduction towards other CSs compared with extinction-only. In addition, generalization of fear reduction has only been measured regarding US expectancy, which may not be an accurate measure for ImRs. Generalization of fear reduction may also be reflected in other, more evaluative CS-related measures, such as CS distress and CS aversiveness (see Boddez et al., 2013).

The aim of the current study was to assess working mechanisms as well as the generalization of fear reduction towards other CSs (including evaluative CS-related measures) between $ImRs_{exp}$ and extinction. Specifically, we aimed to test 1) US revaluation and expectancy learning as working mechanisms of $ImRs_{exp}$ and extinction, respectively; and 2) whether $ImRs_{exp}$ can improve the generalization of fear reduction to other CSs. We compared $ImRs_{exp}$-only, extinction-only (EXT-only) and $ImRs_{exp}$ with extinction ($ImRs_{exp}$ + EXT). Participants first viewed an aversive film clip as an ethically acceptable model for a distressing event (James et al., 2016; Kunze et al., 2015). We tested our research question using a one-day fear conditioning paradigm with three phases (Lonsdorf et al., 2017). In Phase 1, fear acquisition took place with two CS+s, one of which was extinguished and the other was unextinguished (i.e., CS+E and CS+U, respectively) and a CS- that was never paired with the US. Phase 2 consisted of the experimental manipulation ($ImRs_{exp}$-only, EXT-only, and $ImRs_{exp}$ + EXT). The generalization test took place in Phase 3, including both CS+s and the CS-. Note that we used two CS+s for two reasons: 1) to have multiple feared CS+s after acquisition, better mimicking clinical practice; and 2) to determine the relative generalization of the CS+U versus the CS+E (see Pace-Schott et al., 2009). US revaluation was assessed with US distress and US aversiveness ratings. Generalization of fear reduction was measured with US expectancy, CS distress, and CS aversiveness ratings on the CS+E

versus the CS+U. If changes in responses on these ratings to the CS+E also take place for the CS+U (so-called "secondary extinction", Pavlov, 1927), this is evidence for the generalization of fear reduction.

We first predicted that the $ImRs_{exp}$ conditions ($ImRs_{exp}$-only and $ImRs_{exp}$ + EXT) would show more US revaluation as reflected in stronger decreases in US distress and US aversiveness ratings than the EXT-only condition (Hypothesis 1). Second, we predicted that the extinction conditions ($ImRs_{exp}$ + EXT and EXT-only) would show stronger decreases of US expectancy than the $ImRs_{exp}$-only condition (Hypothesis 2). Third, in line with Dibbets et al. (2012), we predicted that $ImRs_{exp}$ + EXT would show increased generalization of US expectancy reduction compared with EXT-only (Hypothesis 3). Fourth, we predicted that the $ImRs_{exp}$ conditions ($ImRs_{exp}$-only and $ImRs_{exp}$ + EXT) would show increased generalization of CS distress and CS aversiveness reduction compared to EXT-only (Hypothesis 4). Finally, recent research showed that participants still think of the US even after successful extinction (Zenses et al., 2021). Thus the CS may still trigger negative thoughts and discomfort associated with the US even in the absence of the expectation of its occurrence (Clark, 2005). Reducing thoughts of the US representation may thus be a fruitful treatment target in addition to US expectancy and revaluation. Therefore, we exploratively assessed how $ImRs_{exp}$ and extinction influenced thinking about the US.

Methods

Preregistration and Ethics

The study was preregistered on the open science framework (OSF) before data analysis. The registration can be accessed at https://osf.io/67xwp. The study was approved by the faculties' ethics committees (G-2022-4772, FETC18-133) and was carried out in accordance with the Declaration of Helsinki.

Participants

The sample size calculation using a repeated measures ANOVA between factors with a medium effect size of $f = .25$, and a power of 80% resulted in 120 participants. A total of 121 participants were recruited at KU Leuven ($n = 21$) and Utrecht University ($n = 100$). Exclusion criteria were: age < 18 or > 40, having experienced physical or sexual abuse, presence of PTSD symptoms (i.e., nightmares, avoidance, or guilt regarding an aversive experience; with or without diagnosis), currently being treated for a psychological disorder, and serious medical problems (e.g., cardiovascular problems). Eight participants were excluded from all data analyses for the following reasons: experiment error ($n = 3$), non-compliance to instructions by skipping the break between experiment phases ($n = 4$), unsuccessful acquisition ($n = 1$). The final sample consisted of 113 participants (93 female; $M_{age} = 21.27$, $SD_{age} = 3.06$). Participants were randomly allocated to one of the three conditions: $ImRs_{exp}$-only ($n = 38$), EXT-only ($n = 38$), or $ImRs_{exp}$ + EXT ($n = 37$). Participants gave written informed consent and received 12 euros or course credit for their participation.

Materials

Film

A film clip (1m27s) extracted from *Irréversible* (Noé, 2002) was used. The clip shows men fighting and screaming and ends with a man being beaten up with a fire extinguisher. The film clip induced distress in previous research and has been used in a study with a similar design (e.g., Arnaudova & Hagenaars, 2017; Woelk et al., 2022).

Conditioning Stimuli

Three pictures were used as conditional stimuli (CSs). The reinforced stimuli (CS+s) were two different fire extinguishers (diagonal presentation with a white label, and horizontal presentation with a blue-white label). One of the CS+s was extinguished (CS+E) and one was unextinguished (CS+U). The unreinforced stimulus (CS-) was a fire reel. The CS+S were

meaningfully connected to the unconditional stimulus (US) to enhance ecological validity (Kunze et al., 2015). The US consisted of a 2 s part of the aversive film clip in which the man is being hit with the fire extinguisher. The sound of this US was slightly louder than in the longer film clip to increase its aversiveness. The CS+E, the CS-, and the US have been used in a previous study (Woelk et al., 2022).

Measures

Questionnaires

PTSD Symptoms. The Primary Care PTSD screen for DSM-5 (PC-PTSD; Prins et al., 2016) was used to screen for PTSD symptoms prior to participation. The questionnaire consists of six questions that can be answered with 'yes' or 'no'. Participants were only excluded if they indicated that they experienced nightmares, avoidance, or guilt regarding the aversive experience.

Neuroticism. Neuroticism is associated with increased fear learning (e.g., Hur et al., 2016). Therefore, we used the neuroticism subscale of the brief version of the Eysenck Personality Questionnaire (EPQ) to assess neuroticism levels for the randomization check (Eysenck & Eysenck, 1975; Sanderman et al., 1991, 2012). Participants were instructed to answer the 12 items (e.g., Do you ever feel miserable for no reason?) with either *yes* (1) or *no* (0). For the analysis, a sum score of all yes's was created ranging from 0 (low neuroticism) to 12 (high neuroticism). Internal consistency in the current sample was acceptable, Cronbach's $\alpha = .77$.

Imagery Vividness. Given the prominent role of mental imagery in ImRs and the potential effect of individual mental imagery vividness on the $ImRs_{exp}$ manipulation (see Dibbets et al., 2012), we used the Dutch version of the Plymouth Sensory Imagery Questionnaire (PsiQ-NL; Andrade et al., 2014; Woelk et al., 2024) to measure the vividness of mental imagery for the randomization check. The PsiQ-NL consists of 35 items assessing

vividness of mental imagery on seven modalities (vision, sound, smell, touch, taste, bodily sensation, and emotion). Participants were instructed to imagine objects or situations (e.g., Imagine the appearance of a cat climbing a tree) and to rate the vividness of their imagery on a scale ranging from 0 (*no image at all*) to 10 (*as vivid as real life*). A mean score across modalities was created for the analysis. Internal consistency in the current sample was excellent, Cronbach's α = .94.

$ImRs_{exp}$ compliance checks. Participants in the $ImRs_{exp}$ conditions rated how well they were able to imagine the new script and the credibility of the script on two VAS ranging from 0 (*not at all*) to 10 (*very much*), and which image (the original film clip or the new imagined script) was strongest on a VAS anchored -5 (*the film*), 0 (*both equal*), and 5 (*the script*).

Stimulus Ratings

US Ratings. Participants were first asked to imagine the US as vividly and detailed as possible, and then rate the subjective distress and aversiveness of the US on a VAS ranging from 0 (*not aversive/distressed at all*) to 100 (*very aversive/distressed*).

CS Ratings. Participants were shown each of the CSs and asked to rate its associated distress and aversiveness on two VAS (presented under the CSs) ranging from 0 (*not aversive/distressed at all*) to 100 (*very aversive/distressed*).

US Expectancy. Participants rated the extent to which they expected the film clip to occur on a visual analogue scale (VAS) ranging from 0 (*not at all expecting the film clip*) to 100 *(definitely expecting the film clip*).

US Thinking. The extent to which participants thought of the US upon seeing the CS was assessed as an explorative measure and therefore not reported in the current paper. See https://osf.io/67xwp/ for all information and results regarding US thinking.

Procedure

See Figure 1a for a graphical representation of the procedure. The entire experiment took place in the laboratory.

Typical Conditioning Trial

Across the experiment, each conditioning trial consisted of a CS presentation (12 s) and an inter-trial interval (ITI; 6-8 s) during which a black screen with a white fixation cross was shown. US expectancy was rated during the first 6 s of CS presentation, US thinking was rated during the last 6 s of each CS presentation. In CS+ trials, the US was shown immediately at CS offset. All trials were presented in random order within blocks consisting of three stimuli (1 CS+E, 1 CS+U, 1 CS-). Participants never saw the same CS more than two times in a row. The number of blocks differed per phase, see below.

[Insert Figure 1 approximately here]

Preacquisition Phase

This phase was the same for all conditions. Upon entering the lab, participants received instructions, read and signed the informed consent form, and put headphones on. The experiment started with the CS ratings, the EPQ, and the PsiQ-NL. This was followed by a mental imagery exercise to prepare participants for the rescripting phase. Participants read a script on preparing lunch and were instructed to imagine this script as vividly and detailed as possible (see Hagenaars & Arntz, 2012). The experimenter guided them through their imagery, focusing on first-person perspective, present tense, and sensory details. To keep the procedure as similar as possible for each condition, EXT-only participants also went through this exercise. After the exercise, participants watched the aversive film clip, with the instruction to immerse oneself in the film as if they were present at the scene. Then, the CS ratings were completed again. All further instructions were presented on the screen.

Practice Phase

Participants (in all conditions) were told that two of the three CSs would be followed by a part of the film clip (US). They were instructed to learn which CSs were reinforced (i.e., followed by the US) and to rate their US expectancy and US thinking with a mouse click on the VAS during CS presentation. This was followed by a practice round to get used to the procedure (i.e., using the VAS slider with time limit), during which two pictures (i.e., a swing and a slide) were shown as practice CSs without the film clip.

Acquisition Phase

The acquisition phase (the same in all conditions) started directly after the practice trials. All CSs were presented eight times in total. The CS+s were followed by the US six out of eight times, with the first and the fifth trial of each CS being unreinforced (i.e., reinforcement rate of 75%; see Lonsdorf et al., 2017). The CS- was never followed by the US. After the acquisition phase, participants completed the CS and US ratings, followed by a 5 min break during which they could read neutral magazines (i.e., travel, interior design, engineering).

Manipulation Phase

See Figure 1b for a schematic representation of trial timing for each manipulation. All instructions were provided before the manipulation started unless otherwise specified.

ImRs$_{exp}$-only. Participants in this condition were given a script related to the aversive film but with a positive ending. In the script, the man is being hit with the fire extinguisher, but the participant calls security immediately and the victim is only mildly injured (see Supplemental Material on OSF for the full script). Note that in line with previous research, the positive script in the ImRs conditions still contained an image of the US to activate the mental representation of the aversive film (e.g., Dibbets et al., 2018; Kunze et al., 2019). Participants were instructed to close their eyes and imagine this script as vividly and detailed as possible upon seeing the following text on the screen: “Imagine the script now”, and to end

their imagery when they heard a tone (i.e., after 8 seconds). They repeated the imagery of the script 12 times (i.e., 12 times 8 seconds; see Woelk et al., 2022).

EXT-only. Participants in this condition were told that they would see two of the three pictures (not specified which ones) from the acquisition phase (i.e., one of the CS+s and the CS-; CS+s counterbalanced) and that they had to indicate their US expectancy and US thinking again. Both CSs were presented 12 times each (i.e., 24 trials in total), all unreinforced.

ImRs$_{exp}$ + EXT. Similar to extinction-only, participants in the ImRs$_{exp}$ + EXT condition were told that they would see two of the three pictures from the acquisition phase (i.e., one of the CS+s and the CS-; CS+s counterbalanced) and were instructed to provide US expectancy and US thinking ratings after CS presentation. In addition, they were instructed to imagine the script at each CS+ offset until they heard a tone (i.e., after 8 seconds). Both CSs were presented 12 times (i.e., 24 trials), and they imagined the script 12 times.

After the manipulation, CS and US ratings were assessed in all conditions. This was followed by a 5 min break of reading the magazines.

Spontaneous Recovery Phase

Participants in all conditions were told that they would see the three pictures from the beginning of the experiment again (i.e., CS+E, CS+U, CS-) and were instructed to rate their US expectancy and US thinking. Four trial blocks were presented with both CS+s and the CS- to assess generalization; all stimuli unreinforced. After this phase, participants completed the CS and US ratings again.

Reinstatement Phase

No further instructions were given after the ratings in the spontaneous recovery phase. A black screen with a white fixation cross appeared for 6-8 seconds, followed by three unexpected US presentations (ITI 6-8 seconds). Four unreinforced trial blocks (CS+E, CS+U,

CS-) followed 6-8 seconds after the last unexpected US presentation. The experiment ended with the final CS and US ratings and the $ImRs_{exp}$ checks.

Data Analysis[1]

For all analyses, significance was set at an α level of .05. Greenhouse-Geisser ($\varepsilon < .75$) or Huyn-Feldt ($\varepsilon \geq .75$) correction was applied in case of a violation of sphericity (Field, 2013). For exploratory analyses, significant main effects were followed by post-hoc tests with Bonferroni correction ($\alpha = .05/3 = .017$). Effect sizes are reported for significant effects only. There were no outliers. Participants with missing data were excluded from the respective analysis. Missing data was only present for US expectancy ratings. These were imputed using regression coefficients based on stimulus-specific US expectancy ratings of the two adjacent trials (e.g., trial 1 and 3 if trial 2 was missing) (see e.g., Zhang, 2016). Data were not imputed if multiple adjacent trials or first and last trials of a phase were missing. The number of excluded participants differed per analysis and ranged from 3 to 13. Specifics are reported in Table S1 in the Supplemental Material (https://osf.io/67xwp/).

Randomization Checks

Differences in gender distribution (i.e., "With which gender do you identify?") between conditions were assessed with a χ^2 test. Differences between conditions in age, neuroticism (EPQ sum score), imagery vividness (PsiQ-NL mean score), and baseline CS (prefilm) and US (premanipulation-phase) ratings were tested with a one-way analysis of variance (ANOVA) with condition ($ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) as the independent variable.

Manipulation Checks

[1] We did not conduct preregistered explorative mediation analyses because, in hindsight, our design did not adhere to the necessary criteria (with the mediator preceding the outcome variable, as outlined by Kazdin, 2007), nor did it encompass the appropriate variables and statistical power required for such analyses.

CS Ratings. To check the impact of the film and the acquisition phase on CS ratings, two 3 (Stimulus: CS+E, CS+U, CS-) × 2 (Time: prefilm vs. postfilm; postfilm vs. postacquisition-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVAs were conducted on CS distress and CS aversiveness[2] ratings separately.

Acquisition. Differential acquisition was analyzed with a 3 (Stimulus: CS+E, CS+U, CS-) × 8 (Time: Acquisition trials 1-8) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA on US expectancy ratings.

Extinction. To assess whether extinction was successful, a 2 (Stimulus: CS+E, CS-) × 12 (Time: Extinction trials 1-12) × 2 (Condition: EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA was conducted on US expectancy.

Hypothesis 1: US Revaluation

US revaluation was assessed with repeated measures ANOVAs on US aversiveness and US distress ratings separately, at different time points. To test for US revaluation pre versus postmanipulation-phase, a 2 (Time: postacquisition-phase vs. postmanipulation-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA was conducted. To test for US revaluation pre versus postspontaneous recovery-phase, we performed a 2 (Time: postmanipulation-phase vs. postspontaneous recovery-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA. To test for US revaluation pre versus postreinstatement-phase, a 2 (Time: postspontaneous recovery-phase vs. postreinstatement-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA was conducted.

Hypothesis 2: US Expectancy

[2] Because of its conceptual overlap, CS aversiveness was included in the main analyses with CS distress (and not as an exploratory measure, as written in the preregistration), expecting the same pattern for both variables.

US expectancy ratings were analyzed with repeated measures ANOVAs. Changes pre versus postmanipulation were assessed with a 2 (Stimulus: CS+E, CS-) × 2 (Time: Acquisition trial 8 vs. Spontaneous Recovery trial 1) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA. Changes in US expectancy pre versus postspontaneous recovery were not assessed because the $ImRs_{exp}$-only condition did not have any US expectancy ratings in the manipulation phase. Change in US expectancy pre versus postreinstatement was analyzed with a 2 (Stimulus: CS+E, CS-) × 2 (Time: Spontaneous Recovery trial 4 vs. Reinstatement trial 1) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures.

Hypothesis 3: Generalization of US Expectancy

Generalization of US expectancy was analyzed with two repeated measures ANOVAs. Pre versus postmanipulation differences were assessed with a 3 (Stimulus: CS+E, CS+U, CS-) × 2 (Time: Acquisition trial 8 vs. Spontaneous Recovery trial 1) × 2 (Condition: EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA. Generalization of US expectancy pre versus postspontaneous recovery was not analyzed, because of the absence of US expectancy ratings before the spontaneous recovery phase in the $ImRs_{exp}$-only condition. Generalization of US expectancy pre versus postreinstatement was analyzed with a 3 (Stimulus: CS+E, CS+U, CS-) × 2 (Time: Spontaneous Recovery trial 4 vs. Reinstatement trial 1) × 2 (Condition: EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA.

Hypothesis 4: Generalization of CS Ratings

Generalization of CS distress and CS aversiveness was analyzed with repeated measures ANOVAs on CS distress and CS valence separately. To assess changes pre to postmanipulation-phase, two 3 (Stimulus: CS+E, CS+U, CS-) × 2 (Time: postacquisition-phase vs. postmanipulation-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVAs were conducted. Changes pre versus postspontaneous recovery-

phase were analyzed with two 3 (Stimulus: CS+E, CS+U, CS-) × 2 (Time: postmanipulation-phase vs. postspontaneous recovery-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVAs. Generalization of CS distress and CS aversiveness pre versus postreinstatement-phase was analyzed with a 3 (Stimulus: CS+E, CS+U, CS-) × 2 (Time: postspontaneous recovery-phase vs. postreinstatement-phase) × 3 (Condition: $ImRs_{exp}$-only, EXT-only, $ImRs_{exp}$ + EXT) repeated measures ANOVA.

Results

See Table 1 for descriptives of baseline measures. See Figures 2 and 3 or Table S2 (Supplemental Material) for a graphical representation and descriptive statistics of CS and US distress and aversiveness ratings throughout the experiment, respectively.

Randomization Checks

Conditions did not differ significantly regarding gender distribution, χ^2 (2) = 0.861, p = .650, age, EPQ, PsiQ, and baseline CS (prefilm) and US (premanipulation-phase) distress and aversiveness baseline ratings (Fs < 1.57, ps > .212).

[Insert Table 1 approximately here]

Manipulation Checks

CS Ratings

Pre versus Postfilm-phase.

CS Distress. CS distress ratings did not differ significantly among stimuli at baseline, $F(2, 220) = 0.77$, $p = .467$. A stimulus × time interaction revealed a significant differential change in CS distress ratings from pre to postfilm-phase, $F(1.34, 147.34) = 92.59$, $p < .001$, $\eta_p^2 = .457$, 95% CI η_p^2 [.340, .547]. That is, CS distress for both CS+s increased more than CS- distress, ts > 10.13, ps < .001, indicating a successful impact of the film on CS distress ratings. Changes in CS distress did not differ significantly between conditions (stimulus × time × condition), $F(2.69, 147.34) = 0.13$, $p = .925$.

CS Aversiveness. CS aversiveness ratings were not significantly different among stimuli at baseline, $F(1.87, 205.22) = 0.05$, $p = .946$. A stimulus × time interaction showed a differential change in CS aversiveness pre versus postfilm, $F(1.33, 145.80) = 105.26$, $p < .001$, $\eta_p^2 = .489$, 95% CI η_p^2 [.375, .576]. Planned contrasts showed that aversiveness for both CS+s increased more than CS- aversiveness, ts > 10.40, ps < .001, reflecting a successful impact of the film on CS aversiveness. There was no evidence for any differences between conditions (stimulus × time × condition), $F(2.65, 145.80) = 0.16$, $p = .907$.

Pre versus Postacquisition-phase.

CS Distress. The stimulus × time interaction revealed a differential change in CS distress ratings pre versus postacquisition-phase, $F(1.48, 162.89) = 59.71$, $p < .001$, $\eta_p^2 = .352$, 95% CI η_p^2 [.236, .448]. Distress increased for both CS+s (ts > 6.45, ps < .001) and decreased for the CS- ($t = 3.87$, $p < .001$), indicating successful differential acquisition of CS distress. There were no significant differences among conditions (stimulus × time × condition, time × condition), Fs < 1.44, ps > .242.

CS Aversiveness. There was a stimulus × time interaction for CS aversiveness ratings, indicating a differential change pre versus postacquisition-phase, $F(1.59, 174.33) = 60.96$, $p < .001$, $\eta_p^2 = .357$, 95% CI η_p^2 [.245, .450]. Aversiveness ratings increased for both CS+s (ts > 6.34, ps < .001) and decreased for the CS- ($t = 4.69$, $p < .001$), suggesting successful acquisition of CS aversiveness. There was no evidence for any differences among conditions (stimulus × time × condition, time × condition), Fs < 2.12, ps > .096.

[Insert Figure 2 approximately here]

US Expectancy

See Figure 4 for a graphical representation of US expectancy ratings throughout the experiment.

Acquisition. There was a significant stimulus × time interaction, $F(7.53, 775.70) = 57.61$, $p < .001$, $\eta_p^2 = .359$, 95% CI η_p^2 [.302, .401], showing that US expectancy increased over time for both CS+s (Fs > 55.70, ps < .001) and decreased over time for the CS- (F[7, 763] = 15.71, $p < .001$). Acquisition did not differ significantly among conditions (stimulus × time × condition, time × condition), Fs < 1.16, ps > .297. Thus, acquisition of US expectancy was successful.

Extinction. There was a significant stimulus × time interaction, $F(4.23, 295.96) = 69.85$, $p < .001$, $\eta_p^2 = .499$, 95% CI η_p^2 [.417, .558]. Subsequent ANOVAs per stimulus revealed main effects of time, showing a decrease in US expectancy over time for the CS+E (F[3.97, 278.14] = 98.94, $p < .001$) and for the CS- (F[5.20, 379.29] = 3.58, $p = .003$) across the two extinction conditions, indicating successful extinction. The extinction pattern over time did not differ between conditions (time × condition, stimulus × time × condition), Fs < 0.34, ps > .858. However, a stimulus × condition interaction, $F(1, 70) = 5.51$, $p = .022$, $\eta_p^2 = .073$, 95% CI η_p^2 [.001, .207], indicated that, after correcting for CS- responding (i.e., CS+E minus CS- US expectancy), $ImRs_{exp}$ + EXT showed higher US expectancy ratings throughout the extinction phase than EXT-only. This indicates a larger discrepancy between CS+E and CS- responding in $ImRs_{exp}$ + EXT.

Hypothesis 1: US Revaluation

Pre versus Postmanipulation-phase

US Distress. The time × condition interaction was significant, $F(2, 110) = 3.60$, $p = .030$, $\eta_p^2 = .062$, 95% CI η_p^2 [<.001, .154]. Planned contrasts showed a larger decrease in US distress pre versus postmanipulation-phase in $ImRs_{exp}$ + EXT than in $ImRs_{exp}$-only, $t(73) = 2.44$, $p = .017$. No other differences between conditions were significant, ts < 1.50, ps > .115. Postmanipulation-phase US distress ratings did not differ between conditions, $F(2, 110) = 0.45$, $p = .638$.

US Aversiveness. There was a significant time × condition interaction, $F(2, 110) = 4.11$, $p = .019$, $\eta_p^2 = .069$, 95% CI η_p^2 [.001, .165]. Planned contrasts showed a larger decrease in US aversiveness pre versus postmanipulation-phase in $ImRs_{exp}$ + EXT compared with $ImRs_{exp}$-only, $t(73) = 2.53$, $p = .014$, and EXT-only, $t(73) = 2.28$, $p = .026$. $ImRs_{exp}$-only and EXT-only did not differ significantly from each other, $t(74) = 0.88$, $p = .384$. Postmanipulation-phase US aversiveness ratings did not differ between conditions, $F(2, 110) = 1.03$, $p = .359$.

Pre versus Postspontaneous Recovery-phase

US Distress. The time × condition interaction was not significant, $F(2, 110) = 0.18$, $p = .832$. A main effect of time indicated a decrease in US distress pre versus postspontaneous recovery-phase across conditions, $F(1, 110) = 21.51$, $p < .001$, $\eta_p^2 = .164$, 95% CI η_p^2 [.056, .285].

US Aversiveness. The time × condition interaction was not significant, $F(2, 110) = 0.34$, $p = .716$. A main effect of time showed a decrease in US aversiveness pre versus postspontaneous recovery-phase across conditions, $F(1, 110) = 24.12$, $p < .001$, $\eta_p^2 = .180$, 95% CI η_p^2 [.067, .302].

Pre versus Postreinstatement-phase

US Distress. There was a significant time × condition interaction, $F(2, 110) = 4.93$, $p = .009$, $\eta_p^2 = .082$, 95% CI η_p^2 [.006, .182]. Planned contrasts showed a larger increase in US distress in $ImRs_{exp}$ + EXT compared with $ImRs_{exp}$-only ($t[73] = 2.91$, $p = .005$) and EXT-only ($t[73] = 2.17$, $p = .033$). $ImRs_{exp}$-only and EXT-only did not differ significantly from each other, $t(74) = 0.94$, $p = .349$. Postreinstatement-phase US distress was still significantly lower than premanipulation-phase US distress in all conditions ($ts > 3.98$, $ps < .001$) and did not differ significantly between conditions ($F[2, 110] = 0.54$, $p = .582$).

US Aversiveness. The time × condition interaction was not significant, $F(2, 110) = 1.93$, $p = .151$. A main effect of time showed a significant increase in US aversiveness from pre to postreinstatement-phase across conditions, $F(1, 110) = 16.10$, $p < .001$, $\eta_p^2 = .128$, 95% CI η_p^2 [.034, .245]. Postreinstatement-phase US aversiveness was still lower than premanipulation-phase US aversiveness across conditions (ts > 3.86, ps < .001) and did not differ significantly between conditions (F[2, 110] = 0.15, $p = .865$).

In sum, we found no evidence for more US revaluation in the two ImRs conditions than in the EXT-only condition. Rather, we found more US revaluation from pre to postmanipulation-phase for $ImRs_{exp}$ + EXT relative to $ImRs_{exp}$-only on both US distress and aversiveness ratings, and relative to EXT-only for US aversiveness ratings. Note that there was also a larger pre to postreinstatement-phase increase in US distress (but not aversiveness) for $ImRs_{exp}$ + EXT (versus $ImRs_{exp}$-only), although postreinstatement-phase ratings were still below premanipulation-phase level, with no differences between conditions. No differences between conditions were found from pre to postspontaneous recovery-phase.

[Insert Figure 3 approximately here]

Hypothesis 2: US Expectancy

Pre versus Postmanipulation

A significant stimulus × time × condition interaction, $F(2, 107) = 8.35$, $p < .001$, $\eta_p^2 = .135$, 95% CI η_p^2 [.030, .247], and planned contrasts showed a decrease in US expectancy for the CS+E in all conditions (ts > 3.51, ps < .002), with larger decreases in the EXT-only and $ImRs_{exp}$ + EXT conditions compared with the $ImRs_{exp}$-only condition (ts > 3.82, ps < .001). EXT-only and $ImRs_{exp}$ + EXT did not differ significantly from each other, $t(72) = 0.03$, $p = .979$. There was no significant change in US expectancy for the CS-, F (1, 110) = 0.11, p =

.743. Conditions did not differ significantly from each other in changes in US expectancy for the CS-, $F(2, 110) = 0.07$, $p = .937$.

Pre versus Postreinstatement

The stimulus × time × condition and time × condition interactions were not significant, indicating that (differential) changes in US expectancy did not differ between conditions, Fs < 1.27, ps > .285. Reinstatement of US expectancy was present across conditions as evidenced by a stimulus × time interaction ($F[1, 99] = 69.53$, $p < .001$). There was a larger increase in US expectancy for the CS+E than for the CS- across conditions, $t(101) = 8.33$, $p < .001$. Postreinstatement US expectancy ratings for the CS+E were still significantly lower than premanipulation levels across conditions, $t(102) = 9.44$, $p < .001$.

In sum, we found evidence for the expected larger pre to postmanipulation decreases in US expectancy for the CS+E in the two EXT conditions versus the $ImRs_{exp}$-only condition. We did not find significant differences between conditions pre versus postreinstatement, reinstatement took place in all conditions.

[Insert Figure 4 approximately here]

Hypothesis 3: Generalization of US Expectancy

Considering the overlap between the analyses for Hypotheses 2 and 3 (specifically, the addition of the CS+U for Hypothesis 3), the results below exclusively contain information that has not been previously discussed in the section for Hypothesis 2 above. Given that the $ImRs_{exp}$-only condition also showed a significant change in US expectancy for the CS+E pre vs. postmanipulation and pre vs. postreinstatement, we exploratively looked at changes in US expectancy for the CS+U and differential change between the CS+E and the CS+U in this condition for both phases.

Pre versus Postmanipulation

A significant stimulus × time interaction, $F(1.82, 122.04) = 77.98$, $p < .001$, $\eta_p^2 = .538$, 95% CI η_p^2 [.414, .622] with planned contrasts showed a decrease in US expectancy for the CS+U in both $ImRs_{exp}$ + EXT and EXT-only (ts > 6.60, ps < .001). $ImRs_{exp}$ + EXT and EXT-only did not differ significantly in US expectancy change for the CS+U, $t(68) = 0.83$, $p = .410$. Within conditions, paired samples t-tests with difference scores (post minus premanipulation) of both CS+s revealed no significant differences in US expectancy changes between the CS+E and the CS+U in $ImRs_{exp}$ + EXT or EXT-only, ts < 1.43, ps > .162.

Exploratory analyses for the $ImRs_{exp}$-only condition showed a decrease in CS+U US expectancy, $t(36) = 3.33$, $p = .002$. Decreases were larger in EXT-only and $ImRs_{exp}$ + EXT compared with $ImRs_{exp}$-only, both ts > 3.41, both ps < .001. Within the $ImRs_{exp}$-only condition, paired samples t-tests with difference scores (post minus premanipulation) of both CS+s revealed no significant difference in US expectancy changes between the CS+E and the CS+U $t(34) = 0.16$, $p = .875$.

Pre versus Postreinstatement

The stimulus × time × condition and time × condition interactions were not significant, indicating that (differential) changes in US expectancy did not differ significantly between $ImRs_{exp}$ + EXT and EXT-only, Fs < 1.94, ps > .149. A main effect of time showed an increase in US expectancy for the CS+U from pre to postreinstatement in both conditions, $F(1, 60) = 65.22$, $p < .001$. Within each of the two conditions, paired samples t-tests with US expectancy difference scores (post minus prereinstatement) revealed no significant differences in US expectancy changes between the CS+E and the CS+U, both ts < 1.13, both ps > .269.

Exploratory analyses for the $ImRs_{exp}$-only condition revealed an increase in US expectancy for the CS+U, $t(36) = 4.96$, $p < .001$, which was not significantly different

compared to the two extinction conditions, both *t*s < 0.781, both *p*s > .438). Within the $ImRs_{exp}$-only condition, paired samples t-tests with difference scores (post minus prereinstatement) revealed no significant difference in US expectancy changes between the CS+E and the CS+U $t(34) = 0.88$, $p = .384$.

In sum, we did not find the expected increased generalization of US expectancy for $ImRs_{exp}$ + EXT compared with EXT-only. The decrease in US expectancy for the CS+U did not differ from the CS+E in any condition. Similarly, the increase in US expectancy for the CS+U during reinstatement did not differ from the increase in US expectancy for the CS+E in any condition. This indicates generalization of US expectancy in all conditions.

Hypothesis 4: Generalization of CS Ratings

Pre versus Postmanipulation-phase

CS Distress. The stimulus × time × condition interaction was significant, $F(3.26, 179.34) = 3.07$, $p = .026$, $\eta_p^2 = .053$, 95% CI η_p^2 [<.001, .114]. Planned contrasts revealed that $ImRs_{exp}$ + EXT and EXT-only showed a larger decrease in CS+E and CS+U distress ratings than $ImRs_{exp}$-only (*t*s > 2.19, *p*s < .032), and were not significantly different from each other, *t*s < 0.82, *p*s > .417. Conditions showed no significant differences in changes in CS- distress, $F(2, 110) = 0.63$, $p = .536$. The decreases in distress for the CS+E versus the CS+U were not significantly different in any condition, as evidenced by non-significant within-condition paired samples t-tests with difference scores (post minus premanipulation-phase), *t*s < 1.57, *p*s > .125.

CS Aversiveness. The stimulus × time × condition interaction was significant, $F(3.31, 182.25) = 4.18$, $p = .005$, $\eta_p^2 = .071$, 95% CI η_p^2 [.007, .138]. Planned contrasts showed a larger decrease in CS+E ($t[54.33] = 3.67$, $p < .001$) and CS+U ($t[59.61] = 2.53$, $p = .014$) aversiveness in the EXT-only condition compared with the $ImRs_{exp}$-only condition. $ImRs_{exp}$ + EXT showed a larger decrease in CS+E ($t[73] = 2.10$, $p = .039$) but not CS+U ($t[73] = 2.00$, *p*

= .050) aversiveness than $ImRs_{exp}$-only. $ImRs_{exp}$ + EXT and EXT-only did not differ significantly from each other, *t*s < 1.63, *p*s > .107. No significant differences between conditions were found for the CS-, $F(2, 110) = 0.74$, $p = .479$. Within conditions, paired samples t-tests with CS aversiveness difference scores (post minus premanipulation-phase) showed a larger decrease in CS+E aversiveness than CS+U aversiveness within the EXT-only condition, $t(36) = 2.06$, $p = .046$, suggesting limited generalization. This contrast was not significantly different in $ImRs_{exp}$ + EXT and $ImRs_{exp}$-only, *t*s < 0.97, *p*s > .341.

Pre versus Postspontaneous Recovery-phase

CS Distress. There were no significant differences between conditions over time (stimulus × time × condition, time × condition), *F*s < 0.31, *p*s > .864. A stimulus × time interaction ($F[2, 220] = 9.15$, $p < .001$) and subsequent paired samples t-tests within stimulus type showed a significant decrease in CS distress for both CS+s (*t*s > 3.43, *p*s < .001) from pre to postspontaneous recovery-phase in all conditions. There was no significant change in CS- distress over time, $t(112) = 0.68$, $p = .501$.

CS Aversiveness. There were no significant differences between conditions (stimulus × time × condition, time × condition), *F*s < 1.05, *p*s > .384. A significant stimulus × time interaction ($F[2, 220] = 7.33$, $p < .001$) and subsequent paired samples t-tests within stimulus type showed a decrease in aversiveness for the CS+E ($t[112] = 2.41$, $p = .018$) and the CS+U ($t[112] = 4.02$, $p < .001$) from pre to postspontaneous recovery-phase in all conditions. There was no significant change in CS- distress over time, $t(112) = 0.62$, $p = .537$.

***Pre versus Postreinstatement*-phase**

CS Distress. The stimulus × time × condition and time × condition interactions were not significant, *F*s < 0.54, *p*s > .710, indicating no significant differences in CS distress change between conditions. A significant stimulus × time interaction ($F[2, 220] = 9.57$, $p < .001$) and follow-up paired samples t-tests within stimulus type showed an increase in distress

for both CS+s (ts >3.49, ps < .001), indicating that reinstatement took place in all conditions. Ratings after the reinstatement phase were lower than premanipulation-phase (ts > 7.62 ps < .001). No significant change in CS- distress over time was observed, $t(112) = 1.05$, $p = .294$.

CS Aversiveness. There were no significant differences between conditions over time (stimulus × time × condition and time × condition), Fs < 1.62, ps > .202. A main effect of time showed a general increase in CS aversiveness from pre to postreinstatement-phase across conditions and stimuli, $F(1, 110) = 6.29$, $p = .014$, indicating a general reinstatement effect. CS aversiveness ratings were still significantly lower at postreinstatement-phase than premanipulation-phase, $t(112) = 8.24$, $p < .001$.

In sum, we did not find the expected larger generalization of CS distress or aversiveness in the two $ImRs_{exp}$ conditions compared to EXT-only. Similar to the pattern of the US expectancy ratings, we found larger decreases in EXT-only (both CS+E and CS+U distress and aversiveness) and $ImRs_{exp}$ + EXT (CS+E distress and aversiveness, and CS+U distress) compared with $ImRs_{exp}$-only from pre to postmanipulation-phase. No differences in CS distress or aversiveness were observed between conditions in the spontaneous recovery and reinstatement phases. CS distress and aversiveness decreased across conditions from pre to postspontaneous recovery-phase and increased from pre to postreinstatement-phase.

$ImRs_{exp}$ compliance checks

Participants indicated being able to imagine the script well in both $ImRs_{exp}$ conditions ($M = 7.23$, $SD = 2.12$), with no differences between conditions, $t(73) = 0.22$, $p = .823$. Mean credibility of the script was 6.28 ($SD = 2.43$), with no differences between conditions, $t(73) = 0.46$, $p = .646$. After the experiment, most participants (69.3%) indicated that their memory of the film was stronger than their memory of the script, 8% had a stronger memory of the script, and 30.7% indicated that both memories were equally strong. This did not differ significantly between conditions, $\chi^2(2) = 2.15$, $p = .341$.

Discussion

The aim of the current study was to test US revaluation and US expectancy learning as potential mechanisms of ImRs and extinction, respectively, as well as to examine differences between ImRs and extinction in the generalization of fear reduction. We found evidence for US revaluation in participants that received a combination of $ImRs_{exp}$ and extinction, as well as enhanced expectancy learning in the conditions that included an extinction procedure. Against expectations, generalization of decreased US expectancy and CS ratings occurred in all conditions, without any difference.

We found partial evidence for enhanced revaluation as a mechanism for ImRs (Hypothesis 1). That is, participants in the combined $ImRs_{exp}$ and extinction condition reported larger decreases in US aversiveness (versus extinction-only and $ImRs_{exp}$-only) and US distress (versus $ImRs_{exp}$-only) during the manipulation. This is in line with Dibbets et al. (2012), who found a decrease in US aversiveness (premanipulation-phase versus postreturn of fear) in $ImRs_{exp}$ combined with extinction, but not in extinction-only. Note that in our previous study, we did not find differences in US revaluation between $ImRs_{exp}$-only, extinction-only, and $ImRs_{exp}$ and extinction combined (Woelk et al., 2022). However, in our previous study the manipulation took place one day after acquisition, and experimental effects are typically small and short-lasting, possibly limiting the detection of differential effects one day later in our previous study. Together with the findings from Dibbets et al. (2012), the findings from the current study suggest US revaluation as a specific mechanism of $ImRs_{exp}$ when combined with extinction. However, not in line with our hypothesis, this effect was specific for the combination of $ImRs_{exp}$ and extinction, and was not present in the $ImRs_{exp}$-only condition. This may be because in $ImRs_{exp}$-only, the link between the script and the CS+ was not explicitly made during the experimental procedure. Merely mentally rescripting the US representation may not be sufficient to induce US revaluation learning, i.e., seeing actual

CSs are more effective than mental CSs to cause US revaluation. At first sight, this is not in line with previous findings, which have shown that mental images may replace actual stimuli (see Mertens et al., 2020, for a review). However, in those studies, participants imagined the specific CS, US, or CS → US association (e.g., Arntz et al., 1997; Krypotos et al., 2020; Reddan et al., 2018), whereas our participants imagined a complex storyline. Possibly, revaluation learning may require explicit imagination of the CS. Future studies could disentangle the effects of different script contents (e.g., explicit vs. non-explicit CS imagination) on $ImRs_{exp}$ efficacy and US revaluation effects. Relatedly, the script in our $ImRs_{exp}$ + EXT (but not $ImRs_{exp}$-only) condition may have led to some sort of counterconditioning. That is, the script possibly served as a positive US countering the earlier acquired CS+ → negative US association (see Arntz, 2012, for a similar argument). Future research may look into counterconditioning as a potential mechanism of ImRs.

As predicted, we found evidence for enhanced expectancy learning in the extinction conditions compared to $ImRs_{exp}$-only (Hypothesis 2), replicating Woelk et al. (2022). Thus, both studies provide evidence for US expectancy learning as a *specific* effect of extinction. We further found less successful extinction (i.e., an overall higher discrepancy between US expectancy ratings for the CS+E versus the CS-) in $ImRs_{exp}$ combined with extinction than in extinction-only. Thus, $ImRs_{exp}$ seems to interfere with expectancy learning. This is in line with the slower extinction rate in the combined $ImRs_{exp}$ and extinction condition in the two previous studies (Dibbets et al., 2012; Woelk et al., 2022). If reducing US expectancy is the main target in treatment, these results suggest that mere exposure would be the best strategy.

Although we hypothesized that the combination of $ImRs_{exp}$ and extinction would outperform extinction-only in terms of generalization of US expectancy, there were no differences between these two conditions (Hypothesis 3). This seems to contradict previous fear extinction studies finding limited generalization of fear reduction in terms of US

expectancy (e.g., Barry et al., 2016; Vervliet & Geens, 2014; Zbozinek & Craske, 2018). However, note that in the current study, different from those studies, we used two CS+s in the acquisition phase. The use of one of the acquisition CS+s as the extinction stimulus (instead of a novel CS) might have promoted generalization in both extinction conditions by the process of secondary extinction (Pavlov, 1927; see Pace-Schott et al., 2009, for similar results). Indeed, previous research showed that an extinction procedure resulted in less generalization if a CS was used that was not presented during acquisition (Vervliet et al., 2004, 2005; Wong & Lovibond, 2020). Importantly, the $ImRs_{exp}$-only condition also showed generalization of reduced US expectancy even in the absence of any explicit CS+. The presence of a fire extinguisher (i.e., one of the CS+) in the $ImRs_{exp}$ script might have led to a conceptual, less specific mental representation of the acquisition CS+s, leading to a decrease of US expectancy for both stimuli in this condition. It remains unknown whether this decreased US expectancy would also transfer to new stimuli with shared features. Future research could use a paradigm with only one acquisition CS+ and multiple CS+s after the manipulation to further disentangle generalization effects in $ImRs_{exp}$.

Finally, we did not find differences between conditions in generalization of CS distress or CS aversiveness reductions (Hypothesis 4). That is, in all conditions, changes in CS+E ratings did not significantly differ from changes in CS+U ratings. Rather, CS distress and aversiveness ratings followed a similar pattern as US expectancy, with larger decreases for both CS+s in the two extinction conditions. This suggests that the effects of an extinction procedure are not limited to changes in US expectancy, but also influence other properties of the CS, including CS evaluation (see also e.g., van Dis et al., 2019). Future research could investigate the interplay between US expectancy and CS-related evaluative measures to gain more insight into the processes involved in (generalization of) fear reduction.

It should be noted that our $ImRs_{exp}$ + EXT condition is not simply an addition of $ImRs_{exp}$ to extinction. The CS+ was linked to the $ImRs_{exp}$ script during the manipulation in our combined condition, which may have led to additional effects such as counterconditioning. Our combined condition was based on previous research comparing $ImRs_{exp}$ with extinction and kept the same way to enhance comparability (Dibbets et al., 2012; 2018, see also Woelk et al., 2022). Future research might opt to include a phased condition in which $ImRs_{exp}$ of the US is followed by extinction. Whereas Kunze et al. (2019) did employ such procedure, they did not compare the sequence of $ImRs_{exp}$ and extinction to an extinction-only or $ImRs_{exp}$-only condition.

Our study had some limitations that need to be addressed. First, we used an experimental ImRs manipulation that differs from the clinical intervention. With our manipulation, we aimed to obtain high experimental control while at the same time mimicking the clinical ImRs component of mentally rescripting a negative event to something more positive, in line with $ImRs_{exp}$ conditions in other non-clinical studies (e.g., Dibbets et al., 2012, Kunze et al., 2015, Woelk et al., 2022). Adaptations from the clinical intervention included using a standardized script instead of personalized imagery, self-guided rather than therapist-guided sessions, and a shorter duration than clinical ImRs sessions (see e.g., Arntz & Weertman, 1999). The impact of such adaptions is currently unclear. More research should be done to advance our understanding of how findings from experimental ImRs studies may be applied to the clinical ImRs intervention. In addition, the standardized script may have been more relevant to some participants than to others. This might have led to a suboptimal $ImRs_{exp}$ manipulation for some participants, possibly leading to a weaker effect. Second, we employed a 1-day fear conditioning paradigm, which may have resulted in relatively weak consolidation of acquisition and extinction learning. Such a paradigm carries the risk of an 'immediate extinction deficit' (i.e., weak fear reduction after extinction; e.g., Maren, 2014, 2022).

However, our results do not necessarily suggest such an effect. Instead, they appear similar to findings from comparable studies using a multiple-day paradigm (e.g., Kunze et al., 2019; Pace-Schott et al., 2009; Woelk et al., 2022, see also Van Dis et al., 2019). Moreover, experimental effects may be too weak for a multiple-day application of the fear conditioning paradigm.

On the other hand, it is worth highlighting that our study has some strengths. First, we assessed evaluative CS ratings in addition to the commonly used US expectancy ratings. Second, we assessed ratings of all CSs after the manipulation phase. Most studies assessing the generalization of fear reduction only include the unextinguished (or generalization) stimulus after the extinction phase and assess the transfer of US expectancy from the extinction stimulus to the generalization stimulus, which limits the conclusions because it remains unknown how the extinguished stimulus would have been rated relative to the non-extinguished or generalization stimulus.

To conclude, we found evidence for increased US revaluation in $ImRs_{exp}$ combined with extinction compared with both solo manipulations, and increased expectancy learning in extinction. We did not find enhanced generalization of decreased US expectancy or CS ratings in any condition. This study provides more insight into potential mechanisms of ImRs versus exposure. Future research in (sub)clinical samples is recommended to further disentangle the role of US revaluation, expectancy learning, and generalization of fear reduction in ImRs.

Acknowledgements

The authors would like to thank Zuzanna Kowal, Sal Lubberts, Maya Schaap, Eva Knoors, Jasmijn Partridge, and Lieke de Ruiter for their assistance in the data collection.

References

Andrade, J., May, J., Deeprose, C., Baugh, S.-J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth sensory imagery questionnaire. *British Journal of Psychology*, *105*, 547–563. https://doi.org/10.1111/bjop.12050

Arnaudova, I., & Hagenaars, M. A. (2017). Lights … action: Comparison of trauma films for use in the trauma film paradigm. *Behaviour Research and Therapy*, *93*, 67–77. https://doi.org/10.1016/j.brat.2017.02.007

Arntz, A. (2011). Imagery rescripting for personality disorders. *Cognitive and Behavioral Practice*, *18*, 466–481. https://doi.org/10.1016/j.cbpra.2011.04.006

Arntz, A. (2012). Imagery rescripting as a therapeutic technique: Review of clinical trials, basic studies, and research agenda. *Journal of Experimental Psychopathology*, *3*, 189–208. https://doi.org/10.5127/jep.024211

Arntz, A., Spit, S., & Merckelbach, H. (1997). Cued UCS rehearsal and the impact of painful conditioned stimuli: UCS rehearsal increases SCRs but reduces experienced pain. *Behaviour Research and Therapy*, *35*(7), 591–605. https://doi.org/10.1016/S0005-7967(97)00020-X

Arntz, A., & Weertman, A. (1999). Treatment of childhood memories: Theory and practice. *Behaviour Research and Therapy*, *37*, 715–740. https://doi.org/10.1016/S0005-7967(98)00173-9

Barry, T. J., Griffith, J. W., Vervliet, B., & Hermans, D. (2016). The role of stimulus specificity and attention in the generalization of extinction. *Journal of Experimental Psychopathology*, *7*, 143–152. https://doi.org/10.5127/jep.048615

Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: Validity of US expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, *44*, 201–206.

https://doi.org/10.1016/j.jbtep.2012.08.003

Bruce, S. E., Yonkers, K. A., Otto, M. W., Eisen, J. L., Weisberg, R. B., Pagano, M., Shea, M. T., & Keller, M. B. (2005). Influence of psychiatric comorbidity on recovery and recurrence in generalized anxiety disorder, social phobia, and panic disorder: A 12-year prospective study. *American Journal of Psychiatry*, *162*, 1179–1187. https://doi.org/10.1176/appi.ajp.162.6.1179

Clark, D. A. (2005). *Intrusive thoughts in clinical disorders: Theory, research, and treatment*. The Guilford Press.

Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*, 20170025. https://doi.org/10.1098/rstb.2017.0025

Davey, G. C. L. (1989). UCS revaluation and conditioning models of acquired fears. *Behaviour Research and Therapy*, *27*, 521–528.

Davey, G. C. L. (1997). A conditioning model of phobias. In G. C. L. Davey (Ed.), *Phobias - A handbook of theory, research and treatment* (pp. 301–322). Wiley.

Dibbets, P., Lemmens, A., & Voncken, M. (2018). Turning negative memories around: Contingency versus devaluation techniques. *Journal of Behavior Therapy and Experimental Psychiatry*, *60*, 5–12. https://doi.org/10.1016/j.jbtep.2018.02.001

Dibbets, P., Poort, H., & Arntz, A. (2012). Adding imagery rescripting during extinction leads to less ABA renewal. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*, 614–624. https://doi.org/10.1016/j.jbtep.2011.08.006

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. Hodder and Stoughton.

Field, A. (2013). *Discovering statistics using IBM SPSS*. Sage.

Haesen, K., & Vervliet, B. (2015). Beyond extinction: Habituation eliminates conditioned skin conductance across contexts. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2014.11.010

Hagenaars, M. A., & Arntz, A. (2012). Reduced intrusion development after post-trauma imagery rescripting; An experimental study. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*, 808–814. https://doi.org/10.1016/j.jbtep.2011.09.005

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis, second edition: A regression-based approach*. The Guilford Press.

Hofmann, S. G. (2008). Cognitive processes during fear acquisition and extinction in animals and humans: Implications for exposure therapy of anxiety disorders. *Clinical Psychology Review*, *28*, 199–210. https://doi.org/10.1016/J.CPR.2007.04.009

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, *36*, 427–440. https://doi.org/10.1007/s10608-012-9476-1

Hur, J., Iordan, A. D., Berenbaum, H., & Dolcos, F. (2016). Emotion-attention interactions in fear conditioning: Moderation by executive load, neuroticism, and awareness. *Biological Psychology*, *121*, 213–220. https://doi.org/10.1016/j.biopsycho.2015.10.007

James, E. L., Lau-Zhu, A., Clark, I. A., Visser, R. M., Hagenaars, M. A., & Holmes, E. A. (2016). The trauma film paradigm as an experimental psychopathology model of psychological trauma: Intrusive memories and beyond. *Clinical Psychology Review*, *47*, 106–142. https://doi.org/10.1016/j.cpr.2016.04.010

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

Kar, N. (2011). Cognitive behavioral therapy for the treatment of post-traumatic stress disorder: A review. *Neuropsychiatric Disease and Treatment*, *7*, 167–181. https://doi.org/10.2147/NDT.S10389

Kazdin, A. E. (2007). Mediators and Mechanisms of Change in Psychotherapy Research. *Annual Review of Clinical Psychology, 3*, 1–27. https://doi.org/10.1146/ANNUREV.CLINPSY.3.022806.091432

Krypotos, A. M., Mertens, G., Leer, A., & Engelhard, I. M. (2020). Induction of conditioned avoidance via mental imagery. *Behaviour Research and Therapy*, *132*, 103652. https://doi.org/10.1016/J.BRAT.2020.103652

Kunze, A. E., Arntz, A., & Kindt, M. (2015). Fear conditioning with film clips: A complex associative learning paradigm. *Journal of Behavior Therapy and Experimental Psychiatry*, *47*, 42–50. https://doi.org/10.1016/j.jbtep.2014.11.007

Kunze, A. E., Arntz, A., & Kindt, M. (2019). Investigating the effects of imagery rescripting on emotional memory: A series of analogue studies. *Journal of Experimental Psychopathology*, *10*, 1–22. https://doi.org/10.1177/2043808719850733

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Maren, S. (2014). Nature and causes of the immediate extinction deficit: A brief review. *Neurobiology of Learning and Memory*, *113*, 19–24. https://doi.org/10.1016/J.NLM.2013.10.012

Maren, S. (2022). Unrelenting fear under stress: Neural circuits and mechanisms for the immediate extinction deficit. *Frontiers in Systems Neuroscience*, *16*, 888461. https://doi.org/10.3389/FNSYS.2022.888461

Mertens, G., Krypotos, A. M., & Engelhard, I. M. (2020). A review on mental imagery in fear conditioning research 100 years since the 'Little Albert' study. *Behaviour Research and Therapy*, *126*, 103556. https://doi.org/10.1016/j.brat.2020.103556

Morina, N., Lancee, J., & Arntz, A. (2017). Imagery rescripting as a clinical intervention for aversive memories: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *55*, 6–15. https://doi.org/10.1016/j.jbtep.2016.11.003

Noé, G. (2002). *Irréversible*.

Pace-Schott, E. F., Milad, M. R., Orr, S. P., Rauch, S. L., Stickgold, R., & Pitman, R. K. (2009). Sleep promotes generalization of extinction of conditioned fear. *Sleep*, *32*, 19–26. https://doi.org/10.5665/sleep/32.1.19

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford University Press.

Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J., & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical Psychology Review*, *30*, 635–641. https://doi.org/10.1016/J.CPR.2010.04.007

Prins, A., Bovin, M. J., Smolenski, D. J., Marx, B. P., Kimerling, R., Jenkins-Guarnieri, M. A., Kaloupek, D. G., Schnurr, P. P., Kaiser, A. P., Leyva, Y. E., & Tiet, Q. Q. (2016). The Primary Care PTSD Screen for DSM-5 (PC-PTSD-5): Development and evaluation within a veteran primary care sample. *Journal of General Internal Medicine*, *31*, 1206–1211. https://doi.org/10.1007/s11606-016-3703-5

Reddan, M. C., Wager, T. D., & Schiller, D. (2018). Attenuating neural threat expression with imagination. *Neuron*, *100*(4), 994-1005.e4.

https://doi.org/10.1016/J.NEURON.2018.10.047

Sanderman, R., Arrindell, W. A., Ranchor, A. V., Eysenck, H. J., & Eysenck, S. B. G. (2012). *Het meten van persoonlijkheidskenmerken met de Eysenck Personality Questionnaire (EPQ), een handleiding.* (2nd ed.). UMCG / Rijksuniversiteit Groningen, Research Institute SHARE.

Sanderman, R., Eysenck, S. B. G., & Arrindell, W. A. (1991). Crosscultural comparison of personality traits: The Netherlands and England. *Psychological Reports*, *69*, 1091–1096. https://doi.org/10.2466/PR0.69.8.1091-1096

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy*, *86*, 87–94. https://doi.org/10.1016/J.BRAT.2016.08.015

Schottenbauer, M. A., Glass, C. R., Arnkoff, D. B., Tendick, V., & Gray, S. H. (2008). Nonresponse and dropout rates in outcome studies on PTSD: Review and methodological considerations. *Psychiatry*, *71*, 134–168. https://doi.org/10.1521/psyc.2008.71.2.134

Springer, K. S., Levy, H. C., & Tolin, D. F. (2018). Remission in CBT for adult anxiety disorders: A meta-analysis. *Clinical Psychology Review*, *61*, 1–8. https://doi.org/10.1016/j.cpr.2018.03.002

van Dis, E. A. M., Hagenaars, M. A., Bockting, C. L. H., & Engelhard, I. M. (2019). Reducing negative stimulus valence does not attenuate the return of fear: Two counterconditioning experiments. *Behaviour Research and Therapy*, *120*, 103416. https://doi.org/10.1016/j.brat.2019.103416

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

clinpsy-050212-185542

Vervliet, B., & Geens, M. (2014). Fear generalization in humans: Impact of feature learning on conditioning and extinction. *Neurobiology of Learning and Memory*, *113*, 143–148. https://doi.org/10.1016/J.NLM.2013.10.002

Vervliet, B., Vansteenwegen, D., Baeyens, F., Hermans, D., & Eelen, P. (2005). Return of fear in a human differential conditioning paradigm caused by a stimulus change after extinction. *Behaviour Research and Therapy*, *43*, 357–371. https://doi.org/10.1016/J.BRAT.2004.02.005

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory*, *11*(5), 555–558. https://doi.org/10.1101/LM.77404

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2006). Generalization gradients for acquisition and extinction in human contingency learning. *Experimental Psychology*, *53*(2), 132–142. https://doi.org/10.1027/1618-3169.53.2.132

Woelk, M., Hagenaars, M. A., & Krans, J. (2024). Validation of the Dutch version of the Plymouth sensory imagery questionnaire. *European Journal of Psychological Assessment, 40,* 73-83. https://doi.org/10.1027/1015-5759/a000729

Woelk, M., Krans, J., Raes, F., Vervliet, B., & Hagenaars, M. A. (2022). Imagery rescripting versus extinction: Distinct and combined effects on expectancy and revaluation learning. *Clinical Psychological Science*, *10*, 622–639. https://doi.org/10.1177/21677026211055169

Wong, A. H. K., & Lovibond, P. F. (2020). Generalization of extinction of a generalization stimulus in fear learning. *Behaviour Research and Therapy*, *125*, 103535. https://doi.org/10.1016/J.BRAT.2019.103535

Zbozinek, T. D., & Craske, M. G. (2018). Pavlovian extinction of fear with the original

conditional stimulus, a generalization stimulus, or multiple generalization stimuli. *Behaviour Research and Therapy*, *107*, 64–75. https://doi.org/10.1016/J.BRAT.2018.05.009

Zenses, A.-K., Baeyens, F., Beckers, T., & Boddez, Y. (2021). Thought onditioning: Inducing and reducing thoughts about the aversive outcome in a fear-conditioning procedure. *Clinical Psychological Science*, *9*, 252–269. https://doi.org/10.1177/2167702620954222

Author contributions

Mandy Woelk: Conceptualization, Software, Formal analysis, Investigation, Visualization, Writing – Original Draft; **Muriel A. Hagenaars:** Conceptualization, Funding acquisition, Supervision, Writing – Review & Editing; **Filip Raes:** Conceptualization, Project administration, Writing – Review & Editing; **Bram Vervliet:** Conceptualization, Writing – Review & Editing; **Julie Krans:** Conceptualization, Funding acquisition, Supervision, Writing – Review & Editing

Declaration of Interest

None.

Funding

This work was supported by the Research Foundation – Flanders (FWO-Vlaanderen; grant number G069918N) awarded to J. Krans and M. A. Hagenaars.

Table 1

Baseline Characteristics per Condition. N or Means (standard deviations).

Variable	$ImRs_{exp}$-only (*n* = 38)	EXT-only (*n* = 38)	$ImRs_{exp}$ + EXT (*n* = 37)
Gender (m/f/x)	5/33/0	8/30/0	6/31/0
Age	20.97 (2.35)	21.63 (4.08)	21.22 (2.49)
EPQ	4.55 (3.12)	4.95 (3.15)	4.27 (2.50)
PsiQ	6.81 (1.05)	7.27 (1.30)	6.92 (1.37)

Note. ImRs = Imagery Rescripting; EXT = extinction; m = male; f = female; x = non-binary; EPQ = Eysenck Personality Questionnaire; PsiQ = Plymouth Sensory Imagery Questionnaire