Behaviour Research and Therapy 143 (2021) 103848

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Intrusive memories as conditioned responses to trauma cues: An empirically supported concept?

Laila K. Franke [a,*], Julina A. Rattel [a], Stephan F. Miedl [a], Sarah K. Danböck [a], Paul-Christian Bürkner [b], Frank H. Wilhelm [a]

[a] Division of Clinical Psychology and Psychopathology, Department of Psychology, Paris-Lodron-University Salzburg, Austria
[b] Department of Computer Science, Aalto University, Finland

ARTICLE INFO

Keywords:
Posttraumatic stress disorder
Pavlovian conditioning
Exposure therapy
Extinction
Trauma film
Imagery
Intrusive memories

ABSTRACT

Intrusions in posttraumatic stress disorder (PTSD) are clinically understood as conditioned responses (CRs) to trauma-cues; however, experimental evidence for this is limited. We subjected 84 healthy participants to a differential conditioned-intrusion paradigm, where neutral faces served as conditioned stimuli (CSs) and aversive film clips as unconditioned stimuli (USs). While one group only completed acquisition, another group additionally received extinction. Subsequently, participants provided detailed e-diary intrusion reports. Several key findings emerged: First, participants in both groups re-experienced not only USs but also CSs as content of their intrusions. Second, intrusions were elicited by cues resembling CSs, USs, and experimental context. Third, extinction reduced probability and severity of US intrusions, and accelerated their decay, and this was particularly the case in participants showing greater cognitive (US-expectancy) and physiological (SCR) differential responding to CS+ vs. CS- at end of acquisition (i.e., conditionability). Similarly, extinction reduced CS-intrusion probability and severity, but only in participants with greater cognitive conditionability. These results support conditioning's role in re-experiencing in two critical ways: (1) Conditioning during trauma provides cues that not only function as reminder cues, but also as content of intrusions; (2) After strong conditioning, weakening the original CS-US relationship via extinction reduces intrusion formation after analogue-trauma.

1. Introduction

Intrusive memories, i.e., the involuntary and recurrent retrieval of highly aversive events, are common and typically subside over the course of days (Steil & Ehlers, 2000). However, they can persist much longer and be involved in the development and maintenance of mental disorders (Brewin, Gregory, Lipton, & Burgess, 2010). This holds particularly in posttraumatic stress disorder (PTSD), where intrusions of the traumatic event(s) constitute a cardinal symptom (American Psychiatric Association, 2013). But what exactly are intrusions, and how do they arise? From a clinical perspective, intrusive memories have often been conceptualized as conditioned responses (CRs) to cues resembling stimuli present during the traumatic event (trauma-related cues; Ehlers et al., 2002). Here, we aim to experimentally investigate to what extent conditioning processes contribute to intrusive memory formation.

1.1. Cue-driven nature of intrusions

A cue-driven nature of involuntary retrieval is already recognized by autobiographical memory accounts (i.e., memories about one's personal past). These accounts assume that involuntary retrieval relies on associative processes, where cues sharing sensory-perceptual properties with the original event are thought to drive automatic activations of episodic representations (Berntsen, 2009; Conway, 2001). In the same vein, prominent PTSD theories assume that cues sensory-perceptually resembling stimuli that were present during or before the traumatic event re-activate sensory representations of the traumatic event in the form of intrusions (Brewin et al., 2010; Ehlers & Clark, 2000). The warning-signal hypothesis (Ehlers et al., 2002) still goes one step further in postulating that cues (e.g., approaching footsteps) spatiotemporally associated with the traumatic event signaling impending danger (e.g., attack) not only function as reminder cues for intrusions, but are themselves also contents of intrusions (Ehlers et al., 2002). However,

* Corresponding author. University of Salzburg, Department of Psychology, Hellbrunner Straße 34, 5020, Salzburg, Austria.
E-mail address: lailakatharina.franke@sbg.ac.at (L.K. Franke).

https://doi.org/10.1016/j.brat.2021.103848
Received 19 August 2020; Received in revised form 11 February 2021; Accepted 15 March 2021
Available online 2 April 2021

this idea has yet to be tested by experimental research.

1.2. Pavlovian fear conditioning as a framework for understanding intrusions

One framework for understanding the process through which cues associated with a traumatic event become capable of evoking emotional responses, and as we argue here, also intrusive memories, is Pavlovian conditioning theory (Craske, Hermans, & Vansteenwegen, 2006; De Houwer, 2020; Mineka & Oehlberg, 2008). During fear acquisition, originally neutral stimuli (conditioned stimuli, CSs; e.g., neutral faces) become associated with the traumatic event (unconditioned stimulus, US; e.g., exposure to aversive details of a traumatic film) through spatial-temporal proximity. Following, these CSs can elicit strong affective and physiological responses (conditioned responses, CRs) in the absence of the US. In PTSD, re-encountered CSs may continue producing CRs years after the traumatic event. Extinction training aims at gradually weakening this CS-US association and CRs by repeatedly exposing individuals to CSs in the absence of the US (Foa, Steketee, & Rothbaum, 1989; Pitman et al., 2012; Rothbaum & Davis, 2003). Fear extinction is a key procedure of exposure therapy (Vervliet, Craske, et al., 2013), and indeed, better fear extinction learning has been found to predict success of exposure therapy in anxious individuals (Ball, Knapp, Paulus, & Stein, 2017; Forcadell et al., 2017; Waters & Pine, 2016).

In this light, conditioning processes may not only explain how intrusions *arise* after individuals learned a CS-US association (fear acquisition), but also how intrusions can be *reduced* by weakening this association (extinction learning). Taken together, if conditioning processes indeed underlie intrusion formation, individuals should report trauma-related cues (experimental CSs) as trauma reminder cues; moreover, as proposed by the warning-signal hypothesis (Ehlers et al., 2002), they should also report trauma-related cues as content of their intrusions. Further, weakening the CS-US association by introducing an extinction training should reduce intrusive memories.

1.2.1. The relationship between conditionability during trauma and intrusions

According to the cognitive model of PTSD (Ehlers & Clark, 2000), persistent PTSD is associated with enhanced fear learning during the traumatic event: stronger CS-US associations are thought to facilitate trauma-related cues to trigger memories of the event. However, recent findings from extinction research allude to a scenario where exhibiting a stronger CS-US association after trauma could also prove "beneficial" in that it may aid in subsequent exposure-therapy success and durability. Specifically, these findings suggested that the success and durability of extinction was modulated by the degree to which participants expected the US before receiving extinction training; in that a greater expectancy of an aversive event before extinction was thought to maximize the discrepancy between expectancy and reality (i.e., surprise) during extinction, and thereby drive learning/retention of extinction (Culver, Vervliet, & Craske, 2015; Struyf, Hermans, & Vervliet, 2018). In this way, current evidence raises the possibility that on the one hand, a stronger CS-US association during trauma may facilitate subsequent cue-driven re-experiencing in the form of intrusions but, on the other hand, enable successful and durable extinction and thereby reduce intrusions.

The present study therefore aims to investigate whether a stronger CS-US association, here indexed by conditionability (i.e., the degree to which participants respond differently to a danger (CS+) vs safety (CS-) cue at the end of fear acquisition; Orr et al., 2000), facilitates intrusion formation in individuals who only undergo fear-acquisition but, in contrast, reduce intrusions in individuals who receive an immediate extinction training after fear acquisition.

1.2.2. Shortcomings of Pavlovian conditioning in investigating intrusions

While mixed findings from the fear-conditioning field may relate to the use of different methods and samples, it might be the case that whereas standard conditioning procedures address the cue-driven nature of fear responses, they do not tap well into the sensory-perceptual nature of intrusive memories. Specifically, Pavlovian conditioning procedures tend to use aversive USs such as electrical stimulation or aversive sounds, and subsequently focus on CS-invoked fear responses, leaving out clinically-relevant phenomena such as intrusions as possible CRs. As intrusive images of the traumatic event are core symptoms of PTSD, in order to optimally model PTSD within fear-conditioning studies, it seems imperative to extend the current scope of these studies to also include stimuli that allow the study of intrusive imagery (Mertens, Krypotos, & Engelhard, 2020).

1.3. Combining Pavlovian conditioning and trauma-film paradigm to investigate intrusions

In order to overcome these shortcomings, we previously developed the conditioned-intrusion paradigm (Wegerer, Blechert, & Wilhelm, 2013; Wegerer, Blechert, Kerschbaum, et al., 2013). The conditioned-intrusion paradigm combines a Pavlovian conditioning procedure with the trauma-film paradigm. More precisely, the conditioned-intrusion paradigm switches standard shock/aversive tone USs with trauma-films, which have been successfully used to induce intrusive memories (Holmes & Bourne, 2008; James et al., 2016). By integrating a more complex and ecologically-valid US within a differential fear-conditioning procedure, with neutral auditory cues functioning as CSs ("warning-signals"), the conditioned-intrusion paradigm enables studying intrusions as CRs in a controlled experimental setting. In specific, following successful fear acquisition, during the so-called memory triggering task, CSs are embedded in a neutral soundscape to provoke intrusions in a laboratory setting. This task resonates with the idea that intrusions are triggered by cues temporally associated with the traumatic event, and is congruent with a number of other studies that have already successfully elicited intrusions following the trauma-film paradigm with reminder cues such as film stills or auditory cues (Lau-Zhu, Holmes, & Porcheret, 2018).

Using the conditioned-intrusion paradigm, we and others have already shown that 1) CS+ cues presented during a memory triggering task resulted in heightened anxiety and skin conductance responses (SCRs), and most importantly, 2) intrusive memories were reported in daily life up to one week after acquisition (Streb, Conway, & Michael, 2017; Wegerer, Blechert, Kerschbaum, et al., 2013). More recently, we extended these findings by showing that reduced extinction mediated gender differences in intrusion formation during the memory triggering task and in daily life (Rattel, Wegerer, et al., 2019). Similarly, we showed that participants with sustained differential conditioned responding during late extinction in anterior insula and dorsal anterior cingulate cortex, two core nodes of the saliency network (Seeley, 2019), reported more intrusions during daily life (Miedl et al., 2020).

1.4. Intrusions as CRs to trauma-related cues: open questions

While previous studies hint at intrusive memories being triggered by CSs, and at impaired extinction of CRs correlating with intrusion formation, they leave some questions unaddressed. First, it is, to the best of our knowledge, yet unclear whether CSs can figure as content of intrusive memories. Second, while current evidence suggests that CSs may trigger intrusions in the laboratory (Rattel, Wegerer, et al., 2019; Streb et al., 2017; Wegerer, Blechert, Kerschbaum, et al., 2013), whether this also holds in daily life remains unknown. Critically, unlike in the laboratory memory triggering task setting where participants are confronted with CSs identical to the original CSs, in daily life, individuals encounter stimuli that likely only share some features with the original CSs (Barry, Griffith, Vervliet, & Hermans, 2016). Relatedly, it remains to be explored whether, and to what extent, other innocuous stimuli surrounding the analogue-trauma, beyond experimental CSs, may elicit

intrusions in daily life. Lastly, while we have previously shown that attenuated extinction increased intrusion formation (Miedl et al., 2020; Rattel, Miedl, et al., 2019), these findings are limited in that they did not allow to dissociate whether the link between extinction learning and intrusion formation was of causal nature or rather due to a shared underlying vulnerability (e.g., trait anxiety, Clark, Mackay, & Holmes, 2015; Haaker et al., 2015) that rendered participants not only more sensitive to developing intrusions following analogue-trauma, but also more likely to show deficient extinction (see Visser, 2020).

1.5. Current study

In the current study, we aim at shedding light on these remaining questions and thereby extend the experimental investigation of sensory-perceptual intrusive images as CRs. For this purpose, we used an adapted version of the conditioned-intrusion paradigm (Wegerer, Blechert, Kerschbaum, et al., 2013). Here, one neutral face (CS+) was paired with short aversive film clips (USs), and another neutral face (CS-) was paired with short neutral film clips.[1] We opted for using neutral faces instead of the so-far used neutral tones (Wegerer, Blechert, Kerschbaum, et al., 2013) mainly because faces constitute social and evolutionary fear-relevant stimuli (Öhman & Mineka, 2001) that are commonly encountered during daily life, and work better within a noisy MRI setting. Fear-acquisition was followed by an immediate extinction phase. Crucially, in order to experimentally investigate extinction effects on intrusion formation, one group of participants underwent acquisition only (ACQ-only), while another group received extinction (ACQ+EXT). Subsequent to the laboratory part of the experiment, participants were instructed to report CS intrusions and/or US intrusions, together with short content descriptions as well as corresponding reminder cues ("triggers"), for four consecutive days in a smartphone e-diary app. Within this Pavlovian conditioning experimental setting, we formulated a number of hypotheses to further test the assumption that intrusive memories can be conceptualized within a conditioning framework.

1.6. Hypotheses

1.6.1. CSs as content of intrusive memories

Firstly, we hypothesized that participants re-experience not only USs (i.e., aversive film clips) but also CSs (i.e., innocuous faces associated with the USs [CS+]). Given that CRs to cues not associated with the aversive USs [CS-], the so-called "safety-cues" are also possible and have shown to discriminate healthy individuals from individuals with anxiety-related disorders (Craske et al., 2012, 2009; Duits et al., 2015; Lissek & van Meurs, 2015), we expected that participants also re-experience CS- as intrusive memories in daily life.

1.6.2. CSs as reminder cues for intrusive memories

Secondly, we expected that participants occasionally identify experimental CSs, next to other innocuous cues associated with the USs, such as experimental-context cues, as reminder cues for their intrusive memories in daily life.

1.6.3. Extinction effects on intrusive memories

Thirdly, we hypothesized that individuals not receiving extinction (ACQ-only) should maintain stronger CS-US associations, thereby being more likely to form intrusions than participants receiving extinction (ACQ+EXT) following fear acquisition. Further, we expected that group differences (ACQ-only vs. ACQ+EXT) are moderated by conditionability. Specifically, in line with Ehlers and Clark (2000), we assumed that greater conditionability at end of acquisition is associated with greater intrusive memory formation in ACQ-only participants. However, considering findings regarding the positive effect of expectancy-violation on extinction success and durability (Culver et al., 2015; Struyf et al., 2018), we expected that greater conditionability at the end of acquisition would be associated with lower intrusion formation in ACQ+EXT participants. Finally, using multilevel modelling, we also explored whether extinction training accelerates the well-documented decay of intrusions over time (Rattel, Grünberger, et al., 2019; Steil & Ehlers, 2000) and whether participants receiving extinction (ACQ+EXT), in dependence of conditionability, displayed a more accelerated decay of intrusions over testing days than participants not receiving extinction (ACQ-only). We expected to find these effects both with respect to US- as well as CS-intrusion formation.

2. Methods

2.1. Participants

The current sample included German-speaking, female students, aged between 18 and 35 years. Participants were recruited at the local university and through social media as part of a larger project investigating neural mechanisms predictive of intrusive memories. Results on this data set, based on the ACQ+EXT group, are reported in Miedl et al., 2020 and Rattel, Miedl, et al., 2019. Exclusion criteria included current mental or neurological disorder, major medical illnesses, and regular medication use (except for oral contraceptives). Furthermore, participants reporting enhanced (more than 2–3 times a week) consumption of extremely violent media or blood/injury phobia were excluded. The former criterium was used to reduce the odds of including participants in whom exposure to aversive film scenes would potentially elicit insufficient distress. Unlike for instance a painful stimulus, that normally provokes a universal unconditioned response (e.g., fear, pain), aversive film material may not elicit sufficient distress in everyone. Possible candidates in whom this may occur could be participants who are habituated to such material by voluntary and frequent exposure to violent media. All participants provided written informed consent approved by the local Ethics Committee prior to participation, and received Euro 70 or study credits as compensation.

2.2. Sample characteristics

We assessed trait anxiety with the State-Trait-Anxiety Inventory (trait and state form, STAI-T; (Spielberger, Gorsuch, & Lushene, 1970), German version (Laux, Glanzmann, & Schaffner, 1981)), depressive symptoms with the Centre for Epidemiologic Studies Depression Scale (CES-D (Radloff, 1977), German version (Meyer & Hautzinger, 2001)), and lifetime adversity with the Traumatic Life Experiences Questionnaire (TLEQ (Kubany et al., 2000); German version (Teegen, 2003)). The TLEQ assesses 22 types of potentially traumatic events that may meet DSM-5 PTSD criterion A1 definition for a traumatic event. For each endorsed event, participants indicated the number of times the event was experienced: "*never*", "*1*" to "*5*", or "*more than five times*". Further, we assessed conceptual and data-driven processing during analogue-trauma with the Conceptual Processing Questionnaire (CPQ) and peritraumatic dissociation with the Peritraumatic Behavior Questionnaire (Halligan, Michael, Clark, & Ehlers, 2003; Murray, Ehlers, & Mayou, 2002).

2.3. Apparatus and physiological recordings

Stimulus presentation and behavioral data acquisition were controlled by E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, USA). Skin conductance (SC) was measured using Ag/AgCl electrodes filled with isotonic electrode paste (Boucsein et al., 2012); electrodes were placed on the lower palm of the left hand. Recording of SC data was performed with a sampling rate of 1000 Hz using Polybench

[1] Note that although reinforcing the CS- is rather unusual in standard conditioning procedures, in the current study it was necessary for specific fMRI analyses (reported in Miedl et al., 2020 and Rattel, Miedl, et al., 2019).

1.22 (TMSi, Twente Medical Systems International, EJ Oldenzaal, Netherlands), a Porti 32-channels-amplifier (TMSi), and an SC-amplifier (Becker Meditec, Karlsruhe, Germany). ANSLAB 2.6 was used for SC analysis (Blechert, Peyk, Liedlgruber, & Wilhelm, 2016; Wilhelm & Peyk, 2005).

2.4. General procedure

After inclusion, participants completed trait questionnaires and rated 19 neutral faces (Radboud Faces Database; Langner et al., 2010), with the two most neutral in valence and least arousing being later used in the conditioning procedure as CS+ and CS-, individually selected for each participant. Approximately one week later, participants underwent the conditioned intrusion paradigm in a magnetic resonance imaging (MRI) scanner. Part of the sample was randomly assigned to undertake acquisition-only (ACQ-only group, $N = 26$), while another part was allocated to undergo acquisition and immediate extinction (ACQ+EXT group, $N = 58$). The unequal sample size was due to the study's focus on neural processes during extinction (Miedl et al., 2020) and cost-constraints. The conditioning procedure was preceded and followed by 8-min resting periods. During conditioning, we measured participants' skin conductance and acquired experiential ratings. Subsequent to conditioning, participants recorded intrusive memories and associated reminder cues on their smartphones for four consecutive days. On the fifth experimental day, participants re-assessed CSs valence.

2.5. Conditioned-intrusion paradigm

The current study adopted the previously developed conditioned-intrusion paradigm (Wegerer, Blechert, Kerschbaum, et al., 2013) and adapted it for functional magnetic resonance imaging (Miedl et al., 2020; Rattel, Miedl, et al., 2019). For each participant, we randomly assigned the two faces that the participants pre-rated as most neutral, and least arousing, to serve as CS+ and CS-. While the CS+ was subsequently followed by aversive film clips (USs), the CS- was followed by neutral film clips. We selected six different aversive film-clips (16s duration) as USs to prevent habituation to a single film-clip over repeated presentations. These film clips were extracted from commercial movies, depicting interpersonal trauma ("Hostel", 2005, directed by Eli Roth; "Antichrist", 2009, directed by Lars von Trier; and "Scar", 2007, directed by Jed Weintrob), and accidental trauma ("127 h", 2010, by Danny Boyle; "Dantes Peak", 1997, by Roger Donaldson; and "Final Destination", 2000, by James Wong). Neutral film-clips were closely matched with respect to length, number of actors, movements, number of film cuts, and sound pitch of background sounds, and depicted people in non-violent interactions (two scenes from "Die Frau des Polizisten", 2013, by Philip Groening; three scenes from "Trois couleurs: Bleu", 1993, by Krzysztof Kieslowski; and one scene from "Trois couleurs: Blanc", 1994, Krzysztof Kieslowski). Each of the six aversive, and each of the six neutral film clips was presented twice to each participant.

2.5.1. Fear conditioning

Fear conditioning consisted of an acquisition phase for all participants (ACQ-only and ACQ+EXT participants), and subsequent fear extinction for part of the sample (ACQ+EXT participants). At the beginning of acquisition, we told participants that "Two faces will be shown to you. Both faces are occasionally followed by a short film clip. However, only one of the two faces is occasionally followed by an aversive film clip". Fig. 1 depicts the acquisition (a) and extinction (b) phases.

2.5.2. Ratings

Baseline face ratings. Each participant evaluated 19 faces from the Radboud Face Database (Radboud Faces Database; Langner et al., 2010)

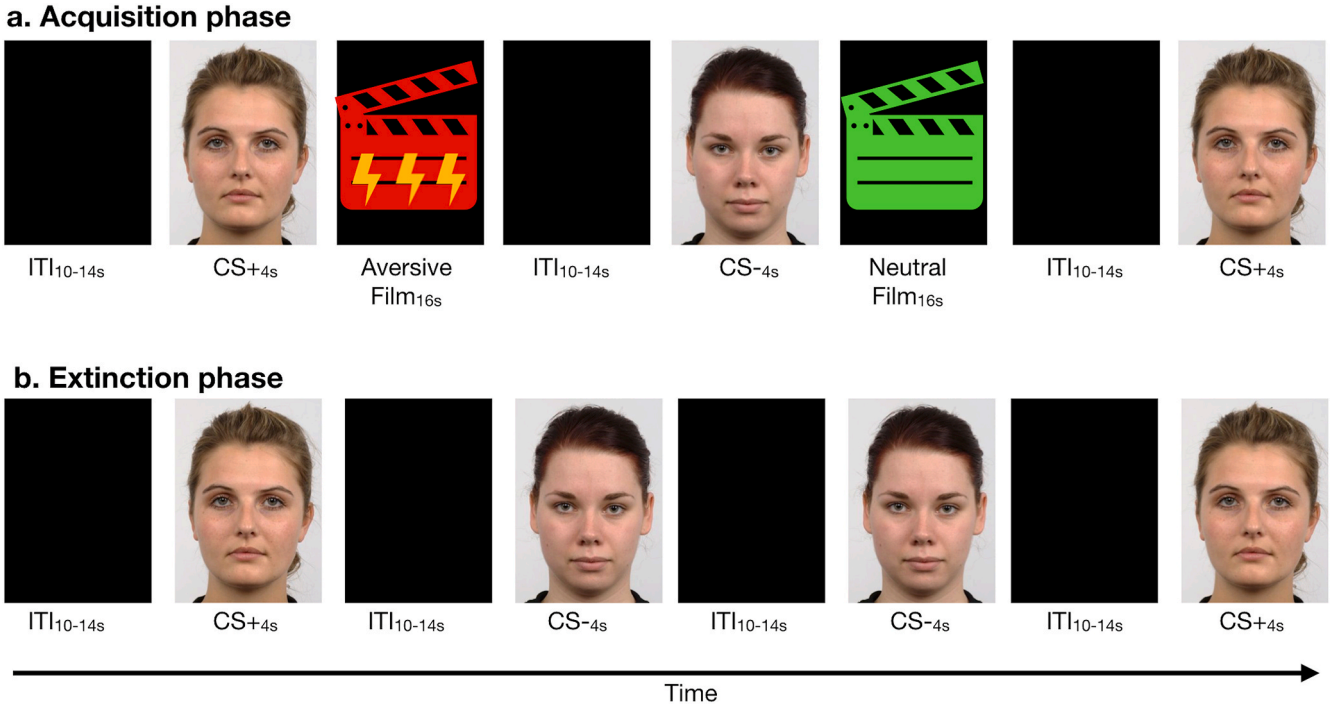

Fig. 1. Examples of conditioning trials of the acquisition (a) and the extinction (b) phases. The acquisition phase (a) started with a blank black screen (ITI of 10–14s), followed by a CS presentation of 4s. In reinforced trials, aversive film clips followed the CS+, whereas neutral film clips followed the CS-. In unreinforced trials, no film clips followed the CS+ or the CS-. In total, there were 12 CS+ and 12 CS- reinforced trials, and 4 CS+ and 4 CS- unreinforced trials (75% reinforcement rate). After the first and second acquisition halves, participants rated CSs on US-expectancy and Valence. The extinction phase (b) immediately followed acquisition in ACQ+EXT participants; 16 CS+ and 16 CS- were presented without subsequent film-clips. After the first and second halves of extinction, participants again rated CSs on US-expectancy and Valence. Order of film clips and trials were pseudorandomized and counterbalanced across participants, with no more than three stimuli of the same type (CS+ and CS-) presented consecutively.

on Valence (0 = *not unpleasant at all*, 100 = *very unpleasant*) and Arousal (0 = *not arousing at all*; 100 = *very arousing*).

Unconditioned responses (UR) to film clips. At the end of the second acquisition phase, participants rated aversive and neutral film-clips for valence and arousal while seeing reminder pictures from the clips. Nine-point Likert-type scales assessed valence (1 = *not unpleasant at all*, 9 = *very unpleasant*) and arousal (1 = *very low*, 9 = *very high*).

Conditioned responses (CR) to CSs. All participants rated valence and US-expectancy to CS+ and CS- after first and second acquisition halves; ACQ+EXT participants also rated valence and US-expectancy after first and second extinction halves. Valence ratings were further provided on the fifth experimental day online from home, four days after conditioning (post-assessments). Such post-assessments were omitted for US-expectancy, as having participants expecting CSs to be followed by an aversive film-clip outside of the experimental context (i.e., outside of the MRI scanner, at home) was deemed highly unlikely.

Again, nine-point Likert-type scales assessed Valence (1 = *not unpleasant at all*, 9 = *very unpleasant*) and US-expectancy (i.e., degree of certainty that the CS would be followed by an aversive film-clip; 1 = *surely not*, 9 = *surely yes*).

2.6. Ambulatory assessment of intrusive memories and reminder cues

After the conditioning procedure, participants were instructed to register intrusive memories on a smartphone app in an event-based manner, i.e., as soon as they occurred (Rattel, Grünberger, et al., 2019) for four consecutive days, starting at the evening of the experiment. Intrusive memories were defined to participants as "memories about the film clips or faces, which could be images, sounds or thoughts about the film-clips or faces, but also recurring thoughts or feelings that had been present during watching" (Ehring, Fuchs, & Kläsener, 2009; Zetsche, Ehring, & Ehlers, 2009). Participants were instructed to report only memories occurring spontaneously, i.e., without deliberate recall. For each intrusive memory entry, participants were requested to first indicate content of intrusion (either to film clips or faces), briefly describe each intrusion (e.g., "man in the mountains"), indicate associated distress on a visual-analogue scale (0 = *not distressing at all*, 100 = *extremely distressing*), as well as modality (*visual, auditory, thought, feeling*). Further, participants were asked whether they identified a reminder cue for the intrusive memory (*yes, no*). In case a cue was identified, participants were prompted to briefly describe it (e.g., "saw tools").

2.7. Data reduction

2.7.1. Unconditioned responses (URs)

URs were operationalized as mean valence and arousal ratings for aversive film clips at the end of acquisition. Skin conductance responses (SCRs) to aversive films (UR_{SCR}) were averaged over the second acquisition half and quantified by subtracting mean baseline skin conductance level (SCL, −2 to 0 s relative to the CS onset) from the maximum SCL during the 16s film clips, considering only the first presentation of each film clip. Average individual UR was calculated by averaging URs over all films and normalizing UR_{SCRs} using the natural logarithm of 1 + SCR. Due to technical problems, data from eight participants (ACQ-only: N = 2; ACQ+EXT: N = 6) are missing for UR_{SCR} analyses.

2.7.2. Conditionability indices

For all three conditioned responses (US-expectancy and valence ratings; SCR), we calculated differential scores by subtracting second acquisition halve (Acq2) CS- values from Acq2 CS+ values (Dunsmoor, Prince, Murty, Kragel, & LaBar, 2011; Wegerer, Blechert, Kerschbaum, et al., 2013). For conditionability as indexed by SCR ($Conditionability_{SCR}$), average pre-CS baseline SCL (−2 to 0 s relative to CS onset) was subtracted from maximum CS SCL (0–5 s relative to the CS onset). $Conditionability_{SCR}$ scores were normalized using the natural logarithm of 1 + SCR.

2.7.3. Classification of intrusive memory content

The partitioning into US and CS intrusions was based on participants reports: upon registering their intrusive memory in the app, participants were prompted to indicate whether their intrusions pertained to the film (US) or face (CS) category. We classified intrusions in two steps. First, based on participants' descriptions provided in the e-diary, intrusive memories were classified into 36 subcategories representing possible intrusive memory contents; second, based on subcategories defined in the first step, intrusive memories were grouped into a total of seven final categories: US-interpersonal trauma, US-accidental trauma, US-mixed, CS+, CS-, CS-mixed. See Appendix A, Table A, for details on the intrusion classification system.

2.7.4. Classification of reminder cue content

We defined a classification system where reminder cues for intrusive memories could be classified in two domains: (A) according to content, and, for descriptive purposes also (B) according to origin. Regarding content, we were primarily interested in discriminating cues according to elements of the experiment (i.e., USs, CSs, wider experimental context cues) they referred to. Accordingly, the classification system included four categories, allowing cues to be categorized into stimuli resembling (1) USs film-clips; (2) CSs faces; (3) wider experimental context; or lastly (4) unspecific elements, not attributable to categories 1–3. Regarding trigger origin, cues could be classified into stimuli of (1) external; (2) internal; or (3) mixed source. See Appendix B, Table B, for details on the cue classification system.

2.8. Statistical analysis

We used JASP (JASP Team, 2020) for frequentist analyses and R-Studio (RStudio JASP Team, 2020) in R (RCore Team, 2019) for estimating Bayesian multilevel models (BMLMs) via the ***brms*** package using Stan (Bürkner, 2017a; Carpenter et al., 2017). In repeated measures analyses of variance (RM ANOVAs) significant main-effects and interactions were followed by post hoc analyses using Bonferroni's correction. The respective effect sizes were partial eta squared (η_p^2) and Cohen's d (*d*). When sphericity was violated as indicated by Mauchly's test, we report the Greenhouse-Geisser corrected *P*-values and degrees of freedom.

2.8.1. Sample characteristics

ACQ-only-, and ACQ+EXT-participants were characterized in terms of their average age, depressive symptoms (ADS-L), trait and state anxiety (STAI-T; STAI-S), the sum score for TLEQ; as well as peritraumatic dissociation (PDEQ), and cognitive processing during analogue-trauma (CPQ). To examine potential differences between groups, we used independent samples *t*-tests.

2.8.2. Manipulation checks

Unconditioned responses. To examine unconditioned responses, we used three separate RM ANOVAs for film valence ratings, film arousal ratings, and SCRs during film viewing as outcome measures. Film-type was added as within-subject factor and, to test for potential group differences, Group was added as a between-subject factor in all analyses, resulting in three separate 2 (Film-type; aversive/neutral) × 2 (Group; ACQ-only/ACQ+EXT) RM ANOVAs.

Baseline CS ratings. To examine whether CS+ and CS- were indeed similar in valence and arousal at baseline, we computed two separate RM ANOVAS for each dependent variable (Valence, Arousal), and added CS-type (CS+/CS-) as within-subject factor, and Group (ACQ-only/ACQ+EXT) as between-subject factor.

Acquisition of conditioned responses. To examine the acquisition of conditioned responses, we computed three separate RM ANOVAs for each dependent variable US-expectancy rating, Valence rating, and SCR for CSs, where CS-type (CS+/CS-) and time (first acquisition [Acq1] and second acquisition [Acq2] halves) were added as within-subject factors.

To examine potential differences in conditioning between ACQ-only and ACQ+EXT groups, Group was added as a between-subject factor in all analyses, resulting in three separate 2 (CS-type; CS+/CS-) × 2 (Time; Acq1/Acq2) × 2 (Group; ACQ-only/ACQ+EXT) RM ANOVAs.

Extinction of conditioned responses. To test for extinction effects on US-expectancy and Valence ratings in the ACQ+EXT group, we extended the above-mentioned RM ANOVAS with ratings from the first extinction halve (Ext1) and second extinction halve (Ext2), resulting in 2 (CS-type; CS+/CS-) × 4 (Time; Acq1/Acq2/Ext1/Ext2) RM ANOVAs.

Long-term retention of conditioned responses (Valence). To assess retention of conditioned responses (i.e., to what extent participants still evaluated the CS+ as more unpleasant than the CS- on the fifth experimental day), Valence analyses also included post-assessment (Post) ratings, resulting in a 2 (CS-type; CS+/CS-) × 2 (Time; Acq2/Post) RM ANOVA.

2.8.3. Extinction effects on overall US and CS intrusions

To examine whether, depending on the degree to which participants acquired differential CRs to CS+ vs. CS- (i.e., depending on participants' *conditionability*), ACQ+EXT participants were less likely to report US intrusions and/or CS intrusions than ACQ-only participants we estimated Bayesian multilevel regression models (BMLMs) (Gelman et al., 2013; Gelman & Hill, 2006). We fitted separate BMLMs for each outcome (US intrusions, CS intrusions). Akin to previous approaches (e.g., Rattel, Miedl, et al., 2019), we operationalized intrusions as "intrusion load" consisting of the product of daily intrusion number and average distress, which is equivalent to the sum of daily intrusive distress. By weighting intrusions for their distress, we procured a more clinically-relevant variable, since persistent PTSD is primarily linked to intrusions perceived as very distressing (Steil & Ehlers, 2000). Based on our theory-based hypotheses, we added three predictors of US intrusions and CS intrusions to our models: Group (ACQ-only/ACQ+EXT), Conditionability (indexed by (I) US-expectancy ratings [$Conditionability_{US\text{-}EXP}$], (II) Valence ratings [$Conditionability_{VAL}$], and (III) SCRs [$Conditionability_{SCR}$]) from end of acquisition phase, and Day (i.e., experimental day on which intrusive memory was registered), as well as the interaction Group × Conditionability × Day. For an overview of fitted models please see Appendix C.

Predictors were centered before being entered in BMLMs: (1) Group was effect coded (ACQ-only = −0.5, ACQ+EXT = +0.5); (2) Conditionability indices were centered to their respective means; and (3) Day was centered on the first 24h day after analogue-trauma, i.e., on the second experimental day. To account for the dependency between observations due to repeated measurement of intrusions over participants, responses by the same person were modelled with varying intercepts. Further, as we expected that the effect of Day (i.e., the decay) on intrusions could vary between participants, we added a varying slope for the effect of day. Although model comparisons using *K-fold* cross-validation suggested a slightly better predictive ability of random-intercept-only than random-slopes models (see Appendix D), as these differences were negligible, we opted to maintain a random slope to model potential variability in intrusion decay over days between participants (Barr, Levy, Scheepers, & Tily, 2013).[2]

To account for the greater than expected (i.e., inflation of) instances of zero intrusions in the data, we modelled our data with a hurdle lognormal distribution (Li, Elashoff, Robbins, & Xun, 2011; Tooze, Grunwald, & Jones, 2002). With this approach, we fitted data in two parts where (A) estimates the probability of not experiencing (i.e. zero) vs. experiencing (i.e. non-zero) intrusions (hurdle part, modelled with a binomial distribution); and where (B) estimates the intrusion load (i.e., severity of) intrusions > 0 (lognormal part, modelled with a lognormal distribution).

For a summary of model parameters, we report regression coefficients and 95% credible intervals (CIs; i.e., Bayesian confidence intervals). Based upon CIs, we can state that there is a 95% probability that the respective parameter falls within this interval, given the evidence provided by the data, prior, and model assumptions. Effects were considered significantly different from zero if the estimate's 95% CIs did not include zero (this would indicate statistical significance on a 5% level). For directed hypotheses, we also estimated the posterior probability that the parameter of interest was in the expected direction (Gelman et al., 2013). As priors we used the weak- or non-informative default priors of brms, which have only negligible influence on the obtained results (Bürkner, 2017a, 2017b). We report Bayesian R^2 as our measure for effect sizes (Gelman, Goodrich, Gabry, & Vehtari, 2019). All Bayesian models converged according to common algorithms-agnostic (Vehtari, Gelman, Simpson, Carpenter, & Bürkner, 2021) and algorithm-specific diagnostics (Betancourt, 2017). Specifically, there were no divergent transitions, $Rhat < 1.01$ and $ESS > 400$ for all relevant parameters.

2.8.4. Exploratory sub-analyses on intrusive memory classification outcomes

Besides our main analyses on the two main intrusion categories of interest (US/CS), we estimated additional explorative BMLMs to investigate whether (1) CS-type (CS+ vs. CS-) influenced the probability and number of CS intrusions, and (2) Movie-type (interpersonal vs. accidental trauma) influenced the probability and number of US intrusions. Intrusions were again operationalized as "load" (i.e., intrusion number × distress). In both instances, we also tested whether potential effects held for ACQ-only and ACQ+EXT participants, and included interaction terms between CS-type and Group, and Movie-type and Group to the respective models. Given the zero inflation in the response variables, we again specified BMLMs with a hurdle lognormal distribution and modelled responses by the same person with varying intercepts. Predictors were effect coded: Group: ACQ-only = −0.5, ACQ+EXT = +0.5/ CS-type: CS- = −0.5, CS+ = +0.5/Movie-type: accidental = −0.5, interpersonal = +0.5.

Lastly, to compare the proportion of individuals reporting US vs. CS intrusions we fitted a Bayesian binomial logistic regression with weakly informative priors (Bürkner, 2017a, 2017b). Here, CS-type (CS- = −0.5, CS+ = +0.5) and Group (ACQ-only = −0.5, ACQ+EXT = +0.5) were entered as predictors, and the proportion of participants reporting intrusions was the outcome variable (i.e., the total number of individuals experiencing an intrusion within the total number of individuals in each Group). For ease of interpretation, we reported exponentiated estimates. As exponentiating 0 = 1, 95% CIs were deemed statistically significant if the respective interval did not contain 1.

3. Results

Results are reported in four sections. The first part describes sample characteristics; the second part describes results of analyses of manipulation checks (unconditioned responses, acquisition of conditioned responses, extinction of conditioned responses, and long-term retention of conditioned responses, Table 2); the third part describes content classification outcomes of intrusive memories (Tables 3–4) and reminder cues (Table 5); and lastly, the fourth part describes results of analyses of extinction effects on US intrusions (Table 6) and CS intrusions (Table 7).

[2] Given our previous observations that intrusions tend to be relatively high during the first experimental days but monotonically flatten out over the last experimental day (Rattel, Grünberger, et al., 2019), we contrasted models assuming a linear relationship with a model assuming a monotonic relationship between Day and intrusions (Bürkner & Charpentier, 2020). Results of sensitivity analyses (see Appendix E) suggested only negligibly better predictions for the monotonic models; and for some CS-intrusion predictions, the linear models slightly outperformed the monotonic model. In this light, we decided to report the simpler linear models here, instead of the more complex monotonic models.

Table 1
Participants' general background information, separately displayed for ACQ-only (N = 26) and ACQ + EXT (N = 58) participants.

	ACQ-only	ACQ+EXT
	M (*SD*)	*M* (*SD*)
Age	21.38 (2.37)	22.90 (4.18)
ADS-L	8.19 (4.48)	9.83 (5.93)
STAI-T	33.58 (5.47)	35.10 (6.90)
Δ STAI-S	1.82 (6.91)	2.88 (9.55)
TLEQ A1 events	5.77 (3.92)	6.41 (5.84)
PDEQ	8.81 (6.68)	11.10 (7.75)
CPQ - CP	16.00 (3.44)	15.29 (3.52)
CPQ - DD	10.62 (5.35)	11.62 (5.93)

Note: Abbreviations: ACQ-only = acquisition-only group; ACQ+EXT = acquisition plus extinction group; ADS-L = General Depression Scale (Long); STAI-T = Spielberger State-Trait Anxiety Inventory-Trait; Δ STAI-S = change score Spielberger State-Trait Anxiety Inventory-State post minus pre analogue-trauma; TLEQ = Traumatic Life Events Questionnaire; A1 event = experience of an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of oneself or others; PDEQ = Peritraumatic Dissociation Questionnaire; CPQ = Cognitive Processing Questionnaire; CP = Conceptual Processing; DD = Data-driven Processing.

3.1. Sample characteristics

A final sample of 84 participants was included in the analyses. Of these, 26 underwent only the fear-acquisition phase (ACQ-only), and 58 underwent this same fear-acquisition plus immediate extinction (ACQ+EXT). Table 1 displays general characteristics of the sample, divided by group. Trait anxiety and depression symptoms were within normal ranges. There were no significant differences (all *P*'s > 0.089) between ACQ-only and ACQ+EXT groups in any of these background or analogue-trauma reactivity variables.

3.2. Manipulation checks

3.2.1. Unconditioned response[3]

Participants experienced aversive film clips as more unpleasant ($M = 7.21$, $SD = 1.06$) than neutral film clips ($M = 1.69$, $SD = 0.87$; $F_{(1,81)} = 1109.32$, $p < .001$, $\eta_p^2 = 0.873$), as well as more arousing ($M = 6.63$, $SD = 1.31$) than neutral film clips ($M = 2.18$, $SD = 1.13$; $F_{(1,81)} = 370.20$, $p < .001$, $\eta_p^2 = 0.820$). Similarly, SCRs were higher in response to aversive than neutral film clips ($F_{(1,74)} = 32.94$, $p < .001$, $\eta_p^2 = 0.308$). These effects held for both ACQ-only and ACQ+EXT participants, as indicated by non-significant Group × Film-type interactions (Valence: $F_{(1,82)} = 1.23$, $p = .264$, $\eta_p^2 = 0.005$; arousal: $F_{(1,81)} = 0.41$, $p = .525$, $\eta_p^2 = 0.001$; SCRs: $F_{(1,74)} = 0.10$, $p = .749$, $\eta_p^2 = 0.001$).

3.2.2. Conditioning procedure

Fig. 2 displays an overview of ACQ-only- and ACQ+EXT-participants' average CS+ and CS- US-expectancy (I) and Valence ratings (II), as well as SCRs (III) over the conditioning procedure. Statistical parameters of analyses are displayed in Table 2.

3.2.2.1. Baseline CSs ratings. At baseline, analyses suggested that CS+ and CS- did not differ in Valence, and that this was the case both in ACQ-only (CS+: $M = 47.10$, $SD = 10.75$; CS-: $M = 45.56$, $SD = 10.28$) and ACQ+EXT (CS+: $M = 44.11$, $SD = 10.33$; CS-: $M = 45.63$, $SD = 11.07$) participants ($F_{(1,\ 79)} = 1.07$, $p = .305$, $\eta p^2 = 0.013$). Analyses further suggested that, at baseline, CS+ and CS- also did not differ in Arousal, again holding for both ACQ-only (CS+: $M = 18.72$, $SD = 15.85$; CS-: $M = 16.16$, $SD = 14.48$) and ACQ+EXT participants (CS+: $M = 19.79$, $SD = 12.41$; CS-: $M = 21.31$, $SD = 15.07$) participants ($F_{(1,\ 79)} = 1.96$, $p = .165$, $\eta p^2 = 0.024$).[4]

3.2.2.2. Acquisition of conditioned responses. ***Self-report.*** US-expectancy (i.e., extent of certainty that an aversive film-clip would follow the CS) analyses showed that over acquisition, participants rated CS+ faces as more likely than CS- faces to be followed by the US ($t_{(82)} = 18.44$, $p_{bonf} < .001$, $d = 2.01$). A CS-type × Time interaction suggested that whereas CS+ US-expectancy was identical in Acq1 and Acq2 ($t_{(82)} = -1.93$, $p_{bonf} = .329$, $d = -0.21$), CS- US-expectancy was higher in Acq1 than in Acq2 ($t_{(82)} = 4.69$, $p_{bonf} < .001$, $d = 0.51$). These effects held for ACQ-only and ACQ+EXT participants, as indicated by non-significant CS-type × Time × Group effects. Also, analyses suggested no group differences in CS- US-expectancy at Acq1 ($t_{(82)} = 0.14$, $p_{bonf} > .999$, $d = 0.02$) and Acq2 ($t_{(82)} = -0.31$, $p_{bonf} > .999$, $d = -0.03$), or CS+ US-expectancy at Acq1 ($t_{(82)} = 0.72$, $p_{bonf} > .999$, $d = 0.08$) and Acq2 ($t_{(82)} = 0.76$, $p_{bonf} > .999$, $d = 0.08$). For details, see Table 2A.I, and Fig. 2.I.

Similarly, Valence (i.e., unpleasantness) analyses suggested that over acquisition, participants rated CS+ cues as more unpleasant than CS- cues ($t_{(82)} = 8.86$, $p_{bonf} < .001$, $d = 0.97$). A CS-type × Time interaction, which was modulated by Group, suggested that whereas ACQ+EXT participants perceived the CS- as more unpleasant in Acq1 than in Acq2 ($t_{(82)} = 3.87$, $p_{bonf} = .004$, $d = 0.42$), ACQ-only participants perceived CS- as equally unpleasant in both acquisition halves ($t_{(82)} = -0.71$, $p_{bonf} > .999$, $d = -0.8$). Despite this difference, ACQ+EXT participants rated CS- as unpleasant as ACQ-only participants at Acq1 ($t_{(82)} = 1.14$, $p_{bonf} > .999$, $d = 0.12$) and Acq2 ($t_{(82)} = -1.19$, $p_{bonf} > .999$, $d = -0.13$). Further, participants perceived CS+ as equally unpleasant in Acq1 and Acq2 (ACQ-only: $t_{(82)} = -0.95$, $p_{bonf} > .999$, $d = -0.10$; ACQ+EXT: $t_{(82)} = -2.29$, $p_{bonf} = .647$, $d = -0.25$), and ACQ+EXT participants rated CS+ as unpleasant as ACQ-only participants at Acq1 ($t_{(82)} = -0.02$, $p_{bonf} > .999$, $d < 0.01$) and Acq2 ($t_{(82)} = 0.40$, $p_{bonf} > .999$, $d = 0.04$). For details, see Table 2A.II, and Fig. 2.II.

Physiology. Regarding physiological responses to CSs, results indicated that over acquisition, participants had higher SCRs to CS+ than CS- ($t_{(74)} = 3.07$, $p_{bonf} < .005$, $d = 0.35$). Further, consistent with strong non-associative effects of habituation on psychophysiological measures (Vervliet, Baeyens, et al., 2013), results indicated a main effect of Time, suggesting that participants' SCRs declined from Acq1 to Acq2 ($t_{(74)} = -2.47$, $p_{bonf} = .016$, $d = -0.28$). There were no group differences in SCRs to CS- at Acq1 ($t_{(74)} = 0.18$, $p_{bonf} < .999$, *Cohens' d* = 0.02) and Acq2 ($t_{(74)} = -0.21$, $p_{bonf} < .999$, *Cohens' d* = −0.03), nor to CS+ at Acq1 ($t_{(74)} = -0.01$, $p_{bonf} < .999$, *Cohens' d* < 0.01) and Acq2 ($t_{(74)} = -0.38$, $p_{bonf} < .999$, *Cohens' d* = −0.04). Unlike self-report data, analyses did not suggest any CS-type × Time or CS-type × Time × Group interactions. For details, see Table2A.III and Fig. 2.III.

3.2.2.3. Extinction of conditioned responses (ACQ+EXT group). ***Self-report.*** Both US-expectancy and Valence rating analyses suggested CS-type × Time interactions: at Ext2, participants rated CS+ cues as less likely to be followed by the US ($t_{(57)} = -2.81$, $p_{bonf} < .001$, $d = -0.37$) and less unpleasant than at Acq2 ($t_{(57)} = -7.89$ $p_{bonf} < .001$, $d = -1.04$), suggesting successful extinction of conditioned responses to CS+. Conversely, CS- US-expectancy ($t_{(57)} = 0.29$, $p_{bonf} > .999$, $d = 0.04$) and Valence ratings ($t_{(57)} = 2.10$, $p_{bonf} > .999$, $d = 0.28$) did not differ from Acq2 to Ext2. For details, see Table 2B.I-II, and Fig. 2.I-II.

Physiology. SCR analyses suggested a main effect of Time, indicating that overall participants' SCR decreased from Acq1 to Ext2 ($t_{(51)} = -3.82$, $p_{bonf} = .002$, $d = 0.53$) and from Acq2 to Ext2

[3] UR valence and arousal ratings were missing for $N = 1$ participant; Analyses are thus based on $N = 83$ participants.

[4] Even though each participant pre-rated CSs on valence and arousal, data of three participants is missing at random (N = 2 in ACQ+EXT group, N = 1 in ACQ-only group).

Table 2
Differences in responding to CS+ and CS- during fear-acquisition, fear-extinction, and retention.

	A - Acquistion				B - Extinction				C - Retention			
	MS	$F_{(df1,\ df2)}$	*p*	η^2_p	MS	$F_{(df1,\ df2)}$	*p*	η^2_p	MS	$F_{(df1,\ df2)}$	*p*	η^2_p
I US-expectancy												
CS-type	2318.40	**$339.88_{(1,82)}$**	**< .001**	.806	3067.35	**$269.80_{(1,57)}$**	**< .001**	.826				
Time	4.57	**$4.33_{(1,82)}$**	**.041**	.050	80.35^{GG}	**$40.60^{GG}_{(2.29,130.77)}$**	**< .001^{GG}**	.416				
CS-type × Time	26.41	**$19.55_{(1,82)}$**	**< .001**	.193	57.51^{GG}	**$22.69^{GG}_{(2.12,119.98)}$**	**< .001^{GG}**	.285				
Group	1.22	$0.56_{(1,82)}$	.456	.007								
CS-type × Group	1.90	$0.28_{(1,82)}$	.599	.003								
Time × Group	0.12	$0.11_{(1,82)}$	.738	.001								
CS-type × Time × Group	0.17	$0.13_{(1,82)}$	.723	.002								
II Valence												
CS-type	684.84	**$78.45_{(1,82)}$**	**< .001**	.489	828.45	**$52.23_{(1,57)}$**	**< .001**	.478	426.22	**$57.68_{(1,79)}$**	**< .001**	.422
Time	0.17	$0.12_{(1,82)}$	.729	.001	15.74^{GG}	**$10.49^{GG}_{(2.42,137.87)}$**	**< .001^{GG}**	.155	25.69	**$13.90_{(1,79)}$**	**< .001**	.150
CS-type × Time	9.07	**$6.66_{(1,82)}$**	**.012**	.075	25.90^{GG}	**$16.99^{GG}_{(2.43,138.68)}$**	**< .001^{GG}**	.230	69.27	**$24.74_{(1,79)}$**	**< .001**	.239
Group	0.12	$0.03_{(1,82)}$	.868	<.001					0.95	$0.17_{(1,79)}$	.680	.002
CS-type × Group	0.18	$0.02_{(1,82)}$	.887	.000					0.06	$0.01_{(1,79)}$	.928	<.001
Time × Group	3.50	$2.51_{(1,82)}$	.117	.030					0.60	$0.33_{(1,79)}$	.570	.004
CS-type × Time × Group	7.22	**$5.30_{(1,82)}$**	**.024**	.061					6.80	$2.43_{(1,79)}$	.123	.030
III SCR												
CS-type	0.04	**$5.64_{(1,74)}$**	**.020**	.071	<0.01	$0.34_{(1,51)}$	.563	.007				
Time	0.05	**$6.99_{(1,74)}$**	**.010**	.086	0.11^{GG}	**$9.52^{GG}_{(2.01,102.39)}$**	**< .001^{GG}**	.157				
CS-type × Time	<.01	$0.60_{(1,74)}$	.442	.008	0.01^{GG}	$2.66^{GG}_{(2.85,145.38)}$	.054^{GG}	.049				
Group	<.01	$0.02_{(1,74)}$	.900	<.001								
CS-type × Group	<.01	$0.07_{(1,74)}$	.799	.001								
Time × Group	<.01	$0.27_{(1,74)}$	.603	.004								
CS-type × Time × Group	<.01	$<.001_{(1,74)}$	.989	<.001								

Note. *F*-statistics significantly different from zero are highlighted in bold. Acquisition analyses (a): self-report analyses based on $N = 84$ participants, SCR-analyses based on $N = 76$ participants. Extinction analyses (b): self-report analyses based on $N = 58$ participants, SCR-analyses based on $N = 52$ participants. Retention analyses (c): analyses based on $N = 79$ participants due to 5 missing values. Abbreviations: Group = ACQ-only, ACQ+EXT; CS-type = CS+, CS-; Time = first and second acquisition halves for Acquisition analyses (a), first and second acquisition and extinction halves for Extinction analyses (b), and second acquisition half and post assessments for retention analyses (c); *MS* = mean square; *df* = degrees of freedom; *GG* = Greenhouse-Geisser corrected values; SCR = skin conductance response; η^2_p = partial eta squared.

($t_{(51)} = -3.30$, $p_{bonf} = .011$, $d = 0.45$); this reduction of SCRs held for both CSs, as indicated by a non-significant CS-type × Time interaction. Possibly due to habituation from Acq1 to Acq2, post-hoc tests revealed no reduction of SCRs from Acq2 to Ext2. For details, see Table2B.II, and Fig. 2.III.

Long-term retention of conditioned responses (Valence). A CS-type x Time interaction suggested that whereas participants perceived CS+ cues as more unpleasant at Acq2 than at post-assessments four days after acquisition learning ($t_{(79)} = 6.21$, $p_{bonf} < .001$, $d = 0.69$), participants perceived CS- as similarly unpleasant at Acq2 and post assessments ($t_{(79)} = -1.51$, $p_{bonf} = .799$, $d = -0.17$). Most importantly though, participants still rated CS+ as more unpleasant than CS- at post assessments ($t_{(79)} = 3.86$, $p_{bonf} < .001$, $d = 0.43$), suggesting retention of differential CRs over testing days. Effects were independent of Group, as indicated by non-significant CS-type × Time × Group interactions. For details, see Table2C.II, and Fig. 2.II.

3.3. CSs as content of intrusive memories

Outcomes of intrusion content classification are detailed in Table 3 (frequency within categories) and Table 4 (associated distress). For better comprehension, we show raw frequency and distress values instead of the weighted product of frequency and distress. In the four days after the conditioned-intrusion paradigm, around two third of participants (77% [$N = 20$] in ACQ-only, 74% [$N = 43$] in ACQ+EXT group) reported at least one intrusive memory. As expected, participants experienced a number of intrusions referring to the USs (ACQ-only: $M = 2.19$, $SD = 1.86$; ACQ+EXT: $M = 1.47$, $SD = 1.86$), but also to the CSs (ACQ-only: $M = 0.92$, $SD = 1.76$; ACQ+EXT: $M = 0.72$, $SD = 1.32$).

Over both ACQ-only and ACQ+EXT participants, CSs were content of almost a third of all reported intrusive memories (see Table 3). Specifically, participants reported intrusions referring to the CS+ (ACQ-only: $M = 0.54$, $SD = 1.27$; ACQ+EXT: $M = 0.29$, $SD = 0.90$), but also to the CS- (ACQ-only: $M = 0.27$, $SD = 0.87$; ACQ+EXT: $M = 0.26$, $SD = 0.98$), and, to a lesser extent, to both CSs or to CSs that could not be clearly distinguished between CS+ and CS- (ACQ-only: $M = 0.12$, $SD = 0.33$; ACQ+EXT: $M = 0.17$, $SD = 0.50$). Exploratory analyses operationalizing intrusions as the product of number and average distress of intrusions indicated that participants reported similar amounts of CS+ and CS- intrusions ($b = 0.50$, 95%CI [-0.86, 1.92]), and this was independent of Group ($b = 0.30$, 95% CI [-2.90, 3.34]). Likewise, the probability of participants reporting intrusions was identical for CS+ and CS- categories ($b = -0.79$, 95% CI [–4.13, 3.58]), and again this was independent of Group ($b = 1.72$, 95% CI [-1.24, 5.54]).

The largest proportion of intrusions however referred to the USs (Table 3). Indeed, analyses suggested that the odds of participants experiencing US intrusions was almost four times greater than the odds of experiencing CS intrusions ($b = 3.82$, 95% CI [1.89, 8.08]). Classification outcomes showed that the vast majority of intrusions (n = 203) referred to aversive, as compared to neutral (n = 5, reported by four individuals) film-clips. In specific, participants reported several intrusions referring to the film clips portraying interpersonal trauma (ACQ-only: $M = 1.12$, $SD = 1.24$; ACQ+EXT: $M = 0.90$, $SD = 1.44$), followed by film clips portraying accidental trauma (ACQ-only: $M = 0.88$, $SD = 1.18$; ACQ+EXT: $M = 0.41$, $SD = 0.80$). Sporadically, participants also reported intrusions that referred to both film-clip types, unclassifiable film clips, or neutral films (ACQ-only: $M = 0.19$, $SD = 0.49$; ACQ+EXT: $M = 0.16$, $SD = 0.52$). Exploratory analyses indicated that participants reported increased intrusion severity (operationalized as the product of number and average distress of intrusions) of film clips portraying interpersonal trauma than intrusions of film clips portraying accidental trauma; ($b = 0.39$, 95%CI [-0.01, 0.78]). This effect of US intrusions mostly referring to interpersonal rather than accidental trauma held for both ACQ-only and ACQ+EXT participants, as indicated

Table 3
Frequency of intrusive memories per content category.

	ACQ-only		ACQ+EXT	
	Proportion of total intrusions ($n = 81$)	Proportion of participants ($n = 26$)	Proportion of total intrusions ($n = 127$)	Proportion of participants ($n = 58$)
US intrusions				
Interpersonal	36% ($n = 29$)	62% ($n = 16$)	41% ($n = 52$)	50% ($n = 29$)
Accidental	28% ($n = 23$)	54% ($n = 14$)	19% ($n = 24$)	28% ($n = 16$)
Mixed US[a]	6% ($n = 5$)	15% ($n = 4$)	7% ($n = 9$)	10% ($n = 6$)
CS intrusions				
CS+	17% ($n = 14$)	23% ($n = 6$)	13% ($n = 17$)	17% ($n = 10$)
CS-	9% ($n = 7$)	12% ($n = 3$)	12% ($n = 15$)	14% ($n = 8$)
Mixed-CSs[b]	4% ($n = 3$)	12% ($n = 3$)	8% ($n = 10$)	12% ($n = 7$)
Total				
US	70% ($n = 57$)	77% ($n = 20$)	67% ($n = 85$)	64% ($n = 37$)
CS	30% ($n = 24$)	38% ($n = 10$)	33% ($n = 42$)	41% ($n = 24$)

Note: Abbreviations: US = unconditioned stimuli; CS = conditioned stimuli; ACQ-only = acquisition-only group; ACQ+EXT = acquisition plus extinction group.

[a] The "mixed US" category includes intrusions referring to a combination of film clips including interpersonal and accidental trauma ($n = 4$), unclassifiable film clips ($n = 5$), or neutral film clips ($n = 5$).

[b] The "mixed CSs" category includes intrusions referring to both the CS+ and the CS- ($n = 3$), as well as intrusive memories not clearly assignable to CS+ or CS- ($n = 10$).

by a non-significant US-type (accidental/interpersonal) × Group interaction ($b = 0.17$, 95% CI [-0.63, 0.99]). Likewise, the probability of intrusion absence was lower for interpersonal-trauma than for accidental-trauma intrusions ($b = -0.77$, 95% CI [-1.56, 0.05]), though this result is associated with some uncertainty.

Lastly, most intrusions were perceived in multiple sensory modalities (44.5%), followed by visual (40.2%), thought (13.9%), feeling (1.0%), and auditory (0.5%) modalities.

3.4. CSs as reminder cues for intrusive memories

Outcomes of the cue content classification are detailed in Table 5. More than half of the participants who reported at least one intrusion identified a corresponding reminder cue (60% [$N = 12$] in ACQ-only, 56% [$N = 24$] in ACQ+EXT group).

Only a few participants of both groups identified CS-reminiscent cues (e.g., "seeing a similar face"; "woman on bike") as reminder cues for their intrusions. However, participants identified experimental-context-related cues (e.g., "switching on the study phone"; "cables") as well as US-related cues (e.g., "numb arm from laying on it, made me think of amputated arm"; „showering and seeing water on the ground") as intrusion reminder cues. Lastly, several cues were of unspecific content, as we could not clearly attribute them to elements resembling CSs, USs, or the experimental context. In what concerns the origin of reported reminder cues, most were of external origin (e.g., "hearing the word pain").

Altogether, although we identified some instances where participants reported reminder cues resembling CSs, these were only few, and thus results only weakly support our hypothesis that CSs occur as trigger to intrusions in daily life. Interestingly though, participants also identified cues that resembled elements of the US and the general experimental context. Implications of these findings will be discussed.

Table 4
Distress (0 = not distressing at all, 100 = extremely distressing) of intrusive memories per content category.

	ACQ-only		ACQ+EXT	
	M	(SD)	M	(SD)
US intrusions				
Interpersonal	51.97	(22.03)	49.88	(22.28)
Accidental	46.33	(26.50)	38.88	(25.96)
Mixed US	17.62	(10.39)	46.72	(26.65)
CS intrusions				
CS+	34.30	(24.00)	53.08	(35.04)
CS-	13.33	(13.77)	17.25	(15.48)
Mixed CSs	28.33	(33.47)	31	(22.69)
Total				
US	48.91	(24.08)	44.51	(24.55)
CS	29.63	(22.14)	30.60	(28.24)

Note: Descriptives for distress include participants reporting at least one intrusion within the respective category. Abbreviations: US = unconditioned stimuli; CS = conditioned stimuli; ACQ-only = acquisition-only group; ACQ+EXT = acquisition plus extinction group; M = mean; SD = standard deviation.

3.5. Extinction effects on US and CS intrusions

We estimated BMLMs to examine the effect of receiving extinction (ACQ+EXT) vs. not receiving extinction (ACQ-only) on US- and CS-intrusions, and the moderating effect of the degree to which

Table 5
Classification outcomes for cues reported in e-diaries over the four experimental days, separately for ACQ-only and ACQ+EXT participants reporting at least one intrusive memory.

	ACQ-only *(N = 20)*		ACQ+EXT *(N = 43)*	
	Proportion of intrusions with reminder cues (n = 24)	Proportion of ACQ-only participants (n = 20)	Proportion of intrusions with reminder cues (n = 44)	Proportion of ACQ+EXT participants (n = 43)
A. Content				
US	25% ($n = 6$)	25% ($n = 5$)	39% ($n = 17$)	26% ($n = 11$)
CS	13% ($n = 3$)	10% ($n = 2$)	7% ($n = 3$)	7% ($n = 3$)
Experimental context	17% ($n = 4$)	15% ($n = 3$)	20% ($n = 9$)	19% ($n = 8$)
Unspecific	46% ($n = 11$)	45% ($n = 9$)	34% ($n = 15$)	21% ($n = 9$)
B. Origin				
External	79% ($n = 19$)	50% ($n = 10$)	86% ($n = 38$)	51% ($n = 22$)
Internal	17% ($n = 4$)	20% ($n = 4$)	9% ($n = 4$)	9% ($n = 4$)
Mixed	4% ($n = 1$)	5% ($n = 1$)	5% ($n = 2$)	5% ($n = 2$)

Note: Descriptive and inferential statistics solely included participants reporting at least one intrusive memory. The "Unspecific" category includes stimuli not resembling the USs, CSs, or wider experimental context. Abbreviations: ACQ-only = acquisition-only group; ACQ+EXT = acquisition plus extinction group.

Table 6
Bayesian multilevel model predicting US intrusions by Group (ACQ-only/ACQ+EXT), Conditionability (indexed by (I) US-expectancy ratings, (II) Valence ratings, (III) SCRs), and Day.

	(A) Hurdle					(B) Lognormal						
	b	95% CI	*SDb*	95% CI	*PP*	*b*	95% CI	*SDb*	95% CI	*PP*	R^2	95% CI
I: Conditionability indexed by US-expectancy											0.31	[0.17, 0.45]
(Intercept)	0.53	[-0.13, 1.21]	1.09	[0.52, 1.69]		**4.40**	**[3.67, 4.41]**	0.70	[0.49, 0.92]			
Day	**0.46**	**[0.02, 0.93]**	0.29	[0.01, 0.78]	0.98	−0.07	[-0.29, 0.15]	0.21	[0.02, 0.44]	0.75		
Group	**0.86**	**[0.04, 1.72]**			0.98	−0.06	[-0.53, 0.42]			0.59		
$Conditionability_{US\text{-}EXP}$	−0.04	[-0.35, 0.25]				0.16	[-0.00, 0.33]					
Group × $Conditionability_{US\text{-}EXP}$	0.05	[-0.30, 0.41]			0.62	**−0.22**	**[-0.42, -0.01]**			0.98		
Group × Day	−0.36	[-0.92, 0.18]				−0.31	[-0.65, 0.01]					
$Conditionability_{US\text{-}EXP}$ × Day	0.16	[-0.03, 0.37]				0.07	[-0.02, 0.16]					
Group × $Conditionability_{US\text{-}EXP}$ × Day	−0.17	[-0.41, 0.07]				−0.07	[-0.21, 0.07]					
II: Conditionability indexed by Valence											0.31	[0.17, 0.45]
(Intercept)	0.51	[-0.14, 1.19]	1.07	[0.50, 1.67]		**4.01**	**[3.65, 4.39]**	0.70	[0.49, 0.93]			
Day	0.45	[-0.01, 0.94]	0.29	[0.01, 0.79]	0.97	−0.09	[-0.32, 0.14]	0.25	[0.03, 0.49]	0.78		
Group	**0.86**	**[0.06, 1.70]**			0.98	−0.02	[-0.50, 0.45]			0.54		
$Conditionability_{VAL}$	−0.04	[-0.26, 0.17]				0.11	[-0.02, 0.23]					
Group × $Conditionability_{VAL}$	0.10	[-0.16, 0.37]			0.78	−0.05	[-0.21, 0.11]			0.71		
Group × Day	−0.35	[-0.91, 0.20]				−0.30	[-0.65, 0.03]					
$Conditionability_{VAL}$ × Day	0.09	[-0.05, 0.24]				0.03	[-0.05, 0.11]					
Group × $Conditionability_{VAL}$ × Day	−0.10	[-0.29, 0.07]				−0.03	[-0.14, 0.09]					
III: Conditionability indexed by SCR											0.36	[0.22, 0.47]
(Intercept)	0.41	[-0.22, 1.07]	0.93	[0.28, 1.58]		**4.00**	**[3.57, 4.42]**	0.80	[0.59, 1.04]			
Day	0.41	[-0.05, 0.91]	0.34	[0.02, 0.87]	0.96	−0.14	[-0.05, 0.91]	0.14	[0.01, 0.38]	0.92		
Group	**0.90**	**[0.11, 1.76]**			0.99	−0.01	[-0.55, 0.53]			0.52		
$Conditionability_{SCR}$	−7.06	[-15.87, 0.72]				−0.64	[-4.94, 3.63]					
Group × $Conditionability_{SCR}$	**11.49**	**[2.06, 22.09]**			0.99	−0.37	[-5.84, 5.03]			0.55		
Group × Day	−0.20	[-0.79, 0.38]				**−0.38**	**[-0.71, -0.06]**					
$Conditionability_{SCR}$ × Day	1.83	[-3.87, 7.85]				0.56	[-1.28, 2.28]					
Group × $Conditionability_{SCR}$ × Day	1.25	[-6.03, 8.62]				**−3.30**	**[-5.91, -0.44]**					

Note. Models *I-II* based on $N = 84$; Model *III* based on $N = 76$ (ACQ-only: $N = 24$; ACQ+EXT: $N = 52$). Coefficients are considered significantly different from zero if the corresponding 95% CI does not contain zero, and are highlighted in bold. The model's hurdle part (A) predicted the probability of US-intrusion absence; The lognormal part of the model (B) predicted US-intrusion severity. Abbreviations: *b* = regression coefficient; CI = credible interval; *SDb* = standard deviation of the varying intercepts and slopes across subjects, higher estimates suggest that intrusions on the first day after analogue-trauma (Intercept) and their change over time (Day) varied substantially across participants; PP = posterior probability that the effect is in the expected direction; ACQ-only = acquisition-only participants; ACQ+EXT: acquisition plus extinction participants; US-EXP=US-expectancy; VAL = valence; SCR = skin conductance response.

participants acquired differential CRs to CS+ vs. CS- (i.e., Conditionability, indexed by (I) US-expectancy ratings [$Conditionability_{US\text{-}EXP}$], (II) Valence ratings [$Conditionability_{VAL}$], and (III) SCRs [$Conditionability_{SCR}$]). We also investigated the effect of Group and Conditionability on intrusion decay over the four experimental days. Regression coefficients and corresponding 95% CIs are given in Table 6 for models predicting US intrusions, and Table 7 for models predicting CS intrusions. Results are split in two parts: one (A) estimating the probability of not experiencing (i.e. zero) vs. experiencing (i.e. non-zero) intrusions (hurdle part: *probability of intrusion absence*); and another (B) estimating the severity of (i.e. non-zero) intrusions (lognormal part: *intrusion severity*).

Table 7
Bayesian multilevel model predicting CS intrusions by Group (ACQ-only/ACQ+EXT), Conditionability (indexed by (I) US-expectancy ratings, (II) Valence ratings, (III) SCRs), and Day.

	(A) Hurdle					**(B) Lognormal**						
	b	95% CI	*SDb*	95% CI	*PP*	*b*	95% CI	*SDb*	95% CI	*PP*	R^2	95% CI
I: Conditionability indexed by US-expectancy											0.40	[0.22, 0.53]
(Intercept)	**2.71**	**[0.68, 4.20]**	2.02	[1.16, 3.18]		**3.38**	**[2.78, 3.98]**	0.64	[0.17, 1.05]			
Day	0.38	[-0.40, 1.25]	0.49	[0.02, 1.23]	0.82	0.17	[-0.26, 0.58]	0.14	[0.01, 0.43]	0.20		
Group	0.23	[-1.21, 1.70]			0.63	−0.04	[-0.79, 0.69]			0.53		
$\text{Conditionability}_{\text{US-EXP}}$	**−0.30**	**[-0.90, -0.22]**				0.23	[-0.06, 0.53]					
Group × $\text{Conditionability}_{\text{US-EXP}}$	**0.63**	**[0.05, 1.31]**			0.98	**−0.32**	**[-0.64, -0.01]**			0.98		
Group × Day	−0.21	[-1.09, 0.65]				−0.22	[-0.73, 0.31]					
$\text{Conditionability}_{\text{US-EXP}}$ × Day	0.11	[-0.23, 0.45]				−0.17	[-0.44, 0.10]					
Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day	−0.18	[-0.55, 0.19]				0.19	[-0.09, 0.48]					
II: Conditionability indexed by valence											0.40	[0.21, 0.56]
(Intercept)	**2.79**	**[1.55, 4.38]**	2.16	[1.27, 3.36]		**3.44**	**[3.65, 4.39]**	0.64	[0.12, 1.10]			
Day	0.41	[-0.42, 1.31]	0.48	[0.02, 1.23]	0.84	−0.15	[-0.64, 0.35]	0.15	[0.01, 0.46]	0.73		
Group	0.16	[-1.39, 1.69]			0.59	0.02	[-0.75, 0.75]			0.48		
$\text{Conditionability}_{\text{VAL}}$	−0.26	[-0.70, 0.15]				0.09	[-0.16, 0.34]					
Group × $\text{Conditionability}_{\text{VAL}}$	0.37	[-0.11, 0.89]			0.94	−0.12	[-0.39, 0.16]			0.80		
Group × Day	−0.26	[-1.16, 0.61]				0.05	[-0.54, 0.64]					
$\text{Conditionability}_{\text{VAL}}$ × Day	0.12	[-0.40, 0.14]				0.09	[-0.11, 0.28]					
Group × $\text{Conditionability}_{\text{VAL}}$ × Day	0.08	[-0.21, 0.39]				−0.09	[-0.30, 0.13]			0.79		
III: Conditionability indexed by SCR											0.40	[0.20, 0.59]
(Intercept)	**2.65**	**[1.42, 4.19]**	2.09	[1.20, 3.36]		**3.51**	**[2.88, 4.13]**	0.63	[0.10, 1.09]			
Day	0.36	[-0.45, 1.25]	0.50	[0.02, 1.27]	0.81	−0.06	[-0.50, 0.40]	0.17	[0.01, 0.50]	0.61		
Group	0.09	[-1.44, 1.61]			0.55	−0.03	[-0.82, 0.72]			0.53		
$\text{Conditionability}_{\text{SCR}}$	−12.60	[-30.49, 2.08]				1.03	[-4.80, 6.77]					
Group × $\text{Conditionability}_{\text{SCR}}$	15.89	[-1.09, 36.21]			0.97	−0.52	[-7.29, 6.27]			0.57		
Group × Day	−0.20	[-1.11, 0.66]				−0.04	[-0.60, 0.49]					
$\text{Conditionability}_{\text{SCR}}$ × Day	2.18	[-6.70, 11.86]				1.36	[-1.89, 4.60]					
Group × $\text{Conditionability}_{\text{SCR}}$ × Day	−4.32	[-15.12, 5.94]				−1.84	[-6.32, 2.50]					

Note. Models *I-II* based on $N = 84$; Model *III* based on $N = 76$ (ACQ-only: $N = 24$; ACQ+EXT: $N = 52$). Coefficients are considered significantly different from zero if the corresponding 95% CI does not contain zero, and are highlighted in bold. The model's hurdle part (A) predicted the probability of US-intrusion absence; The lognormal part of the model (B) predicted US-intrusion severity. Abbreviations: *b* = regression coefficient; CI = credible interval; *SDb* = standard deviation of the varying intercepts and slopes across subjects, higher estimates suggest that intrusions on the first day after analogue-(Intercept) and their change over time (Day) varied substantially across participants; PP = posterior probability that the effect is in the expected direction; ACQ-only = acquisition-only participants; ACQ+EXT: acquisition plus extinction participants; US-EXP=US-expectancy; VAL = valence; SCR = skin conductance response.

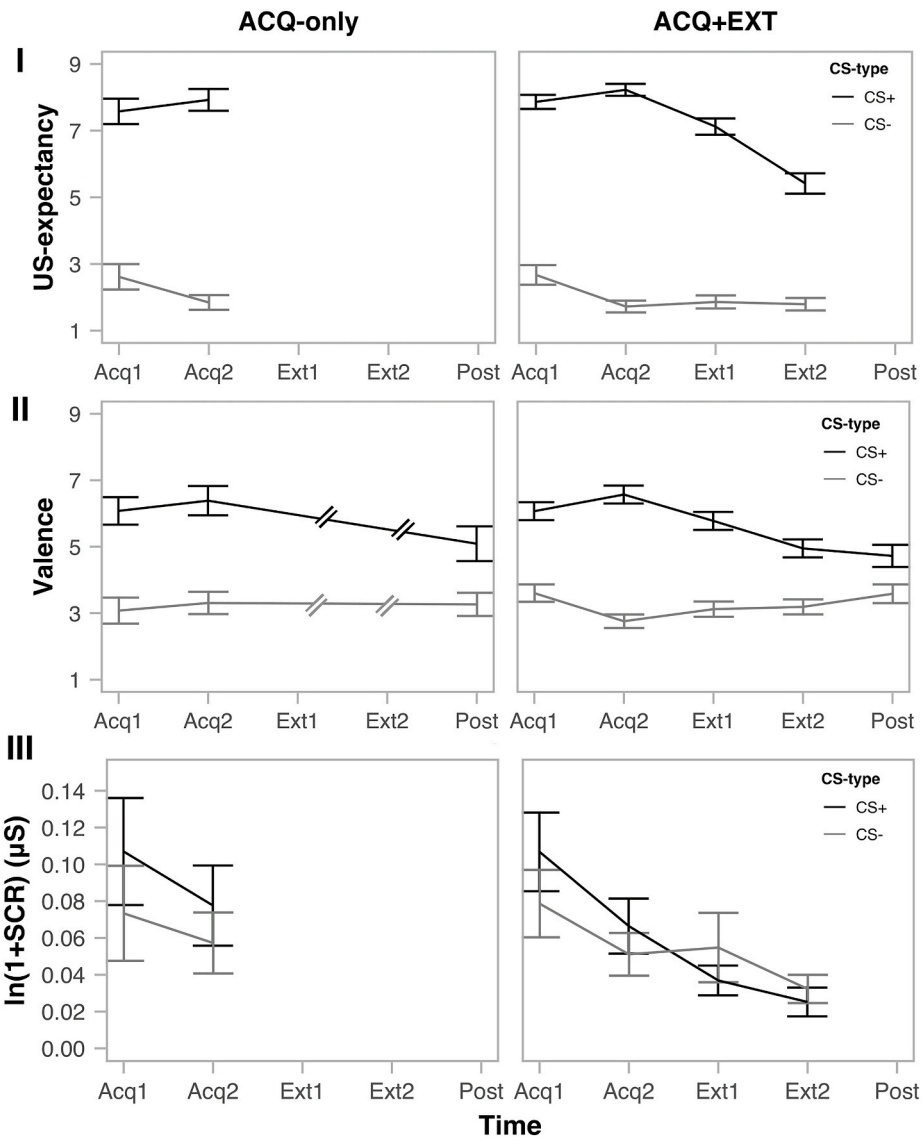

Fig. 2. US-expectancy (I) and Valence (II) ratings, as well as SCR (III) to CS+ vs. CS- over the conditioning procedure, plotted separately for ACQ-only (left panel) and ACQ+EXT participants (right panel). Rating scales for US-expectancy ranged from 1 (*surely not*) to 9 (*surely yes*) and for Valence ranged from 1 (*not unpleasant at all*) to 9 (*very unpleasant*). Error bars represent standard errors of the mean. Abbreviations: ACQ-only = acquisition-only group; ACQ+EXT = acquisition plus extinction group; SCR = skin conductance response; ln = normalized by natural logarithm transformation; µS = microSiemens; Acq1 = first acquisition halve; Acq2 = second acquisition halve; Ext1 = first extinction halve; Ext2 = second extinction halve; Post = five days after analogue-trauma.

3.5.1. *US intrusions*

Analyses estimating the *probability of intrusion absence* suggested that ACQ+EXT participants showed a higher probability of US-intrusion absence following analogue-trauma than ACQ-only participants: This main effect of Group was present in all three models regardless of how Conditionability was indexed ($b_{A.I} = 0.86$, 95%CI [0.04, 1.72]; $b_{A.II} = 0.86$, 95%CI [0.06, 1.70]; $b_{A.III} = 0.90$, 95%CI [0.11, 1.76]). Results revealed no substantial main effects of Conditionability on US-intrusion absence (Table 6A.I-III). However, when Conditionability was indexed by SCR (Table 6A.III), the effect of Group on the probability of US-intrusion absence was moderated by participants' Conditionability ($b_{A.I} = 11.49$, 95%CI [2.06, 22.09]): As can be seen from Fig. 3A. III, while ACQ-only participants with higher $\text{Conditionability}_{SCR}$ displayed a lower probability of US-intrusion absence, ACQ+EXT participants with higher $\text{Conditionability}_{SCR}$ showed a higher probability of US-intrusion absence. When Conditionability was indexed by US-expectancy (Table 6A.I, Fig. 3A. I) or Valence ratings (Table 6A.II, Fig. 3A.II), results suggested no such moderation effects of Conditionability on the relationship between Group and the probability of US-intrusion absence.

Results of analyses estimating the *severity* of US intrusions were too weak to support a main effect of Group or Conditionability (Table 6BI-III). However, when Conditionability was indexed by US-expectancy (Table 6B.I) results indicated that Conditionability moderated the effect of Group on US-intrusion severity ($b_{B.I} = -0.22$, 95%CI [-0.42, −0.01]): As shown in Fig. 3B·I, whereas ACQ-only participants with higher $\text{Conditionability}_{US\text{-}EXP}$ reported more severe US intrusions, ACQ+EXT participants with higher $\text{Conditionability}_{US\text{-}EXP}$ reported less severe US intrusions. When Conditionability was indexed by Valence ratings (Table 6B.II, Fig. 3B.II) or SCR (Table 6B.III, Fig. 3BIII), moderation effects of Conditionability on the relationship between Group and US-intrusion severity yielded approximately the same directions as when indexed by US-expectancy, yet as confidence intervals included zero, effects are associated with large uncertainty.[5]

Regarding the expected US-intrusion decay over the four experimental days, results suggested an increasing probability of *US-intrusion absence* over days, but this was only significant when Conditionability was indexed by US-expectancy ($b_{A.I} = 0.46$, 95%CI [0.02, 0.93]). The same pattern of results, though associated with larger uncertainty, emerged when Conditionability was indexed by Valence (Table 6A.II) or SCR (Table 6A.III). Analyses suggested no interaction effects between

[5] Expected means for whole models, which are a function of both zero (hurdle) and non-zero (lognormal) parts of the models (i.e., US intrusions predicted by Group and Conditionability including both zero and non-zero responses) are depicted in Fig. F·I in Appendix F.

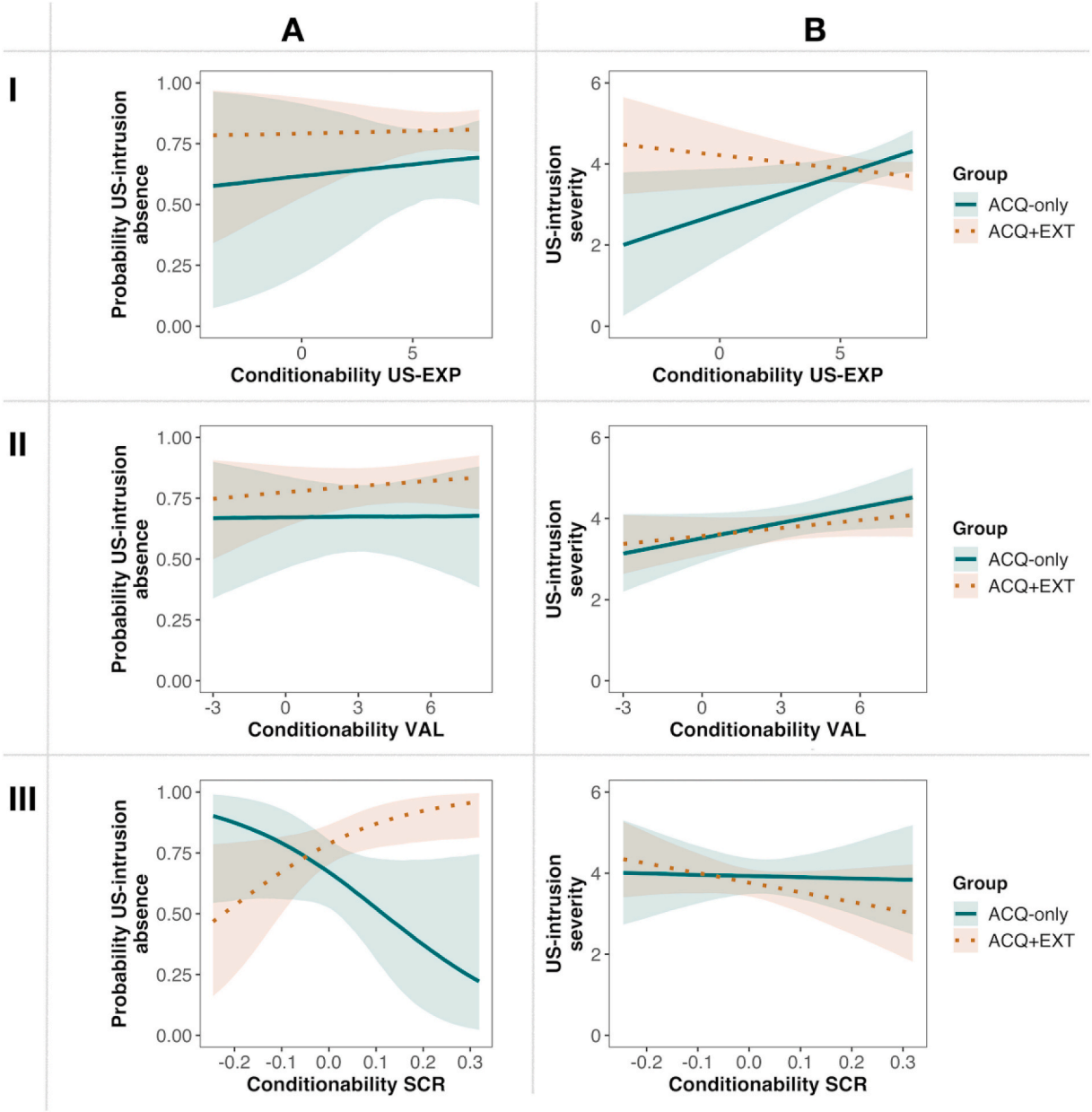

Fig. 3. Solid lines depict fitted values of the US intrusions, Group, and Conditionability regressions. Shaded areas represent 95% credible intervals. Left panels (A) depict the probability of *US-intrusion absence* (estimated by the model's hurdle part; higher values indicate greater probability of US-intrusion absence), predicted by Group and $Conditionability_{US\text{-}EXP}$ (I), $Conditionability_{VAL}$, (II), and $Conditionability_{SCR}$ (III). Right panels (B) depict *US-intrusion severity* (estimated by the model's lognormal part; higher values indicate greater US-intrusion severity) predicted by $Conditionability_{US\text{-}EXP}$ (I), $Conditionability_{VAL}$, (II), and $Conditionability_{SCR}$ (III). For illustrative purposes and better appreciation, plots depict non-mean-centered Conditionability estimates. Abbreviations: ACQ-only = acquisition-only participants; ACQ+EXT = acquisition plus extinction participants; US-EXP = US-expectancy; VAL = Valence; SCR = skin conductance response; US = unconditioned stimuli.

Group, Conditionability, and Day on the probability of US-intrusion absence (Table 6A.I-III, Fig. 4A.I-III). In what concerns the decay of *US-intrusion severity* over days, evidence is uncertain but also points towards the expected decay of US-intrusion severity over days (Table 6B.I-III). Further, when Conditionability was indexed by SCR (Table 6B.III) analyses indicated a Group × Day interaction ($b_{B.III} = -0.38$, 95%CI [-0.71, −0.06]), which was further modulated by $Conditionability_{SCR}$, as indicated by the Group × $Conditionability_{SCR}$ × Day interaction ($b_{B.III} = -3.30$, 95%CI [-5.91, −0.44]). As depicted in Fig. 4B.III, ACQ+EXT participants exhibited a more accelerated decay of US-intrusion severity than ACQ-only participants, but only when displaying higher $Conditionability_{SCR}$. Again, the effects of Group × $Conditionability_{US\text{-}EXP}$ × Day (Table 6B.I, Fig. 4B·I) and Group × $Conditionability_{VAL}$ × Day (Table 6B.II, Fig.B·II) on US-intrusion severity yielded the same direction, but were weaker and associated with larger uncertainty, as indicated by wider CIs.[6]

Together, results regarding the estimation of US intrusions suggested that ACQ+EXT participants, in comparison to ACQ-only participants showed an overall higher probability of *US-intrusion absence*; in part, this effect of extinction was stronger in participants displaying higher physiological (SCR) conditionability. Further, ACQ+EXT participants

[6] Expected means for whole models (i.e., US intrusions predicted by Day, Group, and Conditionability including both zero and non-zero responses) are depicted in Fig.F·II in Appendix F.

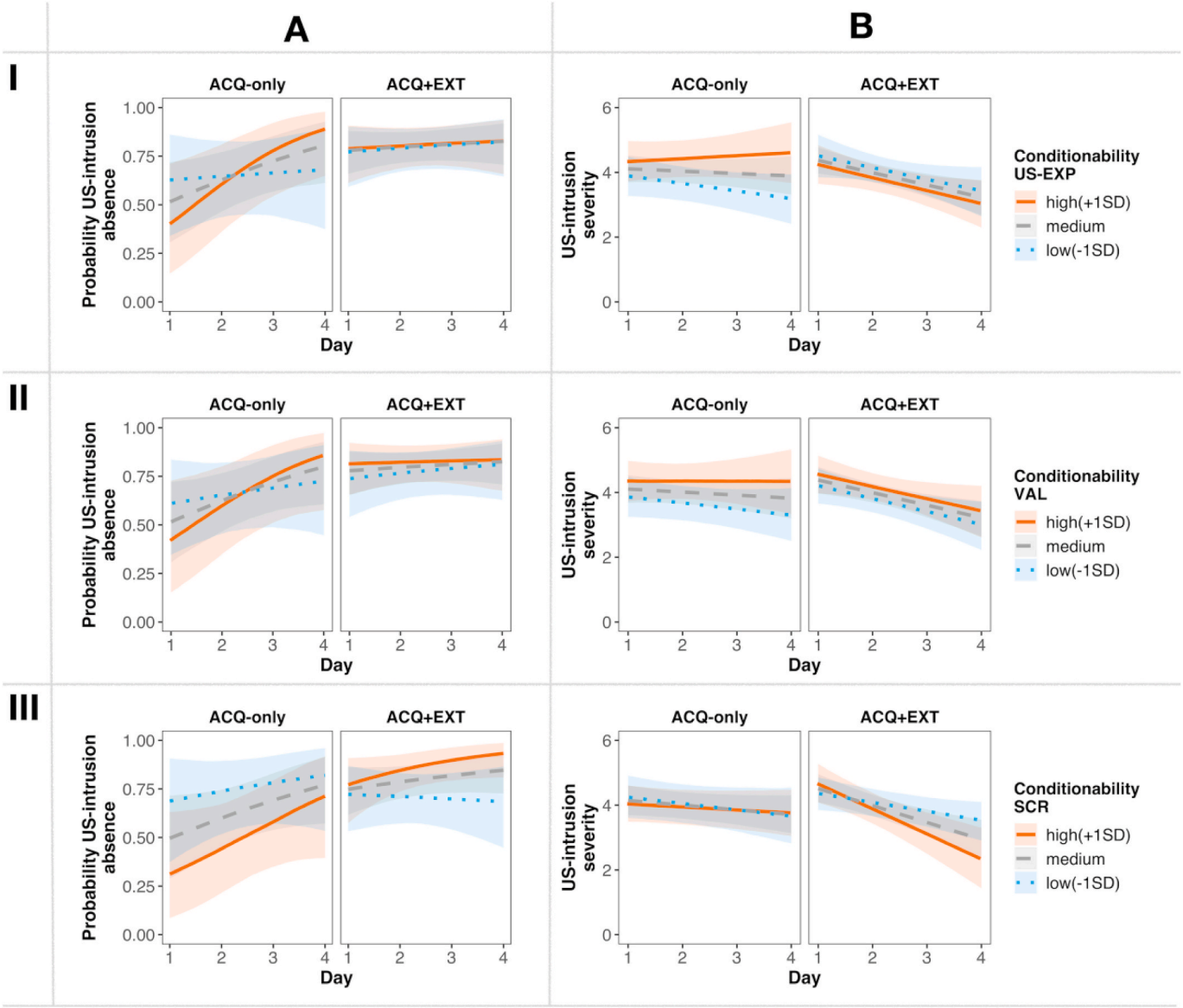

Fig. 4. Solid lines depict fitted values of the US intrusions, Group, and Conditionability regressions. Shaded areas represent 95% credible intervals. Left panels (A) depict the probability of *US-intrusion absence* (estimated by the model's hurdle part; higher values indicate greater probability of US-intrusion absence), predicted by Day, Group and $Conditionability_{US\text{-}EXP}$ (I), $Conditionability_{Val}$, (II), and $Conditionability_{SCR}$ (III). Right panels (B) depict *US-intrusion severity* (estimated by the model's lognormal part; higher values indicate greater US-intrusion severity) predicted by Day, Group, and $Conditionability_{US\text{-}EXP}$ (I), $Conditionability_{VAL}$, (II), and $Conditionability_{SCR}$ (III). Abbreviations: ACQ-only = acquisition-only participants; ACQ+EXT = acquisition plus extinction participants; US-EXP = US-expectancy; VAL = Valence; SCR = skin conductance response; US = unconditioned stimuli.

showed a lower *US-intrusion severity* when displaying higher cognitive (US-expectancy) conditionability. Lastly, ACQ+EXT participants exhibited an accelerated decay of *US-intrusion severity* over experimental days, but only when displaying higher physiological conditionability.

3.5.2. CS intrusions

Regarding the probability of *CS-intrusion absence*, analyses showed no main effect of Group (Table7A.I-III). Instead, when Conditionability was indexed by US-expectancy, results suggested a main effect of $Conditionability_{US\text{-}EXP}$ ($b_{A.I} = -0.30$, 95%CI [-0.90, −0.22]), where participants displaying higher $Conditionability_{US\text{-}EXP}$ at end of fear acquisition had a lower probability of CS-intrusion absence. This effect was further modulated by Group, as indicated by a Group × $Conditionability_{US\text{-}EXP}$ interaction ($b_{A.I} = 0.63$, 95%CI [0.05, 1.31]). As can be seen from Fig. 5A. I, whereas ACQ+EXT participants with higher $Conditionability_{US\text{-}EXP}$ shower a higher probability of CS-intrusion absence, ACQ-only participants showed a slightly lower probability of CS-intrusion absence when displaying higher $Conditionability_{US\text{-}EXP}$. This pattern became more apparent in the lognormal part of the model, when estimating the *severity* of experienced CS intrusions ($b_{B.I} = -0.32$, 95%CI [-0.64, −0.01]): while ACQ+EXT participants with higher $Conditionability_{US\text{-}EXP}$ reported less severe CS intrusions, ACQ-only participants with higher $Conditionability_{US\text{-}EXP}$ reported more severe CS intrusions. Note that although this pattern of effects was also visible when Conditionability was indexed by Valence (Table 7B.II, Fig. 5B·II) and SCRs (Table 7B.III, Fig. 5B.III), effects were weaker and associated with larger uncertainty as respective credible intervals contained zero.[7]

Similar to the results concerning US intrusions, models predicting CS

[7] Expected means for whole models, which are a function of both zero (hurdle) and non-zero (lognormal) parts of the models (i.e., CS intrusions predicted by Group and Conditionability including both zero and non-zero responses) are depicted in Fig. F·II in Appendix F.

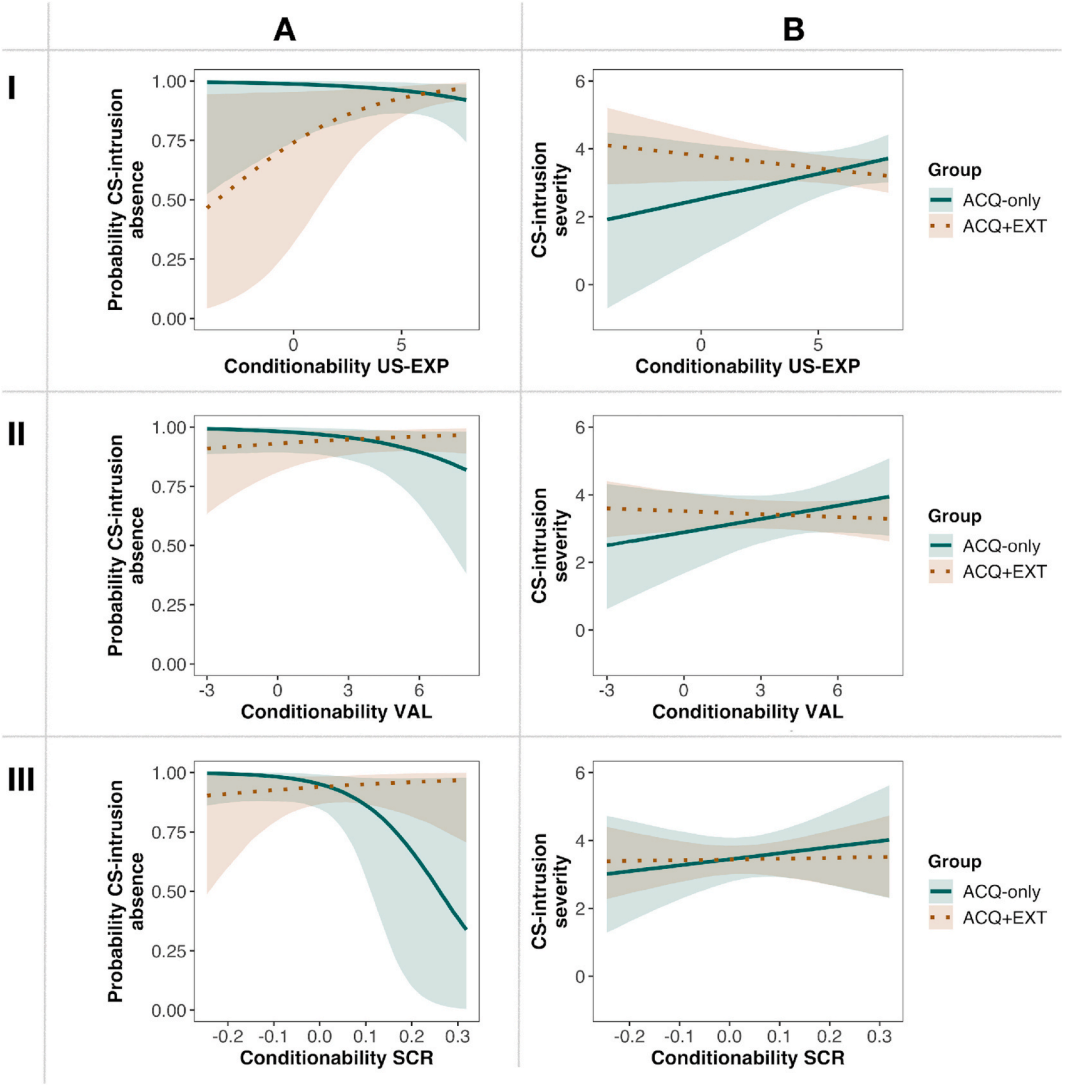

Fig. 5. Solid lines depict fitted values of the CS intrusions, Group, and Conditionability regressions. Shaded areas represent 95% credible intervals. Left panels (A) depict the probability of CS-intrusion absence (estimated by the model's hurdle part; higher values indicate greater probability of CS-intrusion absence), predicted by Group and $Conditionability_{US\text{-}EXP}$ (I), $Conditionability_{VAL}$, (II), and $Conditionability_{SCR}$ (III). Right panels (B) depict CS-intrusion severity (estimated by the model's lognormal part; higher values indicate greater CS-intrusion severity) predicted by predicted by Group and $Conditionability_{US\text{-}EXP}$ (I), $Conditionability_{VAL}$, (II), and $Conditionability_{SCR}$ (III). For illustrative purposes and better appreciation, plots depict non-mean-centered Conditionability estimates. Abbreviations: ACQ-only = acquisition-only participants; ACQ+EXT = acquisition plus extinction participants; US-EXP = US-expectancy; VAL = Valence; SCR = skin conductance response; CS = conditioned stimuli.

intrusions revealed only weak, uncertain effects for CS-intrusion decay over experimental days, suggesting an increased probability of *CS-intrusion absence* over days (Table 7A.I-III), as well as a decreasing *severity* of CS intrusions over days (Table 7B.I-III). However, here the decay of CS intrusions over days was not modulated by Group nor Group × Conditionability.

Overall, analyses predicting CS intrusions largely replicated the pattern of results from analyses predicting US intrusions, albeit effects were generally weaker and associated with larger uncertainty. As such, results only modestly support our hypothesis by suggesting that ACQ+EXT participants with higher cognitive (US-expectancy) Conditionability showed (1) higher probability of *CS-intrusion absence* and (2) less *severe* CS intrusions in the aftermath of analogue-trauma than ACQ-only participants.

4. Discussion

The current study aimed to investigate the role of conditioning processes in the development of intrusive memories. Results largely supported the general hypothesis that intrusions can be successfully conceptualized within a conditioning framework since: (1) CSs were re-experienced as intrusions and (2) occasionally served as reminder cues for intrusions; and (3) ACQ+EXT participants were overall less likely to experiencing US intrusions and showed a steeper decay of US intrusions than ACQ-only participants. In part, these extinction effects were greater in participants showing higher physiological (SCR) and cognitive (US-expectancy) differential responses to CS+ vs. CS- by the end of fear acquisition (i.e., higher conditionability). ACQ+EXT participants only experienced reduced CS intrusions when displaying high cognitive conditionability at the end of acquisition.

4.1. CSs as content of intrusive memories

About one third of all reported intrusions in daily life were about CSs, i.e., cues that signaled aversive film clips (USs) or neutral film-clips during the conditioning procedure. As far as we know, this is the first experimental study showing that previously neutral stimuli, after being temporally associated with an aversive event, re-appear as intrusive

memories during daily life. This finding concurs well with patient studies showing that primarily stimuli signaling the onset of the traumatic event, the so called warning signals of the traumatic event, are re-experienced as intrusive memories (Ehlers et al., 2002; Hackmann, Ehlers, Speckens, & Clark, 2004). Similarly, also reports from PTSD-treatment-seekers suggested that the majority of intrusions referred to hotspots (i.e., worst moments of the traumatic memory) that mainly related to themes representing a general threat of injury and death. For example, one hotspot was about a scaffold smashing a car window during an accident, a moment when the victim thought "I'll be decapitated" (Grey & Holmes, 2008; Holmes, Grey, & Young, 2005), and thereby signaled something bad that was yet about to happen, rather than the incident per se (i.e., the car accident).

Interestingly, we observed that participants' intrusions did not only refer to the conditioned danger cue (CS+), but to some extent also to the conditioned safety cue (CS-), possibly reflecting impaired inhibition of fear responding (Davis, Falls, & Gewirtz, 2000) but, at least in part, also transfer (or generalization) of conditioned fear from the learned danger cue to a resembling stimulus (i.e., with shared features regarding for instance category, shape, size, duration). Indeed, enhanced responding to conditioned safety cues has been shown to be a robust conditioning correlate (Duits et al., 2015; Lissek et al., 2009, 2005; Lissek & van Meurs, 2015) and predictor (Craske et al., 2012) of anxiety-related disorders. As such, considering that CRs to conditioned safety cues seem to be an intrinsic part of anxiety-related disorders, it is not surprising that individuals also re-experienced CS- as intrusive memories. In fact, this suggests that intrusions may constitute CRs not only to stimuli that were directly associated with the traumatic event, but also to stimuli sharing features with the danger cue (here: a similar face; but note that the current study did not include a similarity gradient for more than two stimuli to unambiguously prove this; testing this hypothesis with appropriate generalization stimuli would be an important endeavor for future research).

The majority of reported intrusions referred to USs. In particular, US intrusions mostly alluded to film clips portraying interpersonal trauma, less to film clips portraying accidental trauma. This side finding bridges analogue trauma-film research with a consistent body of clinical evidence showing that interpersonal trauma confers greater risk for the development of PTSD symptomatology than non-interpersonal trauma (Breslau, 2001; Michopoulos et al., 2019). At first sight, the predominance of US intrusions relative to CS intrusions in the current sample may seem to stand in contrast with PTSD patient studies where intrusions mostly represented stimuli that were present shortly before the worst moment (Ehlers et al., 2002; Hackmann et al., 2004). Within the warning-signal hypothesis, through temporal association with the traumatic event, these stimuli became predictors of the event, and thereby induce a sense of serious current threat (Ehlers et al., 2002). An inherent difference between our analogue-trauma study and patient studies relates to the degree of threat experienced during the traumatic event. While highly aversive, our analogue-trauma situation never endangered participants. As such, although our CS+ elicited stronger CRs than the CS- during fear acquisition and participants still evaluated the CS+ as more unpleasant than the CS- on the fifth experimental day, the strength of our CRs was obviously far away from reaching the magnitude of CRs elicited after real-life trauma where, for instance, approaching footsteps signal impending life-threatening torture or abuse. In this way, on the one hand, it does not seem surprising that only a small proportion of intrusions referred to CSs. On the other hand, the fact that even with relatively "harmless" USs, approximately a third of all intrusions in this study did refer to CSs seems remarkable and highlights the importance of associative learning processes during trauma as potential mechanism underlying the nature of intrusive memories.

A reason that may have contributed to the fact that we observed a non-negligible proportion of CS intrusions may relate to our choice of using naturalistic neutral faces, rather than for instance geometric figures, as CSs. In line with evolutionary accounts, faces pertain to a cluster of stimuli that may be highly fear-relevant (Desimone, 1991; Rolls, 1992). In specific, faces can provide information that is vital for a rapid defense recruitment in case of danger (e.g., angry and threatening look on perpetrator's face just before attack elicits fear, but also defensive behaviors that may prove to be essential for survival; Öhman & Mineka, 2001). As such, faces seem to constitute phylogenetically relevant and salient warning signals of danger. Some PTSD patients' reports are indeed in line with this idea of faces constituting salient sources of information of danger vs. safety, and thus also content of intrusive memories. For instance, one patient reported that a particular expression on his assailant's face made him realize that he would never see his wife again; this look on his assailant's face was subsequently content of his intrusions (Hackmann et al., 2004). Thus, there may have been "evolutionary preparedness" that aided our participants in forming a faster and/or more durable association between face CS-US that, in the current study, not only translated in cognitive, evaluative, and physiological CRs, but also in, given the relatively mild US, a number of intrusive-image-CRs.

4.2. CSs as reminder cues for intrusive memories

Participants recognized a reminder cue for almost half of their intrusions. Although some of these cues indeed referred to stimuli resembling CSs, they mostly referred to an array of other stimuli, including elements resembling USs or the wider experimental context. The finding that cues beyond experimentally introduced CSs can trigger intrusions extends previous studies examining intrusions in response to explicit experimental CSs (Rattel, Wegerer, et al., 2019; Streb et al., 2017; Wegerer, Blechert, Kerschbaum, et al., 2013). Although only the CS+ was experimentally paired with USs, from the participants' perspective a broader range of cues was spatiotemporally associated with the aversive elements of the film clips. For instance, within the aversive film clips participants also observed neutral items such as water on a bathroom floor and tools, and were exposed to loud fMRI noises and electric skin conductance cables within the wider experimental context. Since these cues were effectively present during analogue-trauma and thus predicted the occurrence of particularly aversive events within the film clips, they likely also functioned as CSs. Subsequently, cues resembling these CSs, such as one's own bathroom or loud noises, may have triggered intrusions in daily life. Noteworthy, a considerable part of reported reminder cues was not attributable to either CSs, USs, or wider experimental context. Some cues may thus have resulted from individuals generalizing their fear to novel stimuli resembling the CS and/or pairing new stimuli to the CS (i.e., second-order conditioning), which then became able to evoke emotional responses (Keane, Zimering, & Caddell, 1985; Lissek & van Meurs, 2015). Potentially, these processes also partially account for the absence of reported cues for about half of intrusions. The inability to recognize reminder cues for intrusive memories agrees with clinical observations that PTSD patients often fail to spot triggers, as intrusions seem to "come out of the blue" (Ehlers & Clark, 2000; Ehlers, Hackmann, & Michael, 2004). While it is certainly possible that some intrusions appeared without an eliciting cue, the spread of cues beyond the original CSs likely occurs without the individual's awareness, and thus makes it harder for the individual to recognize reminder cues. Concomitantly, as Ehlers and Clark (2000) suggest, there may be particularly strong perceptual priming for stimuli that were spatiotemporally associated with the traumatic event, resulting in a reduced perceptual threshold for these stimuli. Consequently, such cues often remain unnoticed. Altogether, our results support contemporary conceptualizations of PTSD in suggesting that, via temporal and spatial contiguity, individuals' emotional responses became conditioned to a wide array of stimuli present before (CSs, experimental context) or during (stimuli resembling elements of the aversive film clips) the analogue-trauma (Charney, Deutch, Krystal, Southwick, & Davis, 1993; Ehlers & Clark, 2000; Ehlers et al., 2002; Foa et al., 1989; Keane et al., 1985). We argue that these stimuli, by signaling the

occurrence of an aversive event, acquired the ability to activate the "analogue-trauma" memory in the form of intrusions. Overarchingly, our findings also support other theories assuming a cue-driven nature of involuntary retrieval (Berntsen, 2009; Brewin et al., 2010; Conway, 2001).

4.3. Extinction effects on intrusive memories

Receiving immediate extinction reduced the probability and severity of US intrusions and accelerated their decay after analogue-trauma. In part, these extinction effects were stronger in participants who physiologically (SCR) and cognitively (US-expectancy) differentiated better between CS+ and CS-, i.e., showed higher conditionability by the end of fear acquisition. More precisely, extinction reduced the probability of experiencing US intrusions particularly in participants showing high physiological conditionability, and reduced US-intrusion severity in participants showing higher cognitive conditionability. Further, extinction accelerated the decay of US-intrusion severity in participants with higher physiological conditionability. Results indicated the reversed pattern for ACQ-only participants, for whom higher physiological conditionability increased the probability of experiencing US intrusions; and a higher cognitive conditionability increased US-intrusion severity. Also, in ACQ-only participants a higher physiological conditionability slowed down the decay of US-intrusion severity. With regard to CS intrusions, in participants with higher cognitive conditionability, receiving extinction reduced the probability of experiencing CS intrusions and reduced CS-intrusion severity. We elaborate on these findings below.

Extinction effects on intrusion formation may be explained by extinction training forming a new inhibitory CS-noUS association that suppresses the original CS-US association (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014), subsequently weakening CRs to CSs. Considering the conceptualization of intrusive memories as CRs to CSs, it follows that by inhibiting the original CS-US association one would also reduce intrusions. Importantly, extinction effects were nuanced by conditionability levels, a pattern that appears to match previous observations of fear extinction studies where extinction training was more effective (Culver et al., 2015) and more durable (Struyf et al., 2018) when participants showed higher US-expectancy at the start of extinction. In line with the Rescorla-Wagner model (Rescorla & Wagner, 1972), the authors argued that higher US-expectancy at the start of extinction enhanced learning by maximizing prediction errors, i.e., the discrepancy between what was expected to occur (US following CS) and what actually occurred (no US following CS). According to these principles of prediction-error learning, ACQ+EXT participants who were better at discriminating between CS+ and CS- by the end of fear-acquisition may have been more 'surprised' upon US absence, and these larger prediction errors may have potentiated extinction learning.

The more recent latent cause model however challenges the view that larger prediction errors necessarily lead to better extinction learning, let alone to unlearning of the original CS-US association (Dunsmoor, Niv, Daw, & Phelps, 2015; Gershman, Monfils, Norman, & Niv, 2017). Instead, according to this model, prediction errors need to hit a 'sweet spot' where they are just large enough to cause discrepancy in what was expected and what actually occurred so that the original CS-US association is weakened, but small enough to avoid that participants partition acquisition and extinction trials into statistically distinct clusters (i.e., infer distinct 'latent causes'). This idea of inferring 'causes' for observations (here, CSs, USs configurations) lays at the heart of the latent cause model and defines whether individuals will update an existing memory or create a new one. In the case of US absence in extinction, individuals try to infer what caused the surprising event: If a new latent cause is inferred for this experience, a new CS-noUS memory is formed; however, if the latent cause of the old experience (from acquisition) is inferred, the original CS-US memory is updated. Whether a new latent cause is inferred is influenced by the aforementioned prediction errors (Gershman et al., 2017; Gershman & Niv, 2012). In our data, a factor that may have contributed to prediction errors not being too large was our 75% reinforcement rate at acquisition, which may have rendered extinction not too discrepant from acquisition. In this way, the first extinction trials may not have surprised participants enough to partition these trials to a distinct cluster, but instead constituted intervals where participants contemplated about US absence, resulting in gradual weakening of the original CS-US association. Moreover, the fact that extinction trials immediately followed acquisition trials may have 'glued' these trials together as one entity, for which a common, rather than a distinct latent cause would be more likely (contiguity principle, Gershman et al., 2017). It would be interesting for future studies to use computational modelling methods to investigate whether individuals who are more likely to create new inhibitory memories instead of updating the original fear memory are indeed more likely to showing more intrusive memories.

Critically, but also in line with our expectations, this pattern of higher cognitive and physiological conditionability during acquisition functioning as a "resilience factor" for intrusive memory formation in ACQ+EXT participants reversed in ACQ-only participants. More specifically, while greater conditionability at the end of fear-acquisition (likely reflecting stronger original CS-US associations; Ehlers & Clark, 2000; Orr et al., 2000) may have potentiated prediction-error driven extinction learning in those participants who were actually given the opportunity to undergo extinction (ACQ+EXT participants) on the one hand, on the other hand may have left participants who did not undergo extinction (ACQ-only participants) with stronger original CS-US associations that facilitated unintentional cue-driven retrieval of the event. Noteworthy, although the overall pattern in this study suggested a negative association between conditionability at the end of fear acquisition and intrusion formation in participants receiving extinction, this may not invariably hold over experiments, situations, nor individuals. Factors such as the magnitude of prediction-errors might, according to latent cause models, play an important role in determining the durability of extinction (Gershman et al., 2017). Another challenge is that some individuals maintain sustained responding during extinction which, as we have already observed on several occasions (Miedl et al., 2020; Rattel, Wegerer, et al., 2019; Wegerer, Blechert, Kerschbaum, et al., 2013), has been positively associated with intrusion formation. Similarly, also in PTSD, increased conditioned responding during extinction (Blechert, Michael, Vriends, Margraf, & Wilhelm, 2007; Wessa & Flor, 2007) may reflect a tendency to maintain strong CS-US associations that, potentially, facilitate prompter triggering of CRs upon re-encountering trauma-related cues in the environment. Unravelling factors and boundary conditions that may lead to deficient extinction in certain individuals, but not in others, remains an important goal for future studies.

It should be noted that extinction effects were only modulated by conditionability as indexed by US-expectancy ratings and SCR, not by valence ratings. Similarly, a previous study also found that variation in US-expectancy, but not objective and subjective fear responses (as measured by electromyography and ratings of CS fear) positively predicted long-term extinction retention (Brown, LeBeau, Chat, & Craske, 2017). It is well recognized that outcome measures in fear conditioning research diverge (e.g., Lonsdorf et al., 2017). While US-expectancy vs. SCR constitute rather explicit vs. implicit measures of fear learning and dissociations between the two have been reported in a number of paradigms (e.g., Schultz & Helmstetter, 2010), some studies have also reported synchrony between them. In specific, US-expectancy and SCR have shown to diverge from other measures such as startle responses and fear ratings (Sevenster, Beckers, & Kindt, 2014; Soeter & Kindt, 2011). In line with these authors and others (Blechert, Michael, Williams, Purkis, & Wilhelm, 2008; Hamm & Weike, 2005), such divergences may mirror that conditioned responding in the form of US-expectancy ratings and SCR more closely reflects contingency awareness and declarative memory of the CS-US association than measures of emotional response

such as valence ratings. Hypothetically, prediction-error driven extinction learning may thus specifically depend on a strong declarative CS-US association representation, and less on an affective responding.

Interestingly, although we observed an immediate reduction of evaluative conditioning (EC, i.e., change in valence of CS after being paired with the aversive US) immediately after extinction, in the long term (i.e., four days after fear acquisition), all participants, regardless of whether they had received extinction or not, still rated the CS+ as more unpleasant than the CS-. This concurs with some other studies showing that even though extinction may reduce EC effects, this might occur at a slower rate than other forms of Pavlovian conditioning (Blechert et al., 2008; Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). Considering the process of prediction-error driven extinction learning, we speculate that the persistence of EC could relate to the possibility of the mere absence of an aversive US not inducing enough of a "surprise" to update the CS-valence representation. While US-absence may certainly induce enough surprise to update US-expectancy (participants expecting the US are surprised because the US does not follow the CS+ anymore) and SCR, participants may need more than the mere absence of an aversive US (e.g., a positive US) in order to produce large enough discrepancy for prediction errors to drive extinction of negative valence appraisals.

Finally, CS-intrusion effects, although overall reflecting the pattern of results of US intrusions, were weaker and associated with larger uncertainty (i.e., Group effects only became apparent when considering US-expectancy conditionability as a moderator). As discussed above, the overall number of reported CS intrusions was relatively low, and thus we may have lacked statistical power to detect individual differences in CS-intrusion development. We specifically observed some floor effects in ACQ-only participants, which may have arisen from the smaller sample size of this group. Conclusions regarding CS intrusions should thus be interpreted with caution until replicated in a larger sample and/or in a prospective study where individuals encounter more intense and threat-eliciting USs that may, as previously discussed, evoke more CS intrusions than relatively mild, non-endangering USs.

Overarchingly, our results thus suggest that individuals who showed greater conditionability during acquisition, and were then randomly assigned to an extinction group, were protected from intrusion formation following analogue-trauma compared to participants randomly assigned to an ACQ-only group. On the one hand, these findings are consistent with our previous correlational studies (Miedl et al., 2020; Rattel, Wegerer, et al., 2019; Wegerer, Blechert, Kerschbaum, et al., 2013) in that extinction learning indeed seems to play an important role in intrusive memory formation. On the other hand, the novel experimental manipulation of CS-US association strength through extinction supports the hypothesis that conditioning processes may function as a causal igniting factor for intrusive memory formation.

We investigated the effects of receiving immediate extinction where, unlike in exposure therapy, acquired fear memories are not yet consolidated (McKenzie & Eichenbaum, 2011). Even though allowing a longer timeframe between fear memory acquisition and (re-)exposure to fear-eliciting stimuli (CS) would have provided a better analogue model of exposure therapy, studies implicating extinction learning as a central and predictive process of exposure therapy success (Ball et al., 2017; Forcadell et al., 2017; Waters & Pine, 2016) may permit some tentative thoughts on the extent to which our findings can potentially inform clinicians using exposure therapy protocols for treating intrusive memories after trauma. Altogether, if we assume extinction as central process of exposure therapy, our finding that extinction reduced the probability of intrusive memory development after trauma may support exposure therapy as a useful intervention for reducing intrusions after trauma. More importantly though, in line with some exposure therapy protocols (Craske, Hermans, & Vervliet, 2018), our findings partially also suggest that success in reducing intrusions through exposure may depend on the extent to which US absence actually constitutes a violation of expectancies and reliably generates prediction errors during extinction. In line with the latent cause model (Gershman et al., 2017), a future challenge may lie in finetuning those prediction errors so that individuals do not infer a new latent cause for US absence (i.e., believe that "the CS predicts the US as reliably as it did before, but something else is temporarily preventing the US from happening") and form a new memory, but instead infer the old latent cause (i.e., "the CS no longer/less reliably predicts the US") and thus update the original fear memory. In this regard, some techniques that may prevent participants from inferring a new latent cause could be reactivating the traumatic memory pre-exposure (e.g., with reminder cues or imaginal exposure to the fear memory) and approximating the extinction procedure to acquisition (e.g., using an occasional reinforced extinction protocol). While these techniques are already used and recommended in clinical practice to optimize extinction learning (Craske et al, 2014, 2018; Dunsmoor et al., 2015), it remains unclear whether their success is, as here proposed, linked to memory modification or strengthened inhibitory learning. To unravel these questions and advance research in this field, a potential first necessary step might be to detect a neurobiological marker of persistent memory alteration in such experiments.

4.4. Limitations

A number of other potential limitations needs to be contemplated. First, our two experimental groups differed in size. A bigger sample size of the ACQ+EXT group was acquired as part of a larger neural mechanism investigation (reported in Miedl et al., 2020; Rattel, Miedl, et al., 2019), but for practical reasons impossible for the ACQ-only group. As such, statistical power to detect effects in the ACQ-only group may have been reduced relative to the ACQ+EXT group. Current results should thus be regarded with some caution and replicated in a larger sample.

Second, another potential limitation is that, in relation to ACQ-only participants, ACQ+EXT participants spent additional time in the MRI in order to undergo extinction training. We purposely refrained from administering a filler-task or prolonged resting period to avoid potential effects on memory consolidation, i.e., the post-encoding time window of about 6h or more during which memory remains labile before transfer into longer term memory (Nader & Einarsson, 2010; Schiller et al., 2010). The field's infancy prevented us from unanimously excluding the possibility that specific tasks (James et al., 2016), or even prolonged resting-state following the analogue-trauma (Humiston, Tucker, Summer, & Wamsley, 2019), could have influenced memory consolidation in some unwanted ways. Possibly, the reduced time in the MRI could have led to a scenario where ACQ-only participants perceived the analogue-trauma situation as less stressful/aversive, thereby implicating reduced intrusions; this scenario was, however, not supported by the current results.

Third, we opted for immediate extinction to ensure maximal experimental control over the time subsequent to fear-acquisition. Since evidence regarding the timing between acquisition and extinction in humans is mixed (Lonsdorf et al., 2017; Vervliet, Craske, et al., 2013), we cannot be certain whether our results would also replicate if extinction had been delayed. While greater passage of time between acquisition and extinction may allow memory consolidation processes to promote the formation of distinct yet flexible emotional memory traces that confer an ability to recall extinction (i.e., CS-noUS association) and thereby reduce return of fear in the form of intrusive memories, separating acquisition and extinction in time may, as discussed within latent cause models (Gershman et al., 2017), render it more difficult for individuals to infer the "acquisition latent cause" during extinction and hamper updating of the original fear memory. To test this question, it would be important to repeat the current study with a longer time window between acquisition and extinction. Importantly, including a longer time window between acquisition and re-exposure to CS would better approximate extinction to exposure therapy (Forcadell et al., 2017), and thereby scrutinize whether the here hypothesized clinical implications indeed hold true.

Fourth, the generalizability of these findings is limited by a non-clinical sample of young female individuals. We previously observed that men respond differently to aversive film clips than women (Wilhelm et al., 2017) and show reduced intrusions following a conditioning experiment (Rattel, Wegerer, et al., 2019), thus it may be of interest to investigate whether our results also hold in a mixed sample. A related problem is that we did not control for current and past PTSD/intrusion symptomatology. Although the absence of group differences in the number of previous adverse life events, trait anxiety, and depression symptoms, as well as in peritraumatic dissociation and processing style reduce the possibility that residual PTSD-like symptomatology may have played a role in the current results, we cannot with certainty exclude this possibility.

Fifth, the total number of intrusions was relatively low in the present sample. This could be related to our event-based assessment method (i.e., intrusion and associated distress is registered upon occurrence), which depends on participants evaluating whether or not a memory of the analogue-trauma (films or faces) was indeed an "intrusion" and thus merits reporting, or rather just a fleeting, negligible event. Indeed, previous research has suggested possible underestimation of symptoms when using event-based assessments, particularly for less severe symptoms (Takarangi, Strange, & Lindsay, 2014). In contrast however, another study did not support significant differences in intrusion frequency and distress reports when event-based vs. prompted time-based (i.e., intrusion report per observation period) assessment methods were used (Rattel, Grünberger, et al., 2019). Further, even though aversive film clips were rated as more unpleasant and arousing than neutral film clips, another reason for low number of intrusions could be related to insufficiently distressing film clips. However, in this regard, it seems pertinent to ask whether intrusive memories are even tied to aversive distressing events, or whether also neutral (experimental) situations could elicit intrusive memories in certain individuals (e.g., high ruminators, see Xia & Evans, 2020). If this was the case, could it be that we merely captured normative spontaneous memories that have little to do with pathological intrusions as found in PTSD? More precisely, to what extent do our analogue-trauma intrusions reflect clinically-relevant intrusions vs. normative intrusions? And, do findings on our short-lasting analogue-trauma intrusions also contribute to the understanding of long-lasting intrusions in PTSD?

Regarding the first question, - to what extent our analogue-trauma intrusions reflect pathological intrusions -, we believe that it is important to underline that even those intrusions that later persist and are considered pathological, begin by being normal reactions following a traumatic event. Critically, studies suggest that whether or not intrusions persist is not predicted by how frequently they occur after trauma (Kleim, Ehlers, & Glucksman, 2007), but rather by how distressing they are (Clohessy & Ehlers, 1999; Ehlers & Steil, 1995; Michael, Ehlers, Halligan, & Clark, 2005). Following, individuals' degree of distress experienced during intrusions may be a good early indicator of whether intrusions persist and turn into a clinically-relevant phenomenon (if distress is high) or are rather fleeting events (if distress is low). As such, here we went beyond using mere intrusion frequency as outcome of interest, and instead focused on "intrusion load" (i.e., summed daily distress of intrusive memories). In this way, while we investigated intrusions that were still normative, we integrated a more "pathologically-relevant" aspect of these intrusions, and thereby hope to have been able to shedding light on aetiological mechanisms that are also relevant for more persistent, clinically-relevant pathological intrusions. In parallel, we attempted to move closer towards clinically-relevant intrusions by having a four-day-long intrusion assessment period. While this is still a very short time period, an important question is whether, and to what extent, insights on fleeting, relatively mildly distressing intrusions observed in the lab do not already translate to the persistent, extremely distressing intrusions observed in the clinic. Testing this expectation would however be an important next step. For example, in high-risk groups such as police or emergency health care workers, it would be interesting to assess to what degree factors predicting intrusive memory formation and development in an analogue-trauma setting remain relevant/predictive after a real-life traumatic event. More precisely, it could be examined whether stronger conditionability patterns during the acquisition phase of an analogue-trauma influence intrusive memory development after a real-life traumatic event, whether this can be modulated by exposure therapy; and ultimately influences whether or not someone develops PTSD. By these means, the clinical translation of our findings might be tested.

5. Conclusions

Together, our experimental results support the conceptualization of intrusive memories within a conditioning framework: (1) Neutral stimuli (CSs) temporally associated with analogue trauma (US) were re-experienced as intrusive memories and (2) occasionally functioned as reminder cues for intrusive memories. Perhaps most importantly, (3) weakening conditioned responses to CS through extinction reduced the probability of intrusions to film clips, reduced their severity, and accelerated their decay. These extinction effects were modulated by cognitive and physiological conditionability (i.e., differential US-expectancy and SCRs) at the end of acquisition, which may have aided prediction-error driven extinction learning. Extinction also reduced the probability of intrusions to CSs and reduced their severity, but only in participants with higher cognitive conditionability. As such, our experiment provides further support for fear conditioning playing a role in intrusion formation. However, it is obvious that it does not explain all of the variance of reported intrusions, indicating that there are additional processes at work. Further, our series of studies providing cumulative support for a conditioning framework for understanding intrusions and novel insights on mediating mechanisms and boundary conditions such as risk and protective factors (Miedl et al., 2020; Rattel, Miedl, et al., 2019; Rattel, Wegerer, et al., 2019; Wegerer, Kerschbaum, Blechert, & Wilhelm, 2014; Wegerer, Blechert, Kerschbaum, et al., 2013) does not preclude the possibility that future studies may falsify this theoretical assumption.

From a clinical perspective, this study provides novel data and insights on factors contributing to several qualitative and quantitative aspects of intrusion development. Results provide further support for using exposure-based therapy for successful reduction of intrusive memories after trauma. In specific, findings underline the importance of dismantling reminder cues of intrusive memories (Ehlers & Clark, 2011), using strategies that allow expectancy violations during extinction (Craske et al., 2018; Struyf et al., 2018), and approximating extinction to acquisition (e.g., via occasional reinforced extinction, Craske et al., 2018; Gershman et al., 2017).

CRediT authorship contribution statement

Laila K. Franke: Conceptualization, Methodology, Formal analysis, Visualization, Writing – original draft. **Julina A. Rattel:** Conceptualization, Methodology, Investigation, Data curation, Project administration, Writing – review & editing. **Stephan F. Miedl:** Conceptualization, Methodology, Software, Writing – review & editing. **Sarah K. Danböck:** Methodology, Writing – review & editing. **Paul-Christian Bürkner:** Methodology, Formal analysis, Visualization, Writing – review & editing. **Frank H. Wilhelm:** Conceptualization, Methodology, Supervision, Project administration, Writing – review & editing.

Declaration of competing interest

All authors declare no conflict of interest.

Acknowledgments

The second and fourth authors were supported by the Doctoral College "Imaging the Mind" (FWF; W1233-G17, W1233–B; principal investigator Frank H. Wilhelm).

Appendices.

Appendix A. Intrusion classification system

Step 1: Classification of intrusive memories into subcategories

We defined a classification system where intrusive memories could be categorized into 36 subcategories, as detailed in Table A (second column). Based on each brief intrusive memory description provided by participants, the first author and a second independent rater who did not run the study categorized intrusive memories into one of the pre-defined subcategories. Agreement among raters was excellent (*Cohen's kappa* = 0.80). In case of disagreement, the first author's rating was used.

Step 2: Classification of intrusive memories into final categories

Based on the classification into the 36 categories, intrusive memories were subsequently grouped together into seven final, broader categories. Intrusive memories concerning the US were grouped into three categories based on the type of trauma: (1) "interpersonal trauma" (2) "accidental trauma", or (3) "mixed US", when intrusions could not be clearly attributed to either interpersonal or accidental trauma, referred to neutral film clips, or were unclassifiable. Intrusive memories concerning CS cues remained separated into (4) "CS+" and (5) "CS-", with a further category (6) "mixed CS" comprising intrusive memories referring to both CS+ and CS- and/or intrusive memories whose description did not allow for a clear distinction between CS+ and CS- face. A final category (7) "others" comprised intrusive memories whose description did not match any CS face.

Table A
Definition of Intrusive Memory Categories and Examples of Descriptions in the Diary

Final Categories		Comprised subcategories	Example of intrusions reported in e-diary
US	1.Interpersonal trauma	"Hostel"; "Antichrist"; "Scar"; any combination of two or more of these scenes.*	"man tied to chair is being tortured by older man" "woman inserts screws into man's leg" "woman finds dead boyfriend in the bathroom" "man with cut throat on bathroom floor; screwing leg; man tied to a chair"
	2.Accidental trauma	"127-Hours"; "Dantes Peak"; "Final Destination"; any combination of two or more of these scenes*	"man in the mountains cuts off his arm" "war, man and woman in a car" "plane crashing" "man cuts his own flesh, woman screams in car"
	3. Mixed	any combination of interpersonal-trauma scenes and accidental-trauma scenes; neutral film-scenes; any intrusion whose description did not allow clear attribution to a specific film-scene.	"screwing on someone's leg; man is separating his own arm from his body" "woman in café" "blood streaming down body"
CS	4. CS+	intrusion whose description matched individuals' CS+ face.	"woman with starring gaze (brown hair)" "woman shown before the unpleasant movies"
	5. CS-	intrusion whose description matched individuals' CS- face.	"woman with light blonde hair" "neutral woman who was shown mostly before neutral movies"
	6. Mixed	Intrusion referring to both CS+ and CS- together; any intrusion whose description did not allow clear distinction between CS+ and CS- face.	"woman with blond hair and woman with brown hair" "face perceived as more pleasant"
Others		Intrusive memories whose description clearly did not match any of the experimental CS-faces or US-films	"blond tall woman" (individualized CSs did not include a blond woman; height was not perceivable from face pictures).

Note: "Any combination of two or more interpersonal scenes" = e.g., Hostel/Antichrist, Hostel/Scar/Hostel/Antichrist & Scar; "Any combination of two or more accidental scenes" = e.g., 127-Hours/Dantes Peak, 127-Hours/Final Destination, 127-Hours/Dantes Peak/Final Destination. The intrusive memory coded as "others" was excluded from subsequent analyses.

Appendix B. Intrusive memory reminder cue classification system

Based on each brief reminder cue description provided by participants, the first author and a second independent rater who did not run the study categorized cues into one of the four cue content categories, as well as into one of the three cue origin categories. Agreement among raters was good for (A) trigger content (*Cohen's kappa* = 0.64) and excellent for (B) trigger origin (*Cohen's kappa* = 0.73) classification. In case of disagreement, the first author's rating was used.

Table B
Definition of Cue Categories and Examples of Descriptions in the Diary

Category	Definition	Example
A. Content		
1. US	Referents to the US including: - Thoughts/linguistic referents such as hearing someone talk about violence; - sensory and perceptual experiences such as the sound of movie soundtracks or the sight of objects resembling elements of the film clips; - physiological/emotional states.	„showering and seeing water on the ground" (B1) "hearing the word pain" (B1) „numb arm from laying on it, made me think of amputated arm" (B2)
2. CS	Referents to the CS including: - thoughts/linguistic referents such as thinking about the CS faces; - sensory and perceptual experiences such as the sight of persons or faces; - physiological or mood states.	„seeing a similar face" (B1) "woman on bike" (B1) „two people on the sideway" (B1)
3. Experimental context	Referents to the wider experimental context including: - thoughts/linguistic referents, such as hearing someone talk about any study participation; - sensory and perceptual experiences such as the sight of elements from the experimental context (e.g. smartphones) or the sound of the MRI scanner; - physiological or mood states.	„switching on the study phone" (B1) "cables" (B1) „seeing study instruction sheet" (B1)
4.Unspecific elements	Cues not clearly resembling the US, CS, or wider experimental context, including abstract cues, perceptual/sensory cues, and state cues.	„waiting for the train" (B1) "open window in empty house" (B1) "thinking about yesterday" (B2)
B. Origin		
1. External	External source in the environment, including sensory and perceptual experiences	
2. Internal	Internal source, such as bodily sensations (e.g., pain, being cold), thoughts, and emotional states (e.g., feeling fearful)	
3. Mixed	Any combination of an internal and external cue, or when unclear whether cue is internal or external	

Appendix C. Overview of statistical models and respective formula

Table C
Overview of Bayesian Multilevel Models used to predict extinction effects on US and CS intrusions

US intrusions	I	US_intrusions ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × Con × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject)
	II	US_intrusions ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject)
	III	US_intrusions ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject)
CS intrusions	I	CS_intrusions ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject)
	II	CS_intrusions ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject)
	III	CS_intrusions ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject)

Note. Models were fitted in two parts, where the first lognormal part of the model predicted the amount of non-zero intrusions with a lognormal distribution, and the second hurdle part of the model predicted the probability of zero intrusions with a binomial distribution. Abbreviations: US-EXP = US-expectancy; VAL = valence; SCR = skin conductance responses; CS = conditioned stimuli; US = unconditioned stimuli; hu = hurdle.

Appendix D. Results of sensitivity analyses (random-intercept-only vs. random-slopes BMLMs)

Table D
Comparison between random-intercept-only and random-slopes Bayesian Multilevel Models to predict extinction effects on US and CS intrusions

		Formula	Expected log posterior density	
			diff	se
US-ints	I-rs	US_ints ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject)	−2.5	2.0
	I-nrs	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 | Subject)	0	0
	II-rs	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject)	−4	2.2
	II-nrs	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 | Subject)	0	0
	III-rs	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject)	−1.0	2.6
	III-nrs	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 | Subject)	0	0
CS-ints	I-rs	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject)	−3.7	1.2
	I-nrs	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 | Subject)	0	0
	II-rs	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject)	−3.7	1.6
	II-nrs		0	0

(*continued on next page*)

Table D (*continued*)

		Formula	Expected log posterior density	
			diff	se
		CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 | Subject)		
	III-rs	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject)	−4.5	1.7
	III-nrs	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 | Subject)	0	0

Note. Comparisons carried out with *K*-fold cross-validation. Data were randomly portioned in K subsets of equal size. Then the model was refit 10 times, each time leaving out one of the K subsets. K was equal to the number of observations (N = 336). Higher expected log posterior density values indicate better fit. Abbreviations: lin = linear; mon = monotonic; diff = difference; se = standard error; US-EXP = US-expectancy; VAL = valence; SCR = skin conductance responses; CS = conditioned stimuli; US = unconditioned stimuli; ints = intrusions; hu = hurdle; rs = model with random slopes; nrs = model with no random slopes.

Appendix E. Results of sensitivity analyses (linear vs. monotonic BMLMs)

Table E
Comparison between linear and monotonic Bayesian Multilevel Models to predict extinction effects on US and CS intrusions

		Formula	Expected log posterior density	
			diff	se
US-ints	I-lin	US_ints ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject)	−4.3	4.5
	I-mon	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × mo (Day) + (1 + mo (Day) | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × mo (Day) + (1 + mo (Day) | Subject)	0	0
	II-lin	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject)	−3.7	4.1
	II-mon	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × mo (Day) + (1 + mo (Day) | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × mo (Day) + (1 + mo (Day) | Subject)	0	0
	III-lin	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject)	−0.2	4.2
	III-mon	US_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × mo (Day) + (1 + mo (Day) | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × mo (Day) + (1 + mo (Day) | Subject)	0	0
CS-ints	I-lin	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × Day + (1 + Day | Subject)	−0.4	3.5
	I-mon	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × mo (Day) + (1 + mo (Day) | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{US-EXP}}$ × mo (Day) + (1 + mo (Day) | Subject)	0	0
	II-lin	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × Day + (1 + Day | Subject)	0	0
	II-mon	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{VAL}}$ × mo (Day) + (1 + mo (Day) | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{VAL}}$ × mo (Day) + (1 + mo (Day) | Subject)	−6.2	3.1
	III-lin	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × Day + (1 + Day | Subject)	0	0
	III-mon	CS_ints ~ 1 + Group × $\text{Conditionability}_{\text{SCR}}$ × mo (Day) + (1 + mo (Day) | Subject), Hu ~1 + Group × $\text{Conditionability}_{\text{SCR}}$ × mo (Day) + (1 + mo (Day) | Subject)	−1.0	3.6

Note. Comparison linear and monotonic models carried out with *K*-fold cross-validation. Data were randomly portioned in K subsets of equal size. Then the model was refit 10 times, each time leaving out one of the K subsets. K was equal to the number of observations (N = 336). Higher expected log posterior density values indicate better fit. Abbreviations: lin = linear; mon = monotonic; diff = difference; se = standard error; US-EXP = US-expectancy; VAL = valence; SCR = skin conductance responses; CS = conditioned stimuli; US = unconditioned stimuli; ints = intrusions; hu = hurdle.

Appendix F. Plots depicting expected means for whole models

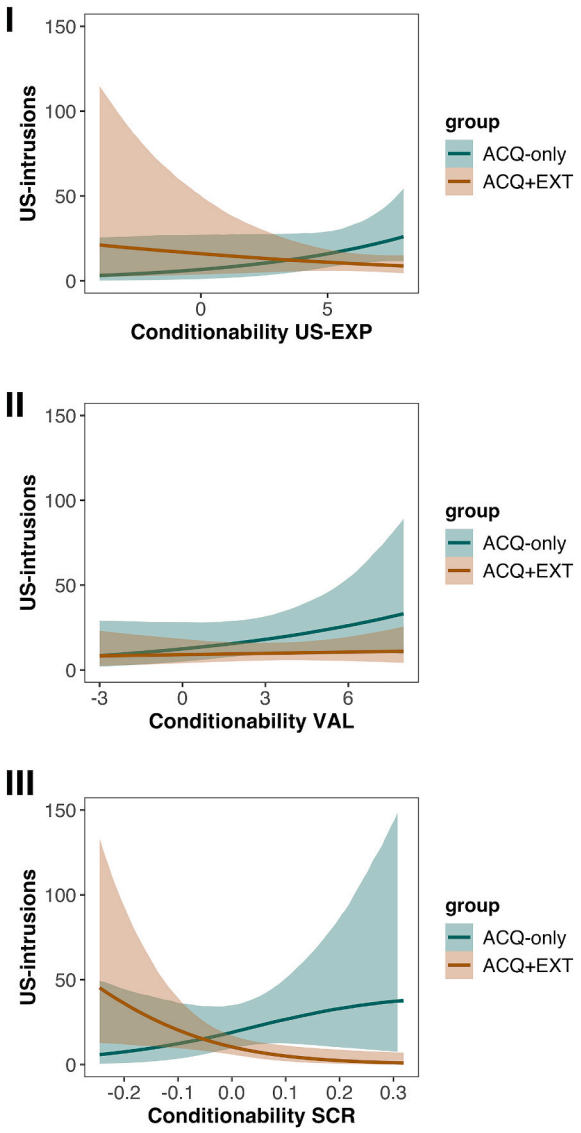

Fig. F.I. Solid lines depict the regression fit, shaded areas represent 95% credible intervals. Expected means for whole models (function of both zero (hurdle) and non-zero (lognormal) parts of the models, i.e., US intrusions predicted by Group and Conditionability$_{\text{US-EXP}}$ (I), Conditionability$_{\text{VAL}}$ (II), and Conditionability$_{\text{SCR}}$ (III) including both zero and non-zero responses). Higher values represent more US intrusions. For illustrative purposes and better appreciation, plots depict non-mean-centered Conditionability estimates. Abbreviations: ACQ-only = acquisition-only participants; ACQ+EXT = acquisition plus extinction participants; US-EXP = US-expectancy; VAL = Valence; SCR = skin conductance response; US = unconditioned stimuli.

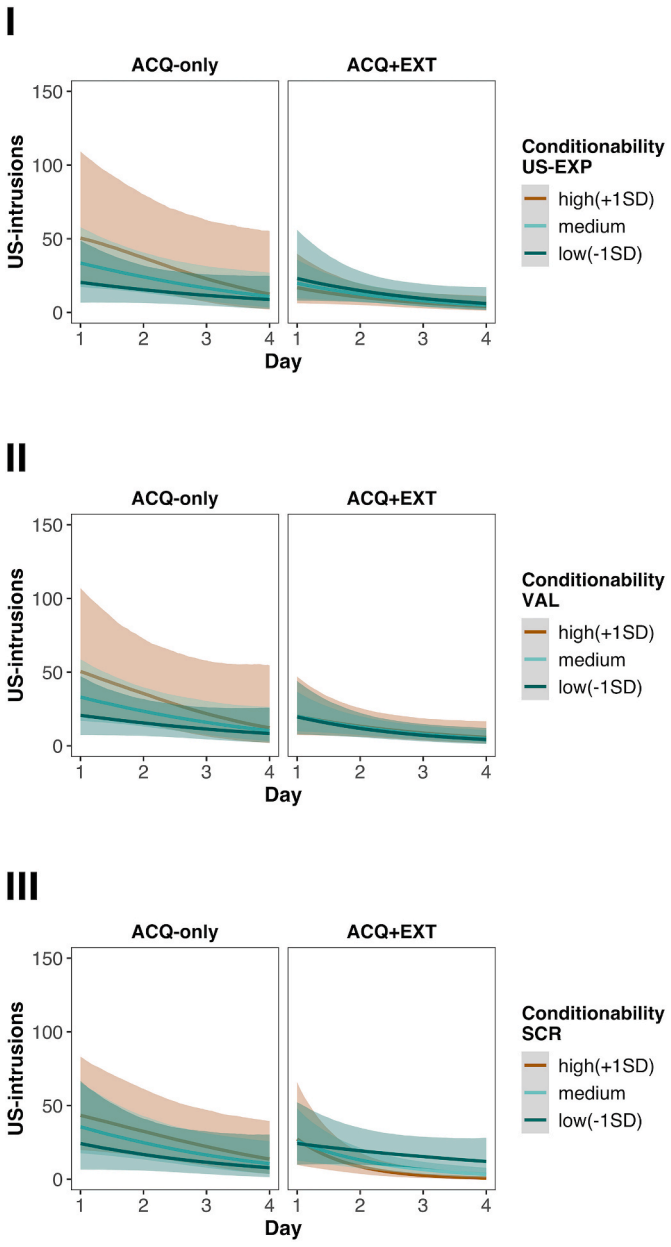

Fig. F.II. Solid lines depict the regression fit, shaded areas represent 95% credible intervals. Expected means for whole models (function of both zero (hurdle) and non-zero (lognormal) parts of the models, i.e., US intrusions predicted by Day, Group and Conditionability$_{US\text{-}EXP}$ (I), Conditionability$_{VAL}$ (II), and Conditionability$_{SCR}$ (III) including both zero and non-zero responses). Higher values represent more US intrusions. For illustrative purposes and better appreciation, plots depict non-mean-centered Conditionability estimates. Abbreviations: ACQ-only = acquisition-only participants; ACQ+EXT = acquisition plus extinction participants; US-EXP = US-expectancy; VAL = Valence; SCR = skin conductance response; US = unconditioned stimuli.

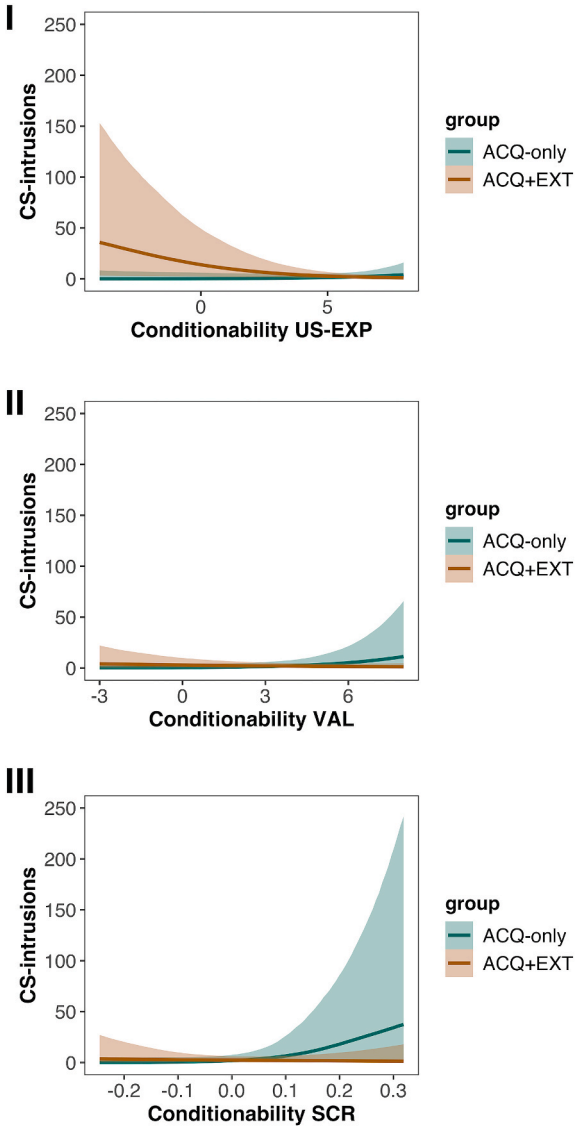

Fig. F.III. Solid lines depict the regression fit, shaded areas represent 95% credible intervals. Expected means for whole models (function of both zero (hurdle) and non-zero (lognormal) parts of the models, i.e., CS intrusions predicted by Group and Conditionability$_{US\text{-}EXP}$ (I), Conditionability$_{VAL}$ (II), and Conditionability$_{SCR}$ (III) including both zero and non-zero responses). Higher values represent more CS intrusions. For illustrative purposes and better appreciation, plots depict non-mean-centered Conditionability estimates. Abbreviations: ACQ-only = acquisition-only participants; ACQ+EXT = acquisition plus extinction participants; US-EXP = US-expectancy; VAL = Valence; SCR = skin conductance response; CS = conditioned stimuli.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596.744053.

Ball, T. M., Knapp, S. E., Paulus, M. P., & Stein, M. B. (2017). Brain activation during fear extinction predicts exposure success. *Depression and Anxiety, 34*(3), 257–266. https://doi.org/10.1002/da.22583.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

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*(1), 143–152. https://doi.org/10.5127/jep.048615.

Berntsen, D. (2009). *Involuntary autobiographical memories: An introduction to the unbidden past*. https://doi.org/10.1017/CBO9780511575921.

Betancourt, M. (2017). *A conceptual introduction to Hamiltonian Monte Carlo*. arXiv preprint arXiv:1701.02434.

Blechert, J., Michael, T., Vriends, N., Margraf, J., & Wilhelm, F. H. (2007). Fear conditioning in posttraumatic stress disorder: Evidence for delayed extinction of autonomic, experiential, and behavioural responses. *Behaviour Research and Therapy, 45*(9), 2019–2033. https://doi.org/10.1016/J.BRAT.2007.02.012.

Blechert, J., Michael, T., Williams, S. L., Purkis, H. M., & Wilhelm, F. H. (2008). When two paradigms meet: Does evaluative learning extinguish in differential fear conditioning? *Learning and Motivation, 39*(1), 58–70. https://doi.org/10.1016/j.lmot.2007.03.003.

Blechert, J., Peyk, P., Liedlgruber, M., & Wilhelm, F. H. (2016). Anslab: Integrated multichannel peripheral biosignal processing in psychophysiological science. *Behavior Research Methods, 48*(4), 1528–1545. https://doi.org/10.3758/s13428-015-0665-1.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., et al. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology, 49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x.

Breslau, N. (2001). The epidemiology of posttraumatic stress disorder: What is the extent of the problem? *Journal of Clinical Psychiatry, 62*(Suppl 17), 16–22.

Brewin, C. R., Gregory, J. D., Lipton, M., & Burgess, N. (2010). Intrusive images in psychological disorders: Characteristics, neural mechanisms, and treatment implications. *Psychological Review, 117*(1), 210–232. https://doi.org/10.1037/a0018113.

Brown, L. A., LeBeau, R. T., Chat, K. Y., & Craske, M. G. (2017). Associative learning versus fear habituation as predictors of long-term extinction retention. *Cognition & Emotion, 31*(4), 687–698. https://doi.org/10.1080/02699931.2016.1158695.

Bürkner, P.-C. (2017a). brms: An *R* package for bayesian multilevel models using *stan*. *Journal of Statistical Software, 80*(1). https://doi.org/10.18637/jss.v080.i01.

Bürkner, P. C. (2017b). Advanced Bayesian multilevel modeling with the R package brms. *R Journal, 10*(1), 395–411. https://doi.org/10.32614/rj-2018-017.
Bürkner, P. C., & Charpentier, E. (2020). Modelling monotonic effects of ordinal predictors in Bayesian regression models. *British Journal of Mathematical and Statistical Psychology, 73*(3), 420–451. https://doi.org/10.1111/bmsp.12195.
Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1). https://doi.org/10.18637/jss.v076.i01.
Charney, D. S., Deutch, A. Y., Krystal, J. H., Southwick, S. M., & Davis, M. (1993). Psychobiologic mechanisms of posttraumatic stress disorder. *Archives of General Psychiatry, 50*(4), 294–305. https://doi.org/10.1001/archpsyc.1993.01820160064008.
Clark, I. A., Mackay, C. E., & Holmes, E. A. (2015). Low emotional response to traumatic footage is associated with an absence of analogue flashbacks: An individual participant data meta-analysis of 16 trauma film paradigm experiments. *Cognition & Emotion, 29*(4), 702–713. https://doi.org/10.1080/02699931.2014.926861.
Clohessy, S., & Ehlers, A. (1999). PTSD symptoms, response to intrusive memories and coping in ambulance service workers. *British Journal of Clinical Psychology, 38*(3), 251–265. https://doi.org/10.1348/014466599162836.
Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. *Philosophical Transactions of the Royal Society B: Biological Sciences, 356*(1413), 1375–1384. https://doi.org/10.1098/rstb.2001.0940.
Craske, M. G., Hermans, D., & Vansteenwegen, D. (2006). *Fear and learning: From basic processes to clinical implications*. American Psychological Association. https://doi.org/10.1037/11474-000.
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*(1742). https://doi.org/10.1098/rstb.2017.0025.
Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006.
Craske, M. G., Waters, A. M., Nazarian, M., Mineka, S., Zinbarg, R. E., Griffith, J. W., et al. (2009). Does neuroticism in adolescents moderate contextual and explicit threat cue modulation of the startle reflex? *Biological Psychiatry, 65*(3), 220–226. https://doi.org/10.1016/j.biopsych.2008.07.020.
Craske, M. G., Wolitzky-Taylor, K. B., Mineka, S., Zinbarg, R., Waters, A. M., Vrshek-Schallhorn, S., et al. (2012). Elevated responding to safe conditions as a specific risk factor for anxiety versus depressive disorders: Evidence from a longitudinal investigation. *Journal of Abnormal Psychology, 121*(2), 315–324. https://doi.org/10.1037/a0025738.
Culver, N. C., Vervliet, B., & Craske, M. G. (2015). Compound extinction: Using the rescorla–wagner model to maximize exposure therapy effects for anxiety disorders. *Clinical Psychological Science, 3*(3), 335–348. https://doi.org/10.1177/2167702614542103.
Davis, M., Falls, W. A., & Gewirtz, J. (2000). Neural systems involved in fear inhibition: Extinction and conditioned inhibition. In M. S. Myslobodsky, & I. Weiner (Eds.), *Contemporary issues in modeling psychopathology* (pp. 113–141). Springer US. https://doi.org/10.1007/978-1-4757-4860-4_8.
De Houwer, J. (2020). Revisiting classical conditioning as a model for anxiety disorders: A conceptual analysis and brief review. *Behaviour Research and Therapy, 127* (January), 103558. https://doi.org/10.1016/j.brat.2020.103558.
Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience, 3*(1), 1–8. https://doi.org/10.1162/jocn.1991.3.1.1.
Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., et al. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety, 32*(4), 239–253. https://doi.org/10.1002/da.22353.
Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking extinction. *Neuron, 88*(1), 47–63. https://doi.org/10.1016/j.neuron.2015.09.028.
Dunsmoor, J. E., Prince, S. E., Murty, V. P., Kragel, P. A., & LaBar, K. S. (2011). Neurobehavioral mechanisms of human fear generalization. *NeuroImage, 55*(4), 1878–1888. https://doi.org/10.1016/j.neuroimage.2011.01.041.
Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behavior Research and Therapy, 38*, 319–345. https://doi.org/10.1016/S0005-7967(99)00123-0.
Ehlers, A., & Clark, D. M. (2011). UKPMC Funders Group Posttraumatic stress disorder : The development of effective psychological treatments. *Psychology, 62*(Suppl 47), 11–18. https://doi.org/10.1080/08039480802315608.Posttraumatic.
Ehlers, A., Hackmann, A., & Michael, T. (2004). Intrusive re-experiencing in post-traumatic stress disorder: Phenomenology, theory, and therapy. *Memory, 12*(4), 403–415. https://doi.org/10.1080/09658210444000025.
Ehlers, A., Hackmann, A., Steil, R., Clohessy, S., Wenninger, K., & Winter, H. (2002). The nature of intrusive memories after trauma: The warning signal hypothesis. *Behaviour Research and Therapy, 40*(9), 995–1002. https://doi.org/10.1016/S0005-7967(01)00077-8.
Ehlers, A., & Steil, R. (1995). Maintenance of intrusive memories in posttraumatic stress disorder: A cognitive approach. *Behavioural and Cognitive Psychotherapy, 23*(3), 217. https://doi.org/10.1017/S135246580001585X.
Ehring, T., Fuchs, N., & Kläsener, I. (2009). The effects of experimentally induced rumination versus distraction on analogue posttraumatic stress symptoms. *Behavior Therapy, 40*(4), 403–413. https://doi.org/10.1016/j.beth.2008.10.001.
Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy, 20*(2), 155–176. https://doi.org/10.1016/S0005-7894(89)80067-X.
Forcadell, E., Torrents-Rodas, D., Vervliet, B., Leiva, D., Tortella-Feliu, M., & Fullana, M. A. (2017). Does fear extinction in the laboratory predict outcomes of exposure therapy? A treatment analog study. *International Journal of Psychophysiology, 121*(September), 63–71. https://doi.org/10.1016/j.ijpsycho.2017.09.001.
Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.
Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for bayesian regression models. *The American Statistician, 73*(3), 307–309. https://doi.org/10.1080/00031305.2018.1549100.
Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press. https://doi.org/10.1017/CBO9780511790942.
Gershman, S. J., Monfils, M. H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *ELife, 6*, 1–41. https://doi.org/10.7554/eLife.23763.
Gershman, S. J., & Niv, Y. (2012). Exploring a latent cause theory of classical conditioning. *Learning & Behavior, 40*(3), 255–268. https://doi.org/10.3758/s13420-012-0080-8.
Grey, N., & Holmes, E. A. (2008). "Hotspots" in trauma memories in the treatment of post-traumatic stress disorder: A replication. *Memory, 16*(7), 788–796. https://doi.org/10.1080/09658210802266446.
Haaker, J., Lonsdorf, T. B., Schümann, D., Menz, M., Brassen, S., Bunzeck, N., et al. (2015). Deficient inhibitory processing in trait anxiety: Evidence from context-dependent fear learning, extinction recall and renewal. *Biological Psychology, 111*, 65–72. https://doi.org/10.1016/j.biopsycho.2015.07.010.
Hackmann, A., Ehlers, A., Speckens, A., & Clark, D. M. (2004). Characteristics and content of intrusive memories in PTSD and their changes with treatment. *Journal of Traumatic Stress, 17*(3), 231–240. https://doi.org/10.1023/B:JOTS.0000029266.88369.fd.
Halligan, S. L., Michael, T., Clark, D. M., & Ehlers, A. (2003). Posttraumatic stress disorder following assault: The role of cognitive processing, trauma memory, and appraisals. *Journal of Consulting and Clinical Psychology, 71*(3), 419–431. https://doi.org/10.1037/0022-006X.71.3.419.
Hamm, A. O., & Weike, A. I. (2005). The neuropsychology of fear learning and fear regulation. *International Journal of Psychophysiology, 57*(1), 5–14. https://doi.org/10.1016/j.ijpsycho.2005.01.006.
Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*(3), 390–421. https://doi.org/10.1037/a0018916.
Holmes, E. A., & Bourne, C. (2008). Inducing and modulating intrusive emotional memories: A review of the trauma film paradigm. *Acta Psychologica, 127*(3), 553–566. https://doi.org/10.1016/j.actpsy.2007.11.002.
Holmes, E. A., Grey, N., & Young, K. A. D. (2005). Intrusive images and "hotspots" of trauma memories in Posttraumatic Stress Disorder: An exploratory investigation of emotions and cognitive themes. *Journal of Behavior Therapy and Experimental Psychiatry, 36*(1 SPEC. ISS.), 3–17. https://doi.org/10.1016/j.jbtep.2004.11.002.
Humiston, G. B., Tucker, M. A., Summer, T., & Wamsley, E. J. (2019). Resting states and memory consolidation: A preregistered replication and meta-analysis. *Scientific Reports, 9*(1), 1–9. https://doi.org/10.1038/s41598-019-56033-6.
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.
Jasp Team. (2020). *Jasp* [Computer software] Version 0.12.2. .
Keane, T., Zimering, R., & Caddell, J. (1985). A behavioral formulation of posttraumatic stress disorder in Vietnam veterans. *The Behavior Therapist, 8*(1), 9–12.
Kleim, B., Ehlers, A., & Glucksman, E. (2007). Early predictors of chronic post-traumatic stress disorder in assault survivors. *Psychological Medicine, 37*(10), 1457–1467. https://doi.org/10.1017/S0033291707001006.
Kubany, E. S., Leisen, M. B., Kaplan, A. S., Watson, S. B., Haynes, S. N., Owens, J. A., et al. (2000). Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The Traumatic Life Events Questionnaire. *Psychological Assessment, 12*(2), 210–224. https://doi.org/10.1037/1040-3590.12.2.210.
Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition & Emotion, 24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076.
Lau-Zhu, A., Holmes, E. A., & Porcheret, K. (2018). Intrusive memories of trauma in the laboratory: Methodological developments and future directions. *Current Behavioral Neuroscience Reports, 5*(1), 61–71. https://doi.org/10.1007/s40473-018-0141-1.
Laux, P., Glanzmann, P., & Schaffner, C. D. S. (1981). *Das state-trait-angstinventar (STAI)*. Weinheim, Germany: Beltz.
Li, N., Elashoff, D. A., Robbins, W. A., & Xun, L. (2011). A hierarchical zero-inflated log-normal model for skewed responses. *Statistical Methods in Medical Research, 20*(3), 175–189. https://doi.org/10.1177/0962280208097372.
Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., et al. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy, 43*(11), 1391–1424. https://doi.org/10.1016/J.BRAT.2004.10.007.
Lissek, S., Rabin, S. J., McDowell, D. J., Dvir, S., Bradford, D. E., Geraci, M., et al. (2009). Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. *Behaviour Research and Therapy, 47*(2), 111–118. https://doi.org/10.1016/j.brat.2008.10.017.
Lissek, S., & van Meurs, B. (2015). Learning models of PTSD: Theoretical accounts and psychobiological evidence. *International Journal of Psychophysiology, 98*(3), 594–605. https://doi.org/10.1016/J.IJPSYCHO.2014.11.006.
Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., et al. (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.

McKenzie, S., & Eichenbaum, H. (2011). Consolidation and reconsolidation: Two lives of memories? *Neuron, 71*(2), 224–233. https://doi.org/10.1016/j.neuron.2011.06.037.

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*(January). https://doi.org/10.1016/j.brat.2020.103556.

Meyer, T. D., & Hautzinger, M. (2001). Allgemeine depressions-skala (ADS). *Diagnostica, 47*(4), 208–215. https://doi.org/10.1026//0012-1924.47.4.208.

Michael, T., Ehlers, A., Halligan, S. L., & Clark, D. M. (2005). Unwanted memories of assault: What intrusion characteristics are associated with PTSD? *Behaviour Research and Therapy, 43*(5), 613–628. https://doi.org/10.1016/J.BRAT.2004.04.006.

Michopoulos, V., Beurel, E., Gould, F., Dhabhar, F. S., Schultebraucks, K., Galatzer-Levy, I., et al. (2019). Association of prospective risk for chronic PTSD symptoms with low TNFα and IFNγ concentrations in the immediate aftermath of trauma exposure. *American Journal of Psychiatry, 177*(1), 58–65. https://doi.org/10.1176/appi.ajp.2019.19010039.

Miedl, S. F., Rattel, J. A., Franke, L. K., Blechert, J., Kronbichler, M., Spoormaker, V. I., et al. (2020). Neural processing during fear extinction predicts intrusive memories. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(4). https://doi.org/10.1016/j.bpsc.2019.12.017.

Mineka, S., & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica, 127*(3), 567–580. https://doi.org/10.1016/j.actpsy.2007.11.007.

Murray, J., Ehlers, A., & Mayou, R. A. (2002). Dissociation and post-traumatic stress disorder: Two prospective studies of road traffic accident survivors. *British Journal of Psychiatry, 180*(APR), 363–368. https://doi.org/10.1192/bjp.180.4.363.

Nader, K., & Einarsson, E.Ö. (2010). Memory reconsolidation: An update. *Annals of the New York Academy of Sciences, 1191*, 27–41. https://doi.org/10.1111/j.1749-6632.2010.05443.x.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483–522. https://doi.org/10.1037/0033-295X.108.3.483.

Orr, S. P., Metzger, L. J., Lasko, N. B., Macklin, M. L., Peri, T., & Pitman, R. K. (2000). De novo conditioning in trauma-exposed individuals with and without posttraumatic stress disorder. *Journal of Abnormal Psychology, 109*(2), 290–298. https://doi.org/10.1037//0021-843X.109.2.290.

Pitman, R. K., Rasmusson, A. M., Koenen, K. C., Shin, L. M., Orr, S. P., Gilbertson, M. W., et al. (2012). Biological studies of post-traumatic stress disorder. *Nature Reviews Neuroscience, 13*(11), 769–787. https://doi.org/10.1038/nrn3339.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385–401. https://doi.org/10.1177/014662167700100306.

Rattel, J. A., Grünberger, L. M., Reichenberger, J., Liedlgruber, M., Miedl, S. F., Blechert, J., et al. (2019). Frequency of intrusions and appraisal of related distress after analogue trauma: A comparative ecological momentary assessment methods study. *Cognitive Therapy and Research, 43*(1), 174–184. https://doi.org/10.1007/s10608-018-9941-6.

Rattel, J. A., Miedl, S. F., Franke, L. K., Grünberger, L. M., Blechert, J., Kronbichler, M., et al. (2019). Peritraumatic neural processing and intrusive memories: The role of lifetime adversity. *Biological psychiatry. Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2018.12.010, 0(0).

Rattel, J. A., Wegerer, M., Miedl, S. F., Blechert, J., Grünberger, L. M., Craske, M. G., et al. (2019). Peritraumatic unconditioned and conditioned responding explains sex differences in intrusions after analogue trauma. July 2018 *Behaviour Research and Therapy, 116*, 19–29. https://doi.org/10.1016/j.brat.2019.01.009.

RCore Team. (2019). R: A language and environment for statistical computing. In *Foundation for statistical computing. Foundation for statistical computing*. http://www.r-project.org/index.html.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning. Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton- Century-Crofts.

Rolls, E. T. (1992). Neurophysiology and functions of the primate amygdala. In J. P. Aggleton (Ed.), *The amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction* (pp. 143–165). Wiley-Liss.

Rothbaum, B. O., & Davis, M. (2003). Applying learning principles to the treatment of post-trauma reactions. *Annals of the New York Academy of Sciences, 1008*, 112–121. https://doi.org/10.1196/annals.1301.012.

RStudio Team. (2020). RStudio: Integrated development for R [Computer software] (No. 2). *RStudio*. http://www.rstudio.com/.

Schiller, D., Monfils, M.-H., Raio, C. M., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature, 463*(7277), 49–53. https://doi.org/10.1038/nature08637.

Schultz, D. H., & Helmstetter, F. J. (2010). Classical conditioning of autonomic fear responses is independent of contingency awareness. *Journal of Experimental Psychology. Animal Behavior Processes, 36*(4), 495–500. https://doi.org/10.1037/a0020263.

Seeley, W. W. (2019). The salience network: A neural system for perceiving and responding to homeostatic demands. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 39*(50), 9878–9882. https://doi.org/10.1523/JNEUROSCI.1138-17.2019.

Sevenster, D., Beckers, T., & Kindt, M. (2014). Fear conditioning of SCR but not the startle reflex requires conscious discrimination of threat and safety. *Frontiers in Behavioral Neuroscience, 8*(FEB). https://doi.org/10.3389/fnbeh.2014.00032.

Soeter, M., & Kindt, M. (2011). Noradrenergic enhancement of associative fear memory in humans. *Neurobiology of Learning and Memory, 96*(2), 263–271. https://doi.org/10.1016/j.nlm.2011.05.003.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory*. https://ubir.buffalo.edu/xmlui/handle/10477/2895.

Steil, R., & Ehlers, A. (2000). Dysfunctional meaning of posttraumatic intrusions in chronic PTSD. *Behaviour Research and Therapy, 38*(6), 537–558. http://www.ncbi.nlm.nih.gov/pubmed/10846804.

Streb, M., Conway, M. A., & Michael, T. (2017). Conditioned responses to trauma reminders: How durable are they over time and does memory integration reduce them? *Journal of Behavior Therapy and Experimental Psychiatry, 57*, 88–95. https://doi.org/10.1016/j.jbtep.2017.04.005.

Struyf, D., Hermans, D., & Vervliet, B. (2018). Maximizing the generalization of fear extinction: Exposures to a peak generalization stimulus. *Behaviour Research and Therapy, 111*(September), 1–8. https://doi.org/10.1016/j.brat.2018.09.005.

Takarangi, M. K. T., Strange, D., & Lindsay, D. S. (2014). Self-report may underestimate trauma intrusions. *Consciousness and Cognition, 27*(1), 297–305. https://doi.org/10.1016/j.concog.2014.06.002.

Teegen, F. (2003). *Posttraumatische Belastungsstörungen bei gefährdeten Berufsgruppen : Prävalenz - prävention - Behandlung*. Bern: Huber.

Tooze, J. A., Grunwald, G. K., & Jones, R. H. (2002). Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research, 11*(4), 341–355. https://doi.org/10.1191/0962280202sm291ra.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P. C. (2021). Rank-normalization, folding, and localization: An improved Rb for assessing convergence of MCMC. *Bayesian Analysis, 1*(1), 1–28. https://doi.org/10.1214/20-BA1221.

Vervliet, B., Baeyens, F., Van den Bergh, O., & Hermans, D. (2013a). Extinction, generalization, and return of fear: A critical review of renewal research in humans. *Biological Psychology, 92*(1), 51–58. https://doi.org/10.1016/j.biopsycho.2012.01.006.

Vervliet, B., Craske, M. G., & Hermans, D. (2013b). 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.

Visser, R. M. (2020). Why do certain moments haunt us? Conceptualizing intrusive memories as conditioned responses. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(4), 375–376. https://doi.org/10.1016/j.bpsc.2020.02.007.

Waters, A. M., & Pine, D. S. (2016). Evaluating differences in Pavlovian fear acquisition and extinction as predictors of outcome from cognitive behavioural therapy for anxious children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 57*(7), 869–876. https://doi.org/10.1111/jcpp.12522.

Wegerer, M., Blechert, J., Kerschbaum, H., & Wilhelm, F. H. (2013a). Relationship between fear conditionability and aversive memories: Evidence from a novel conditioned-intrusion paradigm. *PloS One, 8*(11). https://doi.org/10.1371/journal.pone.0079025.

Wegerer, M., Blechert, J., & Wilhelm, F. H. (2013b). Emotionales Lernen: Ein naturalistisches experimentelles Paradigma zur Untersuchung von Angsterwerb und Extinktion mittels aversiver Filme. *Zeitschrift Fur Psychiatrie, Psychologie Und Psychotherapie, 61*(2), 93–103. https://doi.org/10.1024/1661-4747/a000146.

Wegerer, M., Kerschbaum, H., Blechert, J., & Wilhelm, F. H. (2014). Low levels of estradiol are associated with elevated conditioned responding during fear extinction and with intrusive memories in daily life. *Neurobiology of Learning and Memory, 116*, 145–154. https://doi.org/10.1016/j.nlm.2014.10.001.

Wessa, M., & Flor, H. (2007). Failure of extinction of fear responses in posttraumatic stress disorder: Evidence from second-order conditioning TT - versagen der Extinktion von Furchtreaktionen bei der posttraumatischen Belastungsstörung. Evidenz aus Konditionierungen zweiter Ordnung. *American Journal of Psychiatry, 164* (November), 1684–1692. https://doi.org/10.1176/appi.ajp.2007.07030525.

Wilhelm, F. H., & Peyk, P. (2005). Autonomic nervous system laboratory (ANSLAB) (shareware version). Retrieved from the software repository of the society for psychophysiological research:. www.sprweb.org.

Wilhelm, F. H., Rattel, J. A., Wegerer, M., Liedlgruber, M., Schweighofer, S., Kreibig, S. D., et al. (2017). Attend or defend? Sex differences in behavioral, autonomic, and respiratory response patterns to emotion–eliciting films. *Biological Psychology, 130*, 30–40. https://doi.org/10.1016/J.BIOPSYCHO.2017.10.006.

Xia, J., & Evans, L. H. (2020). Neural evidence that disengaging memory retrieval is modulated by stimulus valence and rumination. *Scientific Reports, 10*(1), 1–11. https://doi.org/10.1038/s41598-020-64404-7.

Zetsche, U., Ehring, T., & Ehlers, A. (2009). The effects of rumination on mood and intrusive memories after exposure to traumatic material: An experimental study. *Journal of Behavior Therapy and Experimental Psychiatry, 40*(4), 499–514. https://doi.org/10.1016/j.jbtep.2009.07.001.