



Registered Report

No persistent attenuation of fear memories in humans: A registered replication of the reactivation-extinction effect



Anastasia Chalkia^{a,b}, Natalie Schroyens^{a,b}, Lu Leng^a,
Niels Vanhasbroeck^a, Ann-Kathrin Zenses^{a,b},
Lukas Van Oudenhove^{b,c} and Tom Beckers^{a,b,*}

^a Centre for the Psychology of Learning and Experimental Psychopathology, Faculty of Psychology & Educational Sciences, KU Leuven, Belgium

^b Leuven Brain Institute, KU Leuven, Belgium

^c Laboratory for Brain-Gut Axis Studies (LaBGAS), Translational Research Centre for Gastrointestinal Disorders (TARGID), Department of Chronic Diseases, Metabolism, and Ageing, KU Leuven, Belgium

ARTICLE INFO

Article history:

Received 16 September 2019

Reviewed 01 December 2019

Revised 24 February 2020

Accepted 14 April 2020

Action editor Robert McIntosh

Published online 18 June 2020

Keywords:

Fear learning

Extinction

Memory updating

Reconsolidation

Memory reactivation

ABSTRACT

It has been proposed that memory retrieval can destabilize consolidated memories, after which they need to be reconsolidated in order to be retained. The presentation of relevant information during memory reconsolidation could then result in the modification of a destabilized memory trace, by allowing the memory trace to be updated before being reconsolidated. In line with this idea, Schiller et al. (2010) have demonstrated that memory retrieval shortly before extinction training can prevent the later recovery of conditioned fear responding that is observed after regular extinction training. Those findings have been the subject of considerable controversy, due in part to theoretical reasons but also due to a number of failures to obtain similar results in conceptual replication attempts. Here, we report the results of a highly powered, direct, independent replication of the critical conditions of Schiller et al. (2010, Experiment 1). Due to misrepresentation of the exclusion criteria in the original Schiller et al. (2010) report, data collection was considerably delayed. When we eventually managed to attain our pre-registered sample size, we found that we could not observe any benefit of reactivation-extinction over regular extinction training in preventing recovery of conditioned fear. The results of the present study, along with the mixed findings in the literature and the misreporting in Schiller et al. (2010), give cause to question whether there is robust evidence that reactivation-extinction prevents the return of fear in humans.

© 2020 Elsevier Ltd. All rights reserved.

* Corresponding author. Faculty of Psychology & Educational Sciences, KU Leuven, Tiensestraat 102 – box 3712, 3000 Leuven, Belgium
E-mail address: tom.beckers@kuleuven.be (T. Beckers).

<https://doi.org/10.1016/j.cortex.2020.04.017>

0010-9452/© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

It has been known for almost 50 years that upon retrieval, memory traces can become sensitive to amnesic interventions that interfere with their later expression (Ricchio, Millin, & Bogart, 2006; Sara, 2000). In fear conditioning, when animals are first trained to be afraid of a cue, e.g., by the administration of foot shock in the presence of that cue, a brief re-exposure to said cue the next day will retrieve the fear memory. When retrieval is immediately followed by the administration of electroconvulsive shock to the brain, later expression of fear when the animals are presented with that same cue will be prevented (Misanin, Miller, & Lewis, 1968). Similar effects can be obtained by the administration of amnesic drugs (e.g., anisomycin; Nader, Schafe, & Le Doux, 2000) or other interventions shortly after memory retrieval (for a review, see Beckers & Kindt, 2017). The standard account for such post-retrieval amnesia holds that memory retrieval can bring a firmly consolidated memory trace back into an active, unstable state, in which it critically depends on de-novo protein synthesis in dedicated brain areas in order to get reconsolidated and return to a stable state (Nader, Schafe, & LeDoux, 2000; Przybylski & Sara, 1997). Interventions that prevent protein synthesis during the time-limited reconsolidation period (such as the administration of electroconvulsive shock or pharmacological agents that impair protein synthesis) interfere with the restabilization of the memory trace, resulting in functional memory loss (for an alternative account of post-retrieval amnesia, see Gisquet-Vierrier & Riccio, 2012; Riccio et al., 2006).

More recently, it has been proposed that the phenomenon of memory destabilization and reconsolidation upon retrieval allows for mnemonic outcomes other than the blunt induction of amnesia. More specifically, it has been argued that the reconsolidation process presents a window of opportunity to update the contents of the memory trace, by presenting novel information after reactivation that is of relevance for the memory. One example in kind is the exploitation of reconsolidation to solidify the effects of extinction training after fear conditioning. In regular fear extinction training, a conditioned stimulus (CS) that was previously paired with an aversive unconditioned stimulus (US) is repeatedly presented on its own, resulting in a gradual decline of fear responding to the CS. An extensive body of research suggests that the decline in conditioned fear responding does not reflect a modification of the original fear memory trace but the creation of an additional, context-dependent extinction memory trace that competes with the fear memory for behavioral control (Bouton, 1993). As a consequence, fear tends to recover whenever the physical or temporal context of extinction training is abandoned (Bouton, 2002; Vervliet, Craske, & Hermans, 2013). Recent findings suggest that such recovery can be prevented by the execution of extinction training during a period of instability of the initial fear memory trace, that is, during the retrieval-induced reconsolidation window. More specifically, Monfils and colleagues demonstrated that if rats were given extinction training within a period of six hours after memory reactivation (reactivation being achieved by the presentation of a single CS, unreinforced), they did not exhibit

any return of fear in later fear memory recovery tests (Monfils, Cowansage, Klann, & LeDoux, 2009). In contrast, rats that received extinction training outside of the reconsolidation window showed a recovery of fear when tested one month later (spontaneous recovery), when tested after the presentation of unsignalled shocks (reinstatement) or when tested in a novel context (renewal) after extinction training. The explanation offered by the authors is that the presentation of extinction training during memory reconsolidation allows for the extinction information to be incorporated in the original fear memory trace, rather than it resulting in the formation of a separate extinction memory trace, thereby permanently modifying or updating the original fear memory and effectively preventing the possibility for the later recovery of fear.

Shortly thereafter, Schiller et al. (2010, Experiment 1) reported similar effects in humans. In a Pavlovian differential fear conditioning procedure, participants first received acquisition using two fear-irrelevant stimuli (CS+ and CS-; pictures of a blue and a yellow square). The CS+ was paired with a mild 200-ms shock to the wrist (US), on 37.5% of the trials, while the CS- was never paired with the US. The following day, participants were allocated to one of three groups before receiving extinction training. The first group was presented with a single unreinforced CS+ trial to reactivate the fear memory ten minutes before the start of extinction training. The second group was presented with the reactivation trial six hours before receiving extinction training. In the third group, the fear memory was not reactivated prior to extinction training. Spontaneous recovery of conditioned fear was assessed 24 h later during a re-extinction test session, in which participants were again presented with unreinforced CS+ and CS- trials. In the group that received extinction training within the putative six-hour window of reconsolidation, spontaneous recovery was not observed, unlike in the other two groups. Likewise, in a subsample of the initial participants tested 12 months later, sensitivity to reinstatement was observed in participants from the latter two groups but not in those from the first group. According to the authors' interpretation, those results suggest that also in humans, conducting extinction training during a period of reconsolidation may permanently abolish conditioned fear, by inducing an update of the previously acquired fear memory.

The findings of Schiller et al. (2010), and the accompanying notion of persistent modification of emotional memory through a reactivation-extinction logic, have attracted tremendous attention, not in the least because of their clinical potential for the treatment of post-traumatic stress disorder (PTSD) and other emotional memory disorders (Beckers & Kindt, 2017). However, they have also generated considerable controversy, in part due to theoretical reasons. The reactivation-extinction procedure essentially entails performing extinction training with a 10-min interval between the first and the second extinction trial, and the findings are not necessarily consistent with other observations regarding the spacing of extinction training (Auber, Tedesco, Jones, Monfils, & Chiamulera, 2013; Baker, McNally, & Richardson, 2013). Equally importantly, attempts at conceptual replication of these findings have yielded mixed results. In addition to two published successful replications from the same lab

(Schiller, Kanen, LeDoux, Monfils, & Phelps, 2013; Steinfurth et al., 2014), a few other groups have been able to demonstrate similar results in human aversive learning paradigms (Agren et al., 2012; Agren, Björkstrand, & Fredrikson, 2017; Asthana et al., 2016; Björkstrand et al., 2016; Johnson & Casey, 2015; Liu et al., 2014; Oyarzún et al., 2012).¹ Of note, those studies were not exact replications of the original study by Schiller et al. (2010), as they exhibited variations in US selection procedures, US intensity, reinforcement schedules and others. Other labs have failed to replicate the Schiller et al. (2010) findings (e.g., Fricchione et al., 2016; Golkar, Bellander, Olsson, & Ohman, 2012; Kindt & Soeter, 2013; Klucken et al., 2016; Meir Drexler et al., 2014; Soeter & Kindt, 2011; Warren et al., 2014).² Those failed attempts deviated in important ways from the protocol of the original study as well, e.g., by the addition of fear-potentiated startle (FPS) measurements (necessitating the repeated presentation of loud noise bursts during acquisition, reactivation, extinction and testing), the trial-by-trial assessment of US expectancies, or the use of fear-relevant stimuli during conditioning. Independent replications of the related findings in rats by Monfils et al. (2009) have proven equally challenging, with some labs reporting successes (e.g., Clem & Huganir, 2010; Flavell, Barber, & Lee, 2011; Jones, Ringue, & Monfils, 2013; Olshavsky, Jones, Lee, & Monfils, 2013; Piñeyro, Ferrer Monti, Alfei, Bueno, & Urcelay, 2013; Rao-Ruiz et al., 2011) but others reporting failures (e.g., Chan, 2014; Chan, Leung, Westbrook, & McNally, 2010; Costanzi, Cannas, Saraulli, Rossi-Arnaud, & Cestari, 2011; Ishii et al., 2012; Luyten & Beckers, 2017; Pérez-Cuesta & Maldonado, 2009).

A recent meta-analysis of reactivation-extinction studies in humans and non-human animals suggests that there is no significant benefit of post-retrieval fear extinction over regular fear extinction training in animals, but that there is a small-to-moderate benefit in humans (Kredlow, Unger, & Otto, 2016). The opposite picture emerges for appetitive learning. Whereas the same meta-analysis yields a large and significant effect for preventing the return of appetitive memories in animals (Kredlow et al., 2016), results in humans are so far sparse and heterogeneous (e.g., Bakkour, Schonberg, Hover, & Poldrack, 2015; Xue et al., 2012).

While some have argued that the mixed findings discussed above cast doubts on the very possibility of fear memory updating through a reactivation-extinction procedure, others have suggested that the divergent results may be attributed to untested boundary conditions that may have been fulfilled in the original reports and the successful replications, but not in the failed attempts at conceptual replication (Auber et al., 2013; Kredlow et al., 2016; Schiller & Phelps, 2011). The divergence in procedures between the original reports and the subsequent replication attempts indeed allows for such an argument. For instance, the addition of noise probes or online

US expectancy ratings or the use of fear-relevant CSs may lead to qualitative changes in fear learning that somehow present a challenge for the reactivation-extinction procedure (e.g., by making learning more explicit in nature). Of note, Kredlow et al. (2016) also conducted moderator analyses on the studies included in their meta-analysis, and indeed identified a number of potential moderators for the benefit of reactivation-extinction over regular extinction (e.g., number of acquisition trials, shock duration, inclusion of expectancy ratings). Arguably, however, these potential moderators should be considered tentative, given the limited number and relatively small sample size of studies included in the meta-analysis, the fact that potential moderators identified there have not been systematically tested, and the possibility, as with any meta-analysis, of publication bias. Trim and fill analyses were conducted by Kredlow et al. (2016) to adjust reactivation-extinction effect sizes for publication bias, but no adjustment was deemed necessary. Nevertheless, they observed a tendency for larger effect sizes to be associated with smaller sample size. As such, the argument for boundary conditions as an explanation for the failed replication attempts may be considered premature in the absence of a direct, well-powered, independent replication of the original findings. If the original findings would prove to be robust in a direct, independent replication, a systematic assessment of critical boundary conditions would seem warranted and timely indeed. On the other hand, if the original findings would not be readily reproducible in a direct, independent replication, an empirical search for theoretically inspired boundary conditions may be misguided. Therefore, the aim of the present replication report was to perform an exact replication of the critical conditions and manipulations reported by Schiller et al. in their seminal study (Schiller et al., 2010, Experiment 1).

2. Materials and methods

2.1. Participants

Two hundred forty-six participants within the age range of 17–45 years were recruited through the KU Leuven experiment management system. All participants were screened for the following medical exclusion criteria: pregnancy, current or past heart or cardiovascular disease, lung disease or lung problems, neurological problems, any psychiatric diagnosis (in the past or present), other severe medical problems, current use of anticholinergic or anti-adrenergic medication, benzodiazepines, or non-stable doses of psychotropic medication, request from the general practitioner to avoid stressful situations, any type of electronic implant, and pain or problems located at the hand or wrist ($n = 2$ excluded).³ In line with Schiller et al. (2010), participants that did not show successful fear acquisition (first session) or extinction of fear (second session) in skin conductance responses (SCR) were excluded.

¹ Since the time in-principle acceptance of the current Registered Report was granted, a few more successful conceptual replications have been published (Björkstrand et al., 2017; Hu et al., 2018; Telch et al., 2017).

² Since the time in-principle acceptance of the current Registered Report was granted, a few more failed conceptual replications have been published (Kredlow et al., 2018; Kroes, Dunsmoor, Lin, Evans, & Phelps, 2017; Maples-Keller et al., 2017).

³ Schiller et al. (2010) did not report any medical exclusion criteria. The medical exclusion criteria used in the present study are based on the guidance of our ethics committee and a comment raised during Stage 1 peer review.

In the initial Stage 1 manuscript, we listed criteria for successful acquisition and extinction that were taken from Schiller et al. (2010). However, when we started data collection, we noticed that application of those criteria resulted in a much higher rate of exclusions than what was reported there. Subsequent correspondence with the authors and inspection of the original data underlying Schiller et al. (2010) revealed that the exclusions they performed deviated in important ways from the criteria as initially reported.⁴ We were presented with a corrected set of exclusion criteria, which the authors confirmed to the editors of *Cortex* and us to represent an accurate representation of their original exclusions, on September 21, 2017 (Daniela Schiller & Elizabeth Phelps, personal communication) (see Figure S1 in the Supplementary Information for a flowchart detailing the timeline of events). It is that set of criteria that was applied here. In particular, with respect to acquisition, participants were included if the differential CS+/CS- response was above an individually-standardized cut-off ($.1 \mu\text{S}$ divided by the participant's average square-rooted US response during acquisition) on either all of acquisition averaged, the first half of acquisition trials, the last half of acquisition trials, or the very last trial of acquisition; or if the increase in differentiation from the first trial of acquisition to the last trial of acquisition exceeded this cut-off value; or if the increase in differentiation from the first half to the last half of acquisition trials exceeded this cut-off value. With respect to extinction, participants were included if the CS+/CS- difference was below this same cut-off for either all extinction trials averaged, the last half of extinction trials, or the last trial of extinction; or if the decline in differential responding from the first trial to the last trial of extinction exceeded this cut-off value; or if the decline in differential responding from the first half to the last half of

⁴ According to the *Nature* report, exclusions were made on the basis of differential (CS+/CS-) responding in the second half of acquisition training and the second half of extinction training, respectively. Specifically, participants were excluded if differential responding at the end of acquisition was in the opposite direction (CS- > CS+) or smaller than $.1 \mu\text{S}$, and similarly, if it was in the opposite direction (CS+ > CS-) or larger than $.1 \mu\text{S}$ at the end of extinction.

⁵ On July 26, 2018 an addendum to the original study was published, detailing the exclusions and exclusion criteria that had previously been misreported. The criteria reported in the addendum are similar to the ones confirmed to us by the authors in September 2017, although they do deviate in minor aspects. In particular, two of the acquisition inclusion criteria confirmed to us were not eventually retained in the addendum: 1) Differential responding above cut-off for all of acquisition averaged, and 2) An increase in differentiation from the first half to the last half that exceeds cut-off. One of the extinction criteria was also not retained, namely, a differentiation below cut-off on all extinction trials averaged. Critically, the criteria listed in the addendum are reported to characterize the exclusions made in the original study rather than being the actual basis for the exclusions. In practice, inclusion or exclusion of participants in Schiller et al. (2010) was reported to be based on qualitative SCR assessment during data collection and analysis. Still, the criteria listed in the addendum do not characterize the dataset comprehensively, as 3 of the included participants in Schiller et al. (2010) failed to meet the inclusion criteria and 9 of the excluded participants actually did meet the inclusion criteria described in the addendum.

extinction exceeded this cut-off value.⁵ Participants were excluded if they failed to meet any of the stated inclusion criteria for acquisition ($n = 82$) or if they failed to meet any of the inclusion criteria for extinction ($n = 15$). Another 23 participants were excluded for various other reasons not related to the protocol: not attending the second or third testing sessions ($n = 9$), technical problems or experimenter error resulting in loss of data ($n = 8$), technical malfunction ($n = 2$), foreign-speaking participants ($n = 2$), tested outside of the ± 2 h window relative to the previous day ($n = 1$), interruption of procedure by fire drill ($n = 1$). Excluded participants were replaced until the predetermined sample size was reached, thus, the final sample included 124 participants ($n = 62$ per group; 105 females) with an average age of 20.48 years ($SD = 3.68$). The study was granted full ethical approval by the Social and Societal Ethics Committee of KU Leuven and all procedures were carried out in accordance with the Declaration of Helsinki. All participants gave written informed consent before the start of the study and were reimbursed with 30 euros or partial course credit for their participation.

2.2. Stimuli and behavioral manipulation

A three-day Pavlovian cued fear conditioning paradigm was employed (see Fig. 1). Sessions were conducted on three consecutive days, about 24 h apart (± 2 h). Two pictures of colored squares (yellow and blue) served as the CSs and a mild electric shock to the wrist served as the US. One CS, designated the CS+, was paired with shock on 37.5% of the trials, while the other CS, the CS-, was never paired with shock. Allocation of CSs was counterbalanced across participants. CS pictures were presented for 4 sec, with a random 10–12 sec inter-trial interval (ITI). Shocks were delivered 3800 msec after CS onset and co-terminated with CS offset.

2.3. Psychophysiological measures

2.3.1. Skin conductance response (SCR)

SCR was measured continuously at 200 Hz using a pair of disposable, pre-gelled 8-mm Ag/AgCl electrodes (Biopac Systems, Goleta, California) attached to the index and middle fingers of the left hand, between the first and second phalanges. The electrodes were connected to an isolated skin conductance coupler (LabLinc v71-23, Coulbourn Instruments, Holliston, Massachusetts). The skin conductance module was further connected to a 16-bit AD-converter (National Instruments NI-6221, Austin, Texas), which digitized the raw, analogue SCR signal.

2.3.2. Electrical stimulation

Mild electric shocks to the right inner wrist were delivered through a stimulating bar electrode, composed of two 8-mm stainless steel electrodes with an inter-electrode distance of 30 mm (Digitimer, Hertfordshire, UK). All shocks were given for 200 msec, at 50 pulses per second, and shock delivery was controlled by a constant current stimulator (DS7A, Digitimer, Hertfordshire, UK). Using a shock work-up procedure, the shock intensity level was determined for each participant individually. Specifically, each subject was given a very low stimulus at first (1 mA), and the intensity was gradually

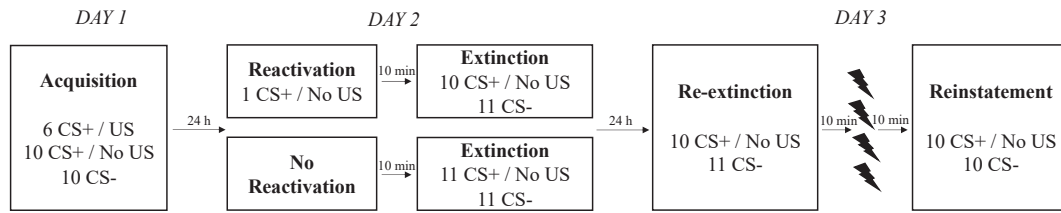


Fig. 1 – Overview of experimental design.

increased in steps of 1 mA until a level was reached that was “uncomfortable, but not painful”. Once selected, the intensity remained the same throughout the experiment, and was not re-calibrated again.

2.4. Subjective assessments

In the original study by Schiller et al. (2010), no subjective assessments were reported. In order to obtain a baseline measure of anxiety, and to confirm that there were no significant differences between groups, the trait version of the State and Trait Anxiety Inventory (STAI-T; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1977) and the short form of the Fear of Pain Questionnaire (FPQ; McNeil & Rainwater, 1998) were administered at the end of the experiment, as to not interfere with the original protocol.

2.5. Procedure

Following attachment of the electrodes, each day began with a 5-min SCR habituation phase for all groups, during which SCR was recorded while no stimuli were presented (see Schiller, Raio, & Phelps, 2012). Two counterbalanced, pseudo-randomized orders of trial presentations were predetermined and used throughout the experiment, with the restriction that no trial type was presented more than two times in a row.⁶

2.5.1. Acquisition

During the initial session, all participants underwent fear acquisition. After electrode placement, the first session began with the shock work-up procedure. Afterwards, and upon completion of the habituation phase (see above), participants were instructed to keep their eyes on the screen, breathe naturally, and refrain from any unnecessary movements. They were informed that they would see some visual stimuli on the screen, and that they would receive occasional shocks. Their task was to focus on the relationship between the images and the shock presentations. Conditioning consisted of 16 CS+ trials (6 reinforced, 10 unreinforced) and 10 CS- trials.

⁶ For one of the two trial orders, an error led to the presentation of 3 CS+ trials in a row, midway through the reinstatement testing phase. However, only the first trial of reinstatement testing was used for the analyses, so this error had no influence on the results or their interpretation.

2.5.2. Reactivation and extinction

Before the start of the second session, participants were allocated randomly to one of two conditions: 1) *Reactivation*, or 2) *No Reactivation*.⁷ Participants in the reactivation group were instructed as on Day 1, and received a single, unreinforced CS+ trial, followed by a 10-min break. Participants in the no reactivation group did not receive a reactivation trial, but proceeded directly to a 10-min break, during which all participants watched a preselected Simpsons episode (see Schiller et al., 2012). All electrodes remained attached (Schiller et al., 2012), but participants were informed explicitly that no shock would be administered during the break.

After the break, both groups proceeded immediately to extinction training. For the reactivation group, extinction consisted of 10 unreinforced CS+ trials and 11 unreinforced CS- trials. For the no reactivation group, extinction consisted of 11 CS+ and 11 CS- trials, thus equating the number of CS presentations across groups.

2.5.3. Spontaneous recovery, re-extinction and reinstatement

At the start of the third session, all participants were instructed as on Day 1 and presented with 10 unreinforced CS+ trials and 11 CS- trials in order to: 1) assess spontaneous recovery of fear, and 2) extinguish any remaining conditioned responding. As can be noticed above, during re-extinction one CS- trial more was presented than there were CS+ trials. The additional CS- trial was always the very first trial; in line with Schiller et al. (2010), it was discarded from analysis to remove the influence of orienting. After the first trial, the order of the stimuli was counterbalanced between CS+/CS- across all participants. Following re-extinction, all participants advanced to a 10-min break, where they were asked to sit quietly in the experimental room without doing anything. All electrodes remained attached, but participants were informed explicitly that no shocks would be administered during the break.

⁷ The study of Schiller et al. (2010) included an additional group, in which memory reactivation was performed prior to extinction training as well but with a 6-h interval between reactivation and extinction training, which is hypothesized to exceed the reconsolidation window. This condition has since been omitted from most replication studies, and was also not included here, because no differences are expected between this group and the no reactivation group; the critical result to replicate is the significant difference in recovery of conditioned fear responding between the regular extinction group and the reactivation-extinction group included here.

Schiller et al. (2010) re-invited participants to the lab one year later to test for reinstatement of fear. If the amnesia induced by their procedure is indeed invulnerable to reinstatement up to one year later, it should surely exhibit insensitivity to reinstatement on a shorter time scale. Therefore, we tested for reinstatement at the end of the third session, avoiding the attrition observed by Schiller et al. (2010) and allowing us to distinguish the presence of genuine reinstatement from the mere spontaneous recovery that may occur with the extended passage of time. This provided an additional opportunity to observe differences between reactivation-extinction and regular extinction training in preventing recovery of fear should there not have been any difference at the start of the spontaneous recovery test. Of note, this test of reinstatement did not aspire to directly replicate the methods of the reinstatement test of Schiller et al. (2010) and should be regarded as a supplementary test; the main measure of interest for the replication test was the degree of spontaneous recovery at the start of the test session. To test for reinstatement, following the 10-min break, four unsignalled shocks were delivered, with a random 10–12 sec ITI. After another 10-min break where all participants watched the same Simpsons episode (see Steinfurth et al., 2014), a test of reinstatement was conducted, consisting of the presentation of 10 unreinforced CS+ trials and 10 unreinforced CS- trials. The session concluded with the administration of the STAI-T and the FPQ-SF.

2.6. Analyses

2.6.1. Processing of psychophysiological data

Processing of SCR data was completed offline, with two separate approaches, using MATLAB 8.4 (The Mathworks Inc., Natick, Massachusetts). First, and as the primary processing approach, raw SCR data were processed according to the methods reported in Schiller et al. (2010). SCRs to the CSs and USs were determined separately, as the dependent variables for conditioned (CR) and unconditioned (UR) responding, respectively. SCRs were determined by taking the base to peak difference in the .5–4.5 sec window following stimulus onset (CS and US onset separately). A minimum response criterion of .02 μ S was used, and any responding below that was scored as 0 and kept in the analysis. Raw SCR data were square-root transformed, and those values were further range-corrected for each participant by dividing the value by that participant's mean square-root transformed US response. Concurrently, a second processing method was applied, for a second set of planned analyses. The mean baseline response in a 2-sec period before CS onset was subtracted from the highest skin conductance response obtained in the .5–4.5 sec window following CS onset. To standardize the data, means and standard deviations from the first testing session were used to calculate within-participant z-scores. This is a method that is more typically applied in the processing of SCR data in our lab and elsewhere (Milad, Orr, Pitman, & Rauch, 2005; Pitman & Orr, 1986).

2.6.2. Statistical analyses

The primary and critical set of statistical analyses copied the analyses reported by Schiller et al. (2010). Only SCR responses

during unreinforced trials were included in the analysis, excluding the reinforced CS+ trials from the acquisition phase, to exclude an influence of unconditioned (shock-elicited) responding. A differential fear score was calculated individually for each participant, by subtracting the values for CS- from those for CS+ for corresponding trials. These differential scores were then averaged across the first and the second half (first 5 vs last 5 trials) of each phase (acquisition and extinction) and subjected to a repeated measures analysis of variance (rm-ANOVA) with Group as a between-subjects factor and Time (early vs late phase) as a within-subjects factor. According to the first author of the initial report (Daniela Schiller, personal communication, September 13, 2017), the first CS- trial of extinction was to be removed from the analysis, as were the first CS+ trial of extinction in the no reactivation group and the CS+ reactivation trial in the reactivation group, leading to a total of 10 CS+ and 10 CS- extinction trials for both groups. To assess the decrease of fear responding from acquisition to extinction an rm-ANOVA was used with Group as a between-subjects factor and Time (late acquisition, last trial of extinction) as a within-subjects factor. Follow-up t-tests were conducted for each group separately to assess successful acquisition (comparing CS+ vs CS- responding during late acquisition), extinction (comparing CS+ vs CS- responding on the last trial of extinction), and the decrease in fear responding (comparing differential responding between late acquisition and the last trial of extinction). Spontaneous recovery was assessed with an rm-ANOVA examining the change in differential responding from the last trial of extinction to the first trial of re-extinction.⁸ Sensitivity to reinstatement was assessed using an rm-ANOVA comparing the change in differential responding from the last trial of re-extinction to the first trial of reinstatement testing. Follow-up t-tests were conducted to evaluate the presence of spontaneous recovery and reinstatement within each group. The critical test to evaluate a successful replication of Schiller et al. (2010) then involved an independent samples t-test that compared the changes in differential responding from the end of extinction to the start of re-extinction (for spontaneous recovery) between the groups. Additionally, and as a secondary outcome, reinstatement was compared between the two groups through an independent samples t-test comparing the change in differential responding from the end of re-extinction to the start of reinstatement testing.

Furthermore, a second set of analyses was also conducted. For these analyses, differential scores were not calculated, but instead, responses for CS+ were compared to those for CS-. Thus, using this method, and mimicking the set of analyses discussed above, Cue (CS+ vs CS-) was included as an

⁸ Schiller et al. (2010) reported to have assessed spontaneous recovery using an rm-ANOVA comparing parts of the re-extinction phase (first 4 trials vs subsequent 4 trials) between groups. It is unclear why this ANOVA was conducted by the original authors and then followed up with t-tests comparing the last trial of extinction to the first trial of re-extinction. Thus, the rm-ANOVA examining spontaneous recovery reported in the current study deviates from what was reported in Schiller et al. (2010) but it is line with the follow-up analysis reported in the original study.

additional within-subjects factor, to allow examination of the specific course of SCR responding throughout the three days of the study. Outliers were defined for each day (z -score > 3) and replaced by linear trend at point. Greenhouse–Geisser corrections were applied in case of violation of sphericity. An alpha level of .05 was set for all analyses, which were conducted using JASP (version 0.9.2) (JASP Team, 2019).

2.6.3. Power analysis

To determine the number of participants required to have sufficient power to replicate the effects reported by Schiller et al. (2010), an a priori power analysis was conducted using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). Based on the overall estimated effect size for the difference in spontaneous recovery between reactivation-extinction and regular extinction groups of $g = .53$ in human fear memory experiments included in the meta-analysis of Kredlow et al. (2016), and adopting an alpha level of .05 (one-tailed), a sample size of 62 participants per group yielded a power of .9 to replicate the benefit of reactivation-extinction over regular extinction in preventing spontaneous recovery in the critical independent t-test described above [the effect size for the actual experiment replicated here (Schiller et al., 2010, Experiment 1) was $g = .73$]. A sample size of 62 per group also clearly fulfills the suggestion that replication efforts include 2.5 times the original sample size, offered by Simonsohn (2015) on the basis of his “small telescopes” approach (which basically entails that if an effect is sufficiently large to have a minimal chance of being detected using a sample size of x , a sample of 3 times x should provide sufficient power to detect it in a replication irrespective of what the actual effect size is).

3. Results

3.1. Participant characteristics

The reactivation group (R) did not differ significantly from the no reactivation group (NR) in terms of age (R: $M = 20.71$, $SD = 4.39$, NR: $M = 20.24$, $SD = 2.80$; $t(122) = .71$, $p = .48$, $d = .12$), gender distribution ($\chi^2(1, N = 124) = 1.55$, $p = .21$), or education level ($\chi^2(2, N = 124) = 1.95$, $p = .38$). Further, the two groups did not differ in terms of selected US intensity (R: $M = 5.76$, $SD = 2.65$, NR: $M = 6.61$, $SD = 3.13$; $t(122) = 1.64$, $p = .10$, $d = .30$), or their subjective rating of the US intensity (R: $M = 8.09$, $SD = 1.22$, NR: $M = 7.90$, $SD = 1.61$; $t(122) = .72$, $p = .47$, $d = .13$). Trait anxiety (STAI-T) was also not significantly different between the groups (R: $M = 38.44$, $SD = 8.19$, NR: $M = 37.36$, $SD = 7.88$; $t(122) = .75$, $p = .46$, $d = .13$), nor was their reported fear of pain (FPQ) (R: $M = 21.47$, $SD = 4.47$, NR: $M = 20.84$, $SD = 4.62$; $t(122) = .77$, $p = .44$, $d = .14$).

3.2. Schiller et al. (2010) analyses

Participants exhibited successful acquisition of fear on the first day (main effect of time, $F(1, 122) = 5.14$, $p = .025$, $\eta_p^2 = .04$), with no significant differences between the groups (group \times

time, $F(1, 122) = 2.94$, $p = .089$, $\eta_p^2 = .02$). Follow-up paired-sample t-tests for each group separately confirmed that by the last half of acquisition, participants in both groups responded more strongly to the CS+ than to the CS- (R: $M_{CS+} = .52$, $SD_{CS+} = .18$, $M_{CS-} = .39$, $SD_{CS-} = .11$; $t(61) = 9.21$, $p < .001$, $d = 1.17$, NR: $M_{CS+} = .52$, $SD_{CS+} = .13$, $M_{CS-} = .39$, $SD_{CS-} = .09$; $t(61) = 10.01$, $p < .001$, $d = 1.27$). The same pattern emerged during extinction training, but in the opposite direction, as participants in both groups effectively extinguished their differential conditioned responding during the second day of the study (main effect of time, $F(1, 122) = 66.27$, $p < .001$, $\eta_p^2 = .35$; group \times time, $F(1, 122) = 3.11$, $p = .08$, $\eta_p^2 = .03$). On the last trial of extinction, the CS+ no longer elicited significant fear responding in comparison with the CS-, either in the reactivation group ($M_{CS+} = .38$, $SD_{CS+} = .27$, $M_{CS-} = .37$, $SD_{CS-} = .25$; $t(61) = .38$, $p = .70$, $d = .05$) or the no reactivation group ($M_{CS+} = .33$, $SD_{CS+} = .21$, $M_{CS-} = .31$, $SD_{CS-} = .20$; $t(61) = .52$, $p = .61$, $d = .07$). Fear responding decreased from late acquisition to the last trial of extinction (main effect of time, $F(1, 122) = 38.01$, $p < .001$, $\eta_p^2 = .24$), across groups (group \times time, $F(1, 122) < 1$). Follow-up paired-sample t-tests for each group separately confirmed these findings, showing fear reduction in the reactivation group (late acquisition: $M = .14$, $SD = .12$, last trial of extinction: $M = .01$, $SD = .18$; $t(61) = 4.42$, $p < .001$, $d = .56$) and the no reactivation group (late acquisition: $M = .13$, $SD = .10$, last trial of extinction: $M = .01$, $SD = .19$; $t(61) = 4.30$, $p < .001$, $d = .55$).

Contrary to the outcome of Schiller et al. (2010), spontaneous recovery was observed across groups on the third day of the study (main effect of time, $F(1, 122) = 44.40$, $p < .001$, $\eta_p^2 = .27$; group \times time, $F(1, 122) < 1$) (see Fig. 2). Differential responding significantly increased from the last trial of extinction to the first trial of re-extinction in the reactivation group (extinction: $M = .01$, $SD = .18$, re-extinction: $M = .22$, $SD = .28$; $t(61) = 4.96$, $p < .001$, $d = .63$) and in the no reactivation group (extinction: $M = .01$, $SD = .19$, re-extinction: $M = .17$, $SD = .24$; $t(61) = 4.46$, $p < .001$, $d = .57$). We also observed sensitivity to reinstatement, with an increase in differential responding from the last trial of re-extinction to the first trial of reinstatement testing (main effect of time, $F(1, 122) = 48.60$, $p < .001$, $\eta_p^2 = .29$; group \times time, $F(1, 122) < 1$). Follow-up paired-sample t-tests confirmed these effects in both the reactivation group (re-extinction: $M = .06$, $SD = .19$, reinstatement: $M = .27$, $SD = .27$; $t(61) = 5.65$, $p < .001$, $d = .72$) and the no reactivation group (re-extinction: $M = .03$, $SD = .18$, reinstatement: $M = .24$, $SD = .27$; $t(61) = 4.41$, $p < .001$, $d = .56$).

To conduct the critical independent samples t-test necessary for evaluating a successful replication of Schiller et al. (2010), we calculated an index of recovery by subtracting the differential responding on the last trial of extinction from the differential responding on the first trial of re-extinction. Likewise, we calculated an index of reinstatement by subtracting the differential responding on the last trial of re-extinction from the differential responding on the first trial of reinstatement testing. The two groups did not significantly differ in their degree of recovery (R: $M = .21$, $SD = .34$, NR: $M = .16$, $SD = .28$; $t(122) = .93$, $p = .35$, $d = .17$), nor in their

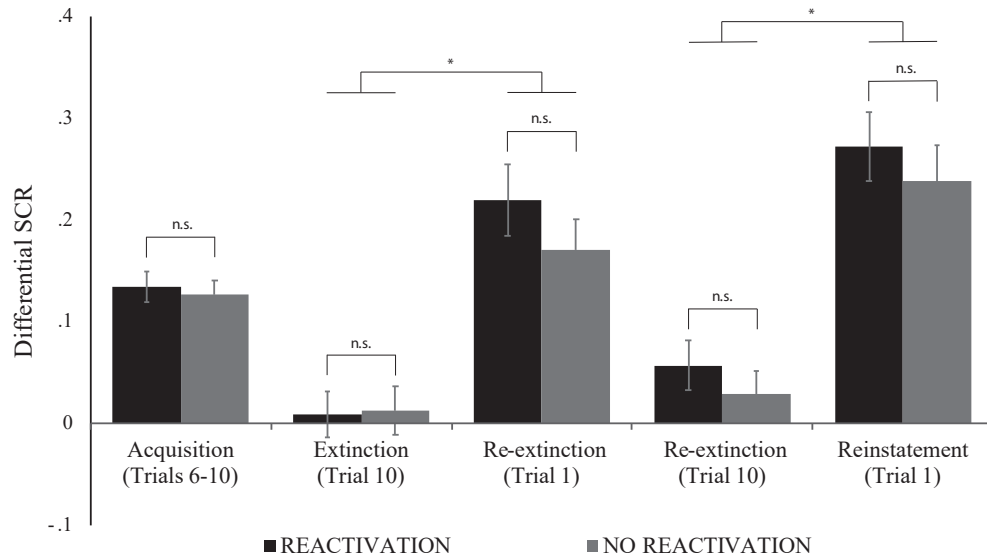


Fig. 2 – Mean differential SCRs (CS+ minus CS-) during acquisition (last half), extinction (last trial), re-extinction (first trial; last trial), and reinstatement (first trial). Error bars represent standard error of the mean. * $p < .001$.

degree of reinstatement (R: $M = .22$, $SD = .30$, NR: $M = .21$, $SD = .38$; $t(122) = .09$, $p = .93$, $d = .02$). Note that if anything, recovery was numerically larger in the reactivation group than in the no reactivation group.

Thus, in contrast with Schiller et al. (2010), we did not observe a benefit of reactivation-extinction over regular extinction in preventing the spontaneous recovery or reinstatement of fear (see Fig. 3 for CS+/CS- responding per trial, across all phases).

3.3. Analyses using z-scores

Using our alternative processing method, once again we observed successful acquisition of fear across groups, with no significant between-group differences (cue \times time, $F(1, 122) = 4.08$, $p = .045$, $\eta_p^2 = .03$; group \times cue \times time, $F(1, 122) = 1.71$, $p = .19$, $\eta_p^2 = .01$). In the last half of acquisition, participants exhibited higher SCRs to the CS+ than to the CS- in both groups (R: $M_{CS+} = -.12$, $SD_{CS+} = .39$, $M_{CS-} = -.69$,

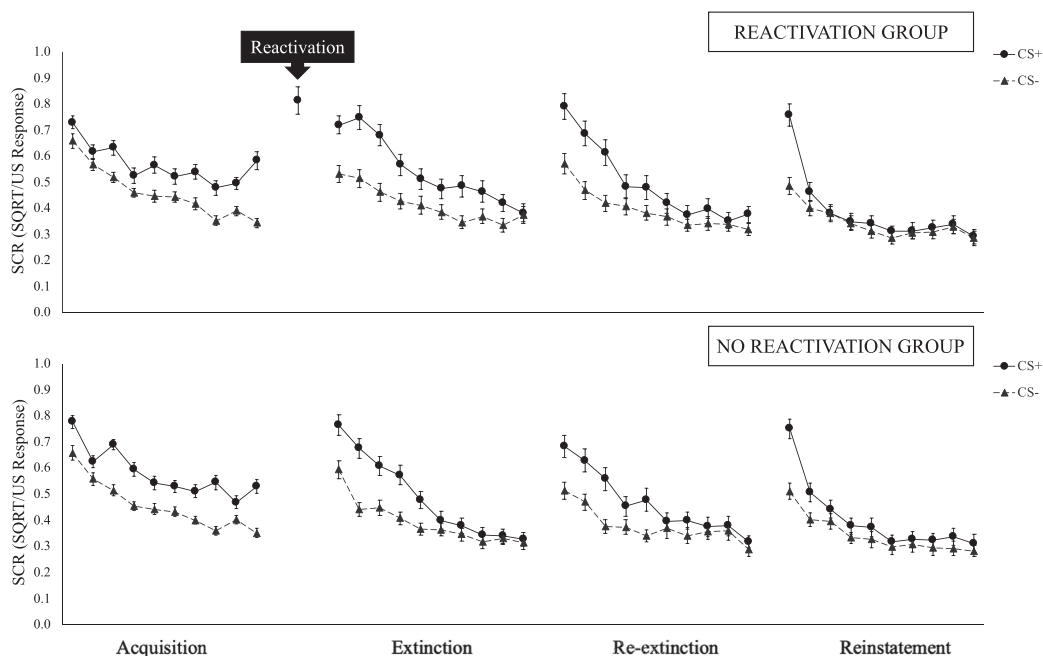


Fig. 3 – Mean SCR per trial, across all experimental phases. Error bars represent standard error of the mean.

$SD_{CS-} = .29$; $t(61) = 8.00$, $p < .001$, $d = 1.02$, NR: $M_{CS+} = -.13$, $SD_{CS+} = .34$, $M_{CS-} = -.71$, $SD_{CS-} = .27$; $t(61) = 9.15$, $p < .001$, $d = 1.16$). Differential responding again extinguished from the first to the last half of extinction across groups (cue \times time, $F(1, 122) = 54.52$, $p < .001$, $\eta_p^2 = .31$; group \times cue \times time, $F(1, 122) = 1.16$, $p = .28$, $\eta_p^2 = .01$). On the last trial of extinction, responding to the CS+ no longer differed from that to the CS- in either the reactivation group ($M_{CS+} = -.34$, $SD_{CS+} = .76$, $M_{CS-} = -.42$, $SD_{CS-} = .70$; $t(61) = .89$, $p = .38$, $d = .11$) or the no reactivation group ($M_{CS+} = -.58$, $SD_{CS+} = .56$, $M_{CS-} = -.65$, $SD_{CS-} = .55$; $t(61) = .91$, $p = .37$, $d = .12$). A reduction of fear from late acquisition to the last trial of extinction was again observed (cue \times time, $F(1, 122) = 44.29$, $p < .001$, $\eta_p^2 = .27$), and while the no reactivation group displayed lower SCR responding in general (main effect of group, $F(1, 122) = 5.12$, $p = .025$, $\eta_p^2 = .04$), the decrease in differential responding did not differ between the groups (group \times cue \times time, $F(1, 122) < 1$).

Spontaneous recovery was detected at the beginning of the third day across groups (cue \times time, $F(1, 122) = 32.35$, $p < .001$, $\eta_p^2 = .21$; group \times cue \times time, $F(1, 122) < 1$). Follow-up paired-sample *t*-tests for each group separately confirmed that while differential responding was absent at the end of extinction for both groups (see above), at the beginning of re-extinction both groups displayed higher responses to the CS+ than to the CS- (R: $M_{CS+} = .89$, $SD_{CS+} = .94$, $M_{CS-} = .02$, $SD_{CS-} = .95$; $t(61) = 4.91$, $p < .001$, $d = .62$, NR: $M_{CS+} = .59$, $SD_{CS+} = .99$, $M_{CS-} = -.15$, $SD_{CS-} = .82$; $t(61) = 5.29$, $p < .001$, $d = .67$). The no reactivation group did again exhibit lower SCR responding in general (main effect of group, $F(1, 122) = 7.87$, $p = .006$, $\eta_p^2 = .06$). Follow-up independent samples *t*-tests showed that in the no reactivation group SCRs to the CS+ and CS- on the last trial of extinction were slightly lower than those of the reactivation group (R: $M_{CS+} = -.34$, $SD_{CS+} = .76$, NR: $M_{CS+} = -.58$, $SD_{CS+} = .56$; $t(122) = 2.05$, $p = .042$, $d = .37$, R: $M_{CS-} = -.42$, $SD_{CS-} = .70$, NR: $M_{CS-} = -.65$, $SD_{CS-} = .55$; $t(122) = 2.01$, $p = .047$, $d = .36$), yet on the first trial of re-extinction the two groups did not significantly differ in terms of CS+ (R: $M = .89$, $SD = .94$, NR: $M = .59$, $SD = .99$; $t(122) = 1.74$, $p = .085$, $d = .31$) or CS- responding (R: $M = .02$, $SD = .95$, NR: $M = -.15$, $SD = .82$; $t(122) = 1.09$, $p = .28$, $d = .20$), even though numerically the no reactivation group again displayed lower SCR.

Lastly, differential reinstatement was observed across groups without significant between-group differences (cue \times time, $F(1, 122) = 69.01$, $p < .001$, $\eta_p^2 = .36$; group \times cue \times time, $F(1, 122) < 1$). Paired-sample *t*-tests for each group separately revealed that CS+/CS- responding was not significantly different at the end of re-extinction for either group (R: $M_{CS+} = -.45$, $SD_{CS+} = .81$, $M_{CS-} = -.58$, $SD_{CS-} = .66$; $t(61) = 1.13$, $p = .26$, $d = .14$, NR: $M_{CS+} = -.54$, $SD_{CS+} = .71$, $M_{CS-} = -.69$, $SD_{CS-} = .60$; $t(61) = 1.55$, $p = .13$, $d = .20$), while at the beginning of reinstatement testing both groups responded more to the CS+ than to the CS- (R: $M_{CS+} = .91$, $SD_{CS+} = .88$, $M_{CS-} = -.40$, $SD_{CS-} = 1.02$; $t(61) = 7.68$, $p < .001$, $d = .98$, NR: $M_{CS+} = .91$, $SD_{CS+} = 1.01$, $M_{CS-} = -.32$, $SD_{CS-} = .88$; $t(61) = 7.54$, $p < .001$, $d = .96$).

In sum, using a different approach for SCR processing (see Figure S2 for trial-by-trial data; Supplementary Information), we were again not able to obtain any evidence for an advantage of reactivation-extinction over regular extinction in preventing the recovery of fear.

4. Discussion

The present study aimed to directly replicate the critical conditions of Schiller et al. (2010, Experiment 1), in order to gauge the robustness of the reactivation-extinction effect in a highly powered, preregistered test. In line with the original study, all included participants successfully acquired and extinguished differential conditioned responding over the first two testing sessions. Yet, despite following the protocol of Schiller et al. (2010) in detail, we observed spontaneous recovery of fear and susceptibility to reinstatement in both the reactivation and no reactivation groups, without any significant between-group differences. Below we discuss our findings in relation to the original results as well as others, and what they imply for future research on the reactivation-extinction effect, taking into account a recently published addendum to the *Nature* report (Schiller et al., 2018).

For nearly a decade, the reactivation-extinction effect has attracted tremendous attention from memory researchers, who have invested considerable resources trying to reproduce it in the lab and to advance it to clinical translation. By now, the variability in success of published conceptual replications is striking. There are about as many successful as failed replication attempts, in humans and non-human animals alike. It has been argued that slight procedural differences and/or boundary conditions might be responsible for those mixed findings (for a recent review see Zuccolo & Hunziker, 2019). However, for such an argument to be valid, successful demonstrations of the effect should at least be upheld in direct replication. The present study shows that this is not the case for the seminal and most influential demonstration of this effect in humans to date. Moreover, the current failure to demonstrate the effect should arguably carry more weight than the initial successful demonstration by Schiller et al. (2010), for a number of reasons. First, the current study involved a considerably larger sample size than the original study (or any of the published conceptual replication attempts), which renders the current results more reliable. Second, the analyses reported here were fully specified in advance and preregistered, removing researchers' degrees of freedom that may reduce the validity of inferential statistics. Finally, and surprisingly perhaps, the alleged benefit of reactivation-extinction over regular extinction training in preventing recovery of fear that Schiller et al. (2010) claimed to have observed was actually dependent on qualitative experimenter evaluations that determined participant exclusion for the original study (see the accompanying Verification Report for details; Chalkia, Van Oudenhove, & Beckers, 2020).

In this respect, it is noteworthy that eight years after the publication of Schiller et al. (2010), an addendum was published (Schiller et al., 2018) to clarify the participant exclusion criteria and the number of excluded participants, which had been misrepresented in the original report. The addendum documents that the actual sample included about 75% more participants ($N = 126$) than were originally reported ($N = 71$). It also contains a list of exclusion criteria (roughly corresponding to the ones supplied to us in September 2017 and used in the current study) that are stated to "characterize the final data set". However, those criteria were not applied strictly, as

arbitrary decisions were made to include participants that should have been excluded by those criteria, and to exclude participants that should have been included. Critically, the observation of a nominal benefit of reactivation-extinction over regular extinction turns out to be highly dependent on those exclusions, which were based on qualitative data evaluations by the experimenters during the process of data collection and analysis. Further, such qualitative evaluations were also made for the data of participants invited for the 1-year follow-up test of reinstatement, and for the data obtained in Schiller et al. (2010; Experiment 2), where 52 out of 70 participants were reportedly excluded (Schiller et al., 2018) (see Chalkia et al., 2020, for details). The qualitative nature of participant exclusion in the original study and its misreporting, and the very weak statistical evidence for a benefit of reactivation-extinction over regular extinction in preventing fear recovery in the original report (see Chalkia et al., 2020), combined with our failure to find any evidence for the effect in the current direct replication, indicate that there is reason to doubt the efficacy and replicability of the original reactivation-extinction protocol.

Considering that there have been some successful demonstrations of the reactivation-extinction effect in conceptual replications, one may question whether it is simply a sample-specific effect, bearing in mind the considerable rate of participant exclusions that may be needed to obtain evidence for it. Not only in Schiller et al. (2010, Experiment 1), but also in several of the later demonstrations (Agren et al., 2017; Asthana et al., 2016; Johnson & Casey, 2015; Steinfurth et al., 2014), participants were highly selected, with some studies excluding up to 74% of the total sample recruited (Schiller et al., 2010, Experiment 2; Schiller et al., 2013). With such rates of participant exclusions, the resulting sample may not be representative of the general population but rather of a specific subset of individuals that have a particular aptitude at learning from sparse contingency information. Indeed, one reason why studies had such high exclusion rates is tied to the very low reinforcement rate employed during acquisition (37.5% in Schiller et al., 2010). Partial reinforcement is known to impair the acquisition of conditioned responses (Poulos & Gormezano, 1974). In the present study this was quite evident, as 67% of our exclusions were due to a failure to meet acquisition criteria. That is to say, not everyone learns readily under such a low reinforcement rate, implying that the participants that were included may be somewhat unique in their learning capabilities. Perhaps, one might be tempted to argue, reactivation-extinction is indeed beneficial for this selected sub-group. In line with this idea, it is important to point out that the individually-standardized cut-off (.1 divided by average US response) employed by Schiller et al. (2010) for inclusion may also be responsible for the selection of a very specific sub-group of participants. It can be noticed that individuals with larger average responses to the US during acquisition yield a smaller cut-off value, while those with smaller responses to the US generate a larger cut-off value. As such, participants with larger US responding are eventually included even if their differential CS+/CS- discrimination during acquisition is small. On the other hand, participants

with smaller responses to the US would need to exhibit considerable discrimination between CS+ and CS- in order to be above said cut-off during acquisition and thus, be included. Yet, if reactivation-extinction was indeed advantageous for certain sub-groups of participants only, we should have been able to observe this benefit in the present study, considering that we used the same partial reinforcement schedule and similar inclusion criteria as in Schiller et al. (2010). Of course, the problem with such extensive data-based selection of participants is that it creates room for researcher degrees of freedom and more generally risks to exert an influence on the results, particularly in view of the lack of clear standards for data-based exclusions in human fear conditioning research (Lonsdorf et al., 2019).

Fear reduction through reactivation-extinction has been suggested to rely on a different mechanism than regular extinction, specifically, on reconsolidation-induced memory updating (Schiller & Phelps, 2011). In line with this idea, several imaging studies have provided evidence for differences in neural network activation and brain connectivity during and following reactivation-extinction versus regular extinction (Agren et al., 2012; Björkstrand et al., 2017, 2016; Schiller et al., 2013). Again, however, all those studies involved a relatively low sample size and considerable selection of participants based on idiosyncratic exclusion criteria (yielding sample sizes of $N = 19$ in Schiller et al., 2013, within-group comparison; $N = 22$ in Agren et al., 2012, $N = 45$ in Björkstrand et al., 2016, and $N = 39$ in Björkstrand et al., 2017, all between-group comparisons), thus necessitating further replication in a larger sample (Button et al., 2013). At the same time, some studies have observed a reversed reactivation-extinction effect, in which extinction was more effective if it was followed rather than preceded by reactivation (Baker et al., 2013; Millan, Milligan-Saville, & McNally, 2013; but see; Telch, York, Lancaster, & Monfils, 2017). This poses a challenge for a reconsolidation view of memory updating, as the theory suggests that memory retrieval triggers the induction of reconsolidation, a transient state where the memory may be modified (Przybylski & Sara, 1997; Sara, 2000). When reactivation follows extinction, reconsolidation processes cannot truly be involved, so these findings question the logic and mechanism of reactivation-extinction in general.

Notwithstanding our failure to directly replicate the original findings of Schiller et al. (2010) and concerns about the reliability of the original report, the fact remains that some conceptual replication attempts have yielded positive results. The meta-analysis of Kredlow et al. (2016) tried to shed light on reactivation-extinction effect sizes, but at the time, it could include only eight human fear conditioning studies, and only three of those studies involved a spontaneous recovery test, for which the meta-analysis found a moderate, significant effect favoring reactivation-extinction over regular extinction. A closer look reveals that the only study of the three to show that reactivation-extinction prevents spontaneous recovery was the original Schiller et al. (2010, Experiment 1) study, while in the other two, Meir Drexler et al. (2014) and Soeter and Kindt (2011), spontaneous recovery was observed following reactivation-extinction. Of note, the study of Meir Drexler et al. (2014) was

reported to show a positive, moderate benefit of reactivation-extinction in the Kredlow et al. (2016) meta-analysis ($g = .59$), contributing to the overall spontaneous recovery effect size, even though in the original report the authors clearly state that they did not find a benefit of reactivation-extinction and spontaneous recovery was found in both groups (Meir Drexler et al., 2014). Further inquiry revealed that this divergence can be attributed to different indices used to evaluate group differences in Meir Drexler et al. (2014) versus the meta-analysis. So, although both groups showed an increase in differential responding from the end of extinction to the Day 3 test, and statistically indistinguishable differential responding was present in the two groups on that test, there were slight differences in the degree of differential responding still present at the end of extinction (Day 2), which inflated the observed amount of spontaneous recovery in the control group (Shira Meir Drexler, personal communication, May 22, 2019). Regardless, if the Schiller et al. (2010) results were to be removed from the meta-analysis and the present replication results along with other recent results (see footnotes 1 and 2) were added, it remains to be seen whether any overall evidence for a benefit of reactivation-extinction over regular extinction in human fear learning would remain.

Reactivation-extinction has been proposed to have great clinical translational potential, as it allows for the opportunity to (non-pharmacologically) augment current therapies for those suffering from anxiety-related disorders (Monfils & Holmes, 2018). Yet, some primary clinical investigations suggest otherwise, as they all failed to observe a benefit of reactivation-extinction with anxious (classified as scoring above 15 on the Beck Anxiety Inventory or above 37 on the Fear Questionnaire; Kredlow, Orr, & Otto, 2018) or phobic individuals (Maples-Keller et al., 2017; Shiban, Brütting, Pauli, & Mühlberger, 2015; but see; Telch et al., 2017). These failed replication and translational attempts in conjunction with the mixed basic findings suggest that the idea of clinical translation may be premature. It may be the case that unknown individual differences moderate reactivation-extinction effects. In an intriguing report, Shumake and colleagues re-analyzed older rodent datasets with the aim of detecting different phenotype subgroups based on their responding during acquisition, extinction, and reinstatement (Shumake, Jones, Auchter, & Monfils, 2018). Only two subgroups (33% of their total sample of $N = 215$) showed a long-term fear reduction, while the other five subgroups showed a return or persistence of fear. Future research should strive to examine such patterns in human data, as this could shed light on specific sample characteristics that differentially influence outcome success.

In conclusion, in a preregistered, highly powered, independent, direct replication of the original reactivation-extinction paradigm in humans, we failed to obtain any evidence that this manipulation prevents the return of fear. Our results complement those of Luyten and Beckers (2017), who also failed to observe a benefit of reactivation-extinction in a direct replication of Monfils et al. (2009) in animals. Considering the results of the present study, along with the mixture of findings in the literature, and in light of the inaccurate reporting in Schiller et al. (2010) that meanwhile emerged (Chalkia et al., 2020; Schiller et al., 2018), one must conclude that at present,

evidence that a reactivation-extinction procedure prevents the return of fear in humans is not very robust.

CRediT author statement

Anastasia Chalkia: Conceptualization, Software, Formal analysis, Investigation, Data curation, Writing – Original draft, Writing – Review & Editing, Visualization. **Natalie Schroyens:** Formal analysis, Investigation, Writing – Review & Editing. **Lu Leng:** Investigation, Data curation, Writing – Review & Editing. **Niels Vanhasbroeck:** Investigation, Data curation, Writing – Review & Editing. **Ann-Kathrin Zenses:** Writing – Review & Editing. **Lukas Van Oudenhove:** Conceptualization, Writing – Review & Editing, Supervision. **Tom Beckers:** Conceptualization, Writing – Original draft, Writing – Review & Editing, Supervision, Funding acquisition.

Funding

Preparation of this manuscript and collection of the data was supported by ERC Consolidator Grant 648176 awarded to Tom Beckers. The funding source was not involved in the study design, data collection and analysis, or in the preparation of this manuscript.

Open practices

The study in this article earned Open Materials, Open Data and Preregistered badges for transparent practices. Materials and data for the study are available at <https://osf.io/8chqu/>.

Declaration of Competing Interest

None.

Acknowledgements

The authors would like to thank Mathijs Franssen and Jeroen Clarysse for providing essential technical support. We also thank Anna Schnell, Wout Vossen, Ludwig Snoeks, Margaux Désirée Sageot and Axelle Bavr  for their help with data collection.

Supplementary information

Supplementary information to this article can be found online at <https://doi.org/10.1016/j.cortex.2020.04.017>.

REFERENCES

Agren, T., Bj rkstrand, J., & Fredrikson, M. (2017). Disruption of human fear reconsolidation using imaginal and in vivo

- extinction. *Behavioural Brain Research*, 319, 9–15. <https://doi.org/10.1016/j.bbr.2016.11.014>.
- Agren, T., Engman, J., Frick, A., Björkstrand, J., Larsson, E.-M., Furmark, T., et al. (2012). Disruption of reconsolidation erases a fear memory trace in the human amygdala. *Science*, 337(6101), 1550–1552. <https://doi.org/10.1126/science.1223006>.
- Asthana, M. K., Brunhuber, B., Mühlberger, A., Reif, A., Schneider, S., & Herrmann, M. J. (2016). Preventing the return of fear using reconsolidation update mechanisms depends on the met-allele of the brain derived neurotrophic factor Val66Met polymorphism. *International Journal of Neuropsychopharmacology*, 19(6), 1–9. <https://doi.org/10.1093/ijnp/pyv137>.
- Auber, A., Tedesco, V., Jones, C. E., Monfils, M. H., & Chiamulera, C. (2013). Post-retrieval extinction as reconsolidation interference: Methodological issues or boundary conditions? *Psychopharmacology*, 226(4), 631–647. <https://doi.org/10.1007/s00213-013-3004-1>.
- Baker, K. D., McNally, G. P., & Richardson, R. (2013). Memory retrieval before or after extinction reduces recovery of fear in adolescent rats. *Learning & Memory*, 20(9), 467–473. <https://doi.org/10.1101/lm.031989.113>.
- Bakkour, A., Schonberg, T., Hover, A. M., & Poldrack, R. A. (2015). Retrieval-extinction within the memory reconsolidation window does not influence appetitive choice. *BioRxiv*. <https://doi.org/10.1101/014316>.
- Beckers, T., & Kindt, M. (2017). Memory reconsolidation interference as an emerging treatment for emotional disorders: Strengths, limitations, challenges and opportunities. *Annual Review of Clinical Psychology*, 13. <https://doi.org/10.1146/annurev-clinpsy-032816-045209>.
- Björkstrand, J., Agren, T., Åhs, F., Frick, A., Larsson, E.-M., Hjorth, O., et al. (2016). Disrupting reconsolidation attenuates long-term fear memory in the human amygdala and facilitates approach behavior. *Current Biology*, 26, 1–6. <https://doi.org/10.1016/j.cub.2016.08.022>.
- Björkstrand, J., Agren, T., Åhs, F., Frick, A., Larsson, E.-M., Hjorth, O., et al. (2017). Think twice, it's all right: Long lasting effects of disrupted reconsolidation on brain and behavior in human long-term fear. *Behavioural Brain Research*, 324, 125–129. <https://doi.org/10.1016/j.bbr.2017.02.016>.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, 114(1), 80–99. <https://doi.org/10.1037/0033-2909.114.1.80>.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, 52(10), 976–986. [https://doi.org/10.1016/S0006-3223\(02\)01546-9](https://doi.org/10.1016/S0006-3223(02)01546-9).
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., et al. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365–376. <https://doi.org/10.1038/nrn3475>.
- Chalkia, A., Van Oudenhove, L., & Beckers, T. (2020). Preventing the return of fear in humans using reconsolidation update mechanisms: A verification report of Schiller et al. (2010). *Cortex*, 129, 510–525. <https://doi.org/10.1016/j.cortex.2020.03.031>.
- Chan, W. Y. M. (2014). *The effects of retrieval - extinction training on the restoration of Pavlovian conditioned fear*.
- Chan, W. Y. M., Leung, H. T., Westbrook, R. F., & McNally, G. P. (2010). Effects of recent exposure to a conditioned stimulus on extinction of Pavlovian fear conditioning. *Learning & Memory*, 17(10), 512–521. <https://doi.org/10.1101/lm.1912510>.
- Clem, R. L., & Huganir, R. L. (2010). Calcium-permeable AMPA receptor dynamics mediate fear memory erasure. *Science*, 330(6007), 1108–1112. <https://doi.org/10.1126/science.1195298.Calcium-Permeable>.
- Costanzi, M., Cannas, S., Sarauli, D., Rossi-Arnaud, C., & Cestari, V. (2011). Extinction after retrieval: Effects on the associative and nonassociative components of remote contextual fear memory. *Learning & Memory*, 18, 508–518. <https://doi.org/10.1101/lm.2175811>.
- Faul, F., Erdfelder, E., Lang, A.-G. G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>.
- Flavell, C. R., Barber, D. J., & Lee, J. L. C. (2011). Behavioural memory reconsolidation of food and fear memories. *Nature Communications*, 2, 504. <https://doi.org/10.1038/ncomms1515>.
- Fricchione, J., Greenberg, M. S., Spring, J., Wood, N., Mueller-Pfeiffer, C., Milad, M. R., et al. (2016). Delayed extinction fails to reduce skin conductance reactivity to fear-conditioned stimuli. *Psychophysiology*, 53, 1343–1351. <https://doi.org/10.1111/psyp.12687>.
- Gisquet-Verrier, P., & Riccio, D. C. (2012). Memory reactivation effects independent of reconsolidation. *Learning & Memory*, 19(9), 401–409. <https://doi.org/10.1101/lm.026054.112>.
- Golkar, A., Bellander, M., Olsson, A., & Ohman, A. (2012). Are fear memories erasable? Reconsolidation of learned fear with fear-relevant and fear-irrelevant stimuli. *Frontiers in Behavioral Neuroscience*, 6(80), 1–10. <https://doi.org/10.3389/fnbeh.2012.00080>.
- Hu, J., Wang, W., Homan, P., Wang, P., Zheng, X., & Schiller, D. (2018). Reminder duration determines threat memory modification in humans. *Scientific Reports*, 8(8848), 1–10. <https://doi.org/10.1038/s41598-018-27252-0>.
- Ishii, D., Matsuzawa, D., Matsuda, S., Tomizawa, H., Sutoh, C., & Shimizu, E. (2012). No erasure effect of retrieval – extinction trial on fear memory in the hippocampus-independent and dependent paradigms. *Neuroscience Letters*, 523, 76–81. <https://doi.org/10.1016/j.neulet.2012.06.048>.
- JASP Team. (2019). *JASP (version 0.9.2)*.
- Johnson, D. C., & Casey, B. J. (2015). Reconsolidation blocks recovery of fear in adolescents. *Scientific Reports*, 5(8863), 1–5. <https://doi.org/10.1038/srep08863>.
- Jones, C. E., Ringuelet, S., & Monfils, M. (2013). Learned together, extinguished apart: Reducing fear to complex stimuli. *Learning & Memory*, 20(12), 674–685.
- Kindt, M., & Soeter, M. (2013). Reconsolidation in a human fear conditioning study: A test of extinction as updating mechanism. *Biological Psychology*, 92(1), 43–50. <https://doi.org/10.1016/j.biopsycho.2011.09.016>.
- Klucken, T., Kruse, O., Schweckendiek, J., Kuepper, Y., Mueller, E. M., Hennig, J., et al. (2016). No evidence for blocking the return of fear by disrupting reconsolidation prior to extinction learning. *Cortex*, 79, 112–122. <https://doi.org/10.1016/j.cortex.2016.03.015>.
- Kredlow, M. A., Orr, S. P., & Otto, M. W. (2018). Exploring the boundaries of post-retrieval extinction in healthy and anxious individuals. *Behaviour Research and Therapy*, 108, 45–57. <https://doi.org/10.1016/j.brat.2018.06.010>.
- Kredlow, M. A., Unger, L. D., & Otto, M. W. (2016). Harnessing reconsolidation to weaken fear and appetitive memories: A meta-analysis of post-retrieval extinction effects. *Psychological Bulletin*, 142(3), 314–336. <https://doi.org/10.1037/bul0000034>.
- Kroes, M. C. W., Dunsmoor, J. E., Lin, Q., Evans, M., & Phelps, E. A. (2017). A reminder before extinction strengthens episodic memory via reconsolidation but fails to disrupt generalized threat responses. *Scientific Reports*, 7(10858), 1–14. <https://doi.org/10.1038/s41598-017-10682-7>.
- Liu, J., Zhao, L., Xue, Y., Shi, J., Suo, L., Luo, Y., et al. (2014). An unconditioned stimulus retrieval extinction procedure to prevent the return of fear memory. *Biological Psychiatry*, 76(11), 895–901. <https://doi.org/10.1016/j.biopsych.2014.03.027>.

- Lonsdorf, T. B., Klingenhöfer-Jens, M., Andreatta, M., Beckers, T., Chalkia, A., Gerlicher, A., et al. (2019). Navigating the garden of forking paths for data exclusions in fear conditioning research. *eLife*, 8(e52465). <https://doi.org/10.7554/eLife.52465>.
- Luyten, L., & Beckers, T. (2017). A preregistered, direct replication attempt of the retrieval-extinction effect in cued fear conditioning in rats. *Neurobiology of Learning and Memory*, 144, 208–215. <https://doi.org/10.1016/j.nlm.2017.07.014>.
- Maples-Keller, J. L., Price, M., Jovanovic, T., Norrholm, S. D., Odenat, L., Post, L., et al. (2017). Targeting memory reconsolidation to prevent the return of fear in patients with fear of flying. *Depression and Anxiety*, 34, 610–620. <https://doi.org/10.1002/da.22626>.
- McNeil, D. W., & Rainwater, A. J. (1998). Development of the Fear of Pain Questionnaire-III. *Journal of Behavioral Medicine*, 21(4), 389–410. <https://doi.org/10.1023/A:1018782831217>.
- Meir Drexler, S., Merz, C. J., Hamacher-Dang, T. C., Marquardt, V., Fritsch, N., Otto, T., et al. (2014). Effects of postretrieval-extinction learning on return of contextually controlled cued fear. *Behavioral Neuroscience*, 128(4), 474–481. <https://doi.org/10.1037/a0036688>.
- Milad, M. R., Orr, S. P., Pitman, R. K., & Rauch, S. L. (2005). Context modulation of memory for fear extinction in humans. *Psychophysiology*, 42(4), 456–464. <https://doi.org/10.1111/j.1469-8986.2005.00302.x>.
- Millan, E. Z., Milligan-Saville, J., & McNally, G. P. (2013). Memory retrieval, extinction, and reinstatement of alcohol seeking. *Neurobiology of Learning and Memory*, 101, 26–32. <https://doi.org/10.1016/j.nlm.2012.12.010>.
- Misanin, J. R., Miller, R. R., & Lewis, D. J. (1968). Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science*, 160, 554–555. <https://doi.org/10.1126/science.160.3827.554>.
- Monfils, M. H., Cowansage, K. K., Klann, E., & LeDoux, J. E. (2009). Extinction-reconsolidation boundaries: Key to persistent attenuation of fear memories. *Science*, 324(5929), 951–955. <https://doi.org/10.1126/science.1167975>.
- Monfils, M. H., & Holmes, E. A. (2018). Memory boundaries: Opening a window inspired by reconsolidation to treat anxiety, trauma-related, and addiction disorders. *The Lancet Psychiatry*, 5(12), 1032–1042. [https://doi.org/10.1016/S2215-0366\(18\)30270-0](https://doi.org/10.1016/S2215-0366(18)30270-0).
- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, 406(6797), 722–726. <https://doi.org/10.1038/35021052>.
- Olshavsky, M. E., Jones, C. E., Lee, H. J., & Monfils, M. H. (2013). Appetitive behavioral traits and stimulus intensity influence maintenance of conditioned fear. *Frontiers in Behavioral Neuroscience*, 7(179), 1–7. <https://doi.org/10.3389/fnbeh.2013.00179>.
- Oyarzún, J. P., Lopez-Barroso, D., Fuentemilla, L., Cucurell, D., Pedraza, C., Rodríguez-Fornells, A., et al. (2012). Updating fearful memories with extinction training during reconsolidation: A human study using auditory aversive stimuli. *Plos One*, 7(6). <https://doi.org/10.1371/journal.pone.0038849>.
- Pérez-Cuesta, L. M., & Maldonado, H. (2009). Memory reconsolidation and extinction in the crab: Mutual exclusion or coexistence? *Learning & Memory*, 16(11), 714–721. <https://doi.org/10.1101/lm.1544609>.
- Piñeyro, M. E., Ferrer Monti, R. I., Alfei, J. M., Bueno, A. M., & Urcelay, G. P. (2013). Memory destabilization is critical for the success of the reactivation-extinction procedure. *Learning & Memory*, 21(1), 46–54. <https://doi.org/10.1101/lm.032714.113>.
- Pitman & Orr. (1986). Test of the conditioning model of neurosis: Differential aversive conditioning of angry and neutral facial expressions in anxiety disorder patients. *Journal of Abnormal Psychology*, 95(3), 208–213. <https://doi.org/10.1037//0021-843x.95.3.208>.
- Poulos, C. X., & Gormezano, I. (1974). Effects of partial and continuous reinforcement on acquisition and extinction in classical appetitive conditioning. *Bulletin of the Psychonomic Society*, 4(3), 197–198. <https://doi.org/10.3758/BF03334244>.
- Przybylski, J., & Sara, S. J. (1997). Reconsolidation of memory after its reactivation. *Behavioural Brain Research*, 84, 241–246. [https://doi.org/10.1016/S0166-4328\(96\)00153-2](https://doi.org/10.1016/S0166-4328(96)00153-2).
- Rao-Ruiz, P., Rotaru, D. C., van der Loo, R. J., Mansvelter, H. D., Stiedl, O., Smit, A. B., et al. (2011). Retrieval-specific endocytosis of GluA2-AMPA receptors underlies adaptive reconsolidation of contextual fear. *Nature Neuroscience*, 14(10), 1302–1308. <https://doi.org/10.1038/nn.2907>.
- Riccio, D. C., Millin, P. M., & Bogart, A. R. (2006). Reconsolidation: A brief history, a retrieval view, and some recent issues. *Learning & Memory*, 13(5), 536–544. <https://doi.org/10.1101/lm.290706>.
- Sara, S. J. (2000). Retrieval and reconsolidation: Toward a neurobiology of remembering. *Learning & Memory*, 7(2), 73–84. <https://doi.org/10.1101/lm.7.2.73>.
- Schiller, D., Kanen, J. W., LeDoux, J. E., Monfils, M. H., & Phelps, E. A. (2013). Extinction during reconsolidation of threat memory diminishes prefrontal cortex involvement. *Proceedings of the National Academy of Sciences of the United States of America*, 110(50), 20040–20045. <https://doi.org/10.1073/pnas.1320322110>.
- 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>.
- Schiller, D., Monfils, M. H., Raio, C. M., Johnson, D. C., LeDoux, J. E., & Phelps, E. A. (2018). Addendum: Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, 562(7727), E21. <https://doi.org/10.1038/s41586-018-0405-7>.
- Schiller, D., & Phelps, E. A. (2011). Does reconsolidation occur in humans? *Frontiers in Behavioral Neuroscience*, 5(24), 1–12. <https://doi.org/10.3389/fnbeh.2011.00024>.
- Schiller, D., Raio, C. M., & Phelps, E. A. (2012). Extinction training during the reconsolidation window prevents recovery of fear. *Journal of Visualized Experiments: JoVE*, (66), e3893. <https://doi.org/10.3791/3893>.
- Shiban, Y., Brütting, J., Pauli, P., & Mühlberger, A. (2015). Fear reactivation prior to exposure therapy: Does it facilitate the effects of VR exposure in a randomized clinical sample? *Journal of Behavior Therapy and Experimental Psychiatry*, 46, 133–140. <https://doi.org/10.1016/j.jbtep.2014.09.009>.
- Shumake, J., Jones, C., Auchter, A., & Monfils, M. H. (2018). Data-driven criteria to assess fear remission and phenotypic variability of extinction in rats. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373. <https://doi.org/10.1098/rstb.2017.0035>.
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, 26(5), 559–569. <https://doi.org/10.1177/0956797614567341>.
- Soeter, M., & Kindt, M. (2011). Disrupting reconsolidation: Pharmacological and behavioral manipulations. *Learning & Memory*, 18(6), 357–366. <https://doi.org/10.1101/lm.2148511>.
- Spielberger, C. D., Gorsuch, R., Lushene, R., Vagg, P., & Jacobs, G. (1977). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press. <https://doi.org/10.1037/t06496-000>.
- Steinfurth, E. C. K., Kanen, J. W., Raio, C. M., Clem, R. L., Huginir, R. L., & Phelps, E. A. (2014). Young and old Pavlovian fear memories can be modified with extinction training during reconsolidation in humans. *Learning & Memory*, 21(7), 338–341. <https://doi.org/10.1101/lm.033589.113>.
- Telch, M. J., York, J., Lancaster, C. L., & Monfils, M. H. (2017). Use of a brief fear memory reactivation procedure for enhancing

- exposure therapy. *Clinical Psychological Science*, 5(2), 367–378. <https://doi.org/10.1177/2167702617690151>.
- Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology*, 9(1), 215–248. <https://doi.org/10.1146/annurev-clinpsy-050212-185542>.
- Warren, V. T., Anderson, K. M., Kwon, C., Bosshardt, L., Jovanovic, T., Bradley, B., et al. (2014). Human fear extinction and return of fear using reconsolidation update mechanisms: The contribution of on-line expectancy ratings. *Neurobiology of Learning and Memory*, 113, 165–173. <https://doi.org/10.1016/j.nlm.2013.10.014>.
- Xue, Y., Luo, Y., Wu, P., Shi, H., Xue, L., Chen, C., et al. (2012). A memory retrieval-extinction procedure to prevent drug craving and relapse. *Science*, 336(6078), 241–245. <https://doi.org/10.1126/science.1215070>.
- Zuccolo, P. F., & Hunziker, M. H. L. (2019). A review of boundary conditions and variables involved in the prevention of return of fear after post-retrieval extinction. *Behavioural Processes*, 162, 39–54. <https://doi.org/10.1016/j.beproc.2019.01.011>.