

Novel approaches for strengthening human fear extinction: The roles of novelty, additional
USs, and additional GSs

Ottmar V. Lipp¹

Allison M. Waters²

Camilla C. Luck¹

Kathrin M. Ryan²

Michelle G. Craske³

¹School of Psychology, Curtin University, Australia

²School of Applied Psychology, Griffith University, Australia

³Department of Psychology, University of California, USA

For the special issue of Behaviour Research and Therapy entitled “100 years of
human fear conditioning”

Guest Editors: Bram Vervliet and Yannick Boddez

Corresponding author: Professor Ottmar Lipp, School of Psychology, Curtin University. GPO
Box U1987, Perth, WA, 6845, Australia; phone: +61 8 9266 5112; fax: +61 8 9266 2464;

Email: ottmar.lipp@curtin.edu.au

Acknowledgments:

This work was supported by Grant APP1156490 from the National Health and Medical
Research Council.

Abstract

Since Watson and Rayner's (1920) initial demonstration that human fear can be learned by means of Pavlovian conditioning, neuroscientific and behavioral studies have provided a thorough understanding of fear acquisition. Less is known about the manner in which we can harness insights from Pavlovian conditioning research to reduce fears and, most importantly, make the reduction of fear lasting and resistant against relapse. The current paper reviews three manipulations that have shown promise in achieving a reduction of conditional fear that is more resistant to relapse than is the reduction of conditional fear after standard extinction: novelty-facilitated extinction training, presentation of conditional-unconditional stimulus pairings or of unpaired unconditional stimuli during extinction, and extinction with additional stimuli that are similar to the original conditional stimuli. It summarizes past research involving human and non-human animal subjects and highlights knowledge gaps in the current literature. Moreover, it discusses potential mechanisms that mediate the reduction of fear seen as a result of these manipulations in an attempt to enhance our understanding of what renders fear extinction less vulnerable to the known pathways to fear relapse. It is hoped that this review will contribute to the achievement of the goal that was denied to Watson and Rayner, the development of experimental techniques that can be utilized to remove conditioned emotional responses permanently.

Key words: Fear conditioning, extinction, novelty-facilitated extinction, gradual extinction, generalization of extinction.

Introduction

Watson and Rayner's (1920) pioneering work almost a century ago has laid the foundation for several fields of psychological science that focus on human emotional learning, including developmental psychology, experimental psychology and experimental psychopathology. The relevance of this work has stood the test of time, surviving paradigm shifts and the development of highly sophisticated methodologies – arguably making their findings more relevant today than they were 50 years ago. On the way, Watson and Rayner's work has yielded outcomes, which in all likelihood would have made Pavlov, who gave his name to the experimental method, which Watson and Rayner employed, very proud indeed. Their work led to the development of theories that predicted undiscovered phenomena (e.g., over prediction, Rescorla & Wagner, 1972) and seeded the contemporary field of behavioral neuroscience, which enabled the realization of Pavlov's dream of using conditioning to understand how the brain works. Perhaps the greatest impact of this work was through the development of evidence-based exposure therapy interventions that have proven to be capable of reducing fear and improving the quality of life of millions of people suffering from anxiety disorders – the most common and costly mental health disorders worldwide.

Watson and Rayner's main aim was to provide experimental evidence for the hypothesis that conditional emotional responses can be established in the laboratory under conditions that mimic a child's early home life. In addition to establishing new emotional responses to stimuli that did not elicit these responses previously, Watson and Rayner's original research plan also included the exploration of methods to reduce these newly acquired emotional responses. However, this aspect of their project was not implemented in their seminal test case with "little Albert", due to his removal from the hospital environment in which his mother worked and he was being assessed. As stated in their section on "Detachment" or removal of conditional emotional response', 'the opportunity of building

up an experimental technique by means of which we could remove the conditioned emotional responses was denied us' (Watson & Rayner, 1920, p. 12). The failure to reduce Albert's conditional emotional response is one of the many aspects on which Watson and Rayner's original work was subsequently criticized. Nevertheless, it should be noted that Watson and Rayner proposed a number of interventions aimed at reducing anxiety that they intended to explore, and that, with the benefit of hindsight, seem to pre-empt later developments that led to now well-established clinical interventions. Their intention to investigate the reduction of fears came to fruition in the work of Mary Cover Jones (1924a, 1924b) who acknowledges Watson's input in her work. In her first study, Jones (1924a) describes the application of several different interventions to reduce fears in a sample of children living in an institution for the temporary care of children. In her summary, she identifies 'direct conditioning' – associating the fear-object with an that the children desire – and 'social imitation' – sharing in the social activity of a specially chosen group of children under controlled conditions – as the two interventions that provided 'unqualified success' Jones (1924a, p. 390). Jones (1924b) details the application of these interventions in a case study of the reduction of fear displayed by another little boy, Peter, who she describes as 'almost to be Albert grown a bit older' (Jones, 1924b, p. 309). Jones documented changes in Peter's fear as a result of her interventions, which included using direct conditioning and extinction methodologies in a controlled laboratory setting. These methodologies included components resembling fear hierarchies (i.e., gradual exposure to fear-provoking stimuli until emotional responses decline) and indirect learning strategies, such as the modelling of non-fear responses to fear-provoking stimuli by other children.

It is very much in this tradition of controlled experimental methodology and scientist-practitioner inquiry (cf. Waters, LeBeau, & Craske 2017) that the approaches currently under investigation in our laboratories are rooted. Reducing fear and anxiety are still major

challenges in clinical psychological practice and while our understanding of the manner in which fear is acquired has improved, thanks in part to Watson and Rayner and those who followed, less is known about how to reduce fear and more importantly, how to make this reduction long lasting and robust against relapse. Our work is based on the now well accepted view that extinction training (the major theoretical model of exposure therapy) results in the acquisition of a new inhibitory association between the conditional stimulus (CS) and the unconditional stimulus (US) that reduces conditional fear responding while the CS-US association, which was acquired during acquisition, is still maintained (Bouton, 2002). This renders the newly acquired reduction of conditional fear fragile, as future behavior will depend on which of the two competing associations is dominant. This view of extinction learning also suggests two potential avenues to long-lasting fear reduction – interventions that weaken or eliminate the original CS-US fear association or interventions that strengthen the newly acquired CS-noUS association. The former approach is the focus of a separate body of research on fear memory reconsolidation (see Schiller et al., 2010; Thompson & Lipp, 2017; see Elsey, Van Ast, & Kindt, 2018, for review). In brief, this approach exploits the observation that memories can become vulnerable and subject to change for a short period of time after retrieval. Using solely behavioral interventions, or a combination of behavioral and pharmacological interventions, recent basic research has yielded promising results (e.g., Schiller et al., 2010; Kindt, Soeter, & Vervliet, 2009). However, as some findings have been difficult to replicate, and translation to clinical psychological practice has proven challenging, more basic and applied research is required (Treanor, Brown, Rissman, & Craske, 2017).

The current overview focusses on three interventions that strengthen the newly acquired inhibitory CS-noUS association. It should be noted that our selection of interventions for discussion is not meant to be exhaustive or to suggest that other approaches would be less likely to prove successful (readers interested in approaches focusing on

interventions that reduce the mental representation of the US like imagery re-scripting, extinction in multiple contexts or counter conditioning are referred to (McEvoy et al., 2017; Neumann, Lipp, & Cory, 2007; or Keller, Hennings, & Dunsmoor, this volume respectively). Rather it reflects on current work conducted in our laboratories and is based on our reading of the literature and successful pilot work. The interventions to be discussed here are derived from an analysis of the extant experimental literature on extinction and extrapolations from clinical practice. Clinical practice has produced a number of interventions that are successful in reducing fear and anxiety and a systematic analysis of these interventions can provide us with ideas that can complement those derived from experimental studies (see Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). Conversely, animal studies of extinction learning have provided a number of promising candidates that may be suitable for translation into clinical practice with humans. Our discussion will focus on three manipulations that have yielded promising results in studies with human and nonhuman animal participants. We note, however, like research on fear memory reconsolidation, that these manipulations require a considerable amount of further research before they can be translated into clinical practice. However, the promising results observed in our proof of concept studies (Lucas, Luck, & Lipp, 2018; Thompson, McEvoy, & Lipp, 2018; Waters, Kershaw, & Lipp, 2018) and the potential that these manipulations offer for augmenting exposure-based therapy suggest that we are well on the way to achieve what Watson and Rayner could not – developing the experimental evidence-base for the long-lasting reduction of human conditional fear responses.

Novelty-facilitated extinction

Rationale for approach

It is well established that learning involving novel CSs is faster than learning that utilizes CSs that are familiar (latent inhibition; Lubow, 1989; Lipp, Sheridan & Siddle, 1994) or CSs that are good predictors of their outcomes (Hall & Pearce, 1980). More recent work on novelty-facilitated extinction expands on the role of stimulus novelty in fear learning, whereby the presentation of novel outcomes that replace the US can strengthen extinction learning and prevent the return of fear (Dunsmoor, Campese, Ceceli, LeDoux, & Phelps, 2015; Dunsmoor, Kroes, Li, Daw, Simpson, & Phelps, 2019; Lucas et al., 2018).

Overview of findings to date

Novelty-facilitated extinction was first described by Dunsmoor et al. (2015). In two experiments, the first involving rodents and the second humans, the authors showed that replacing the aversive electrocutaneous US with a novel non-aversive tone during extinction training reduced spontaneous recovery of the conditional response – freezing in rats and electrodermal responses in humans – when tested 24 hours after extinction training. In humans, when reinstatement was examined after spontaneous recovery, Dunsmoor et al. (2015) also reported a trend towards a reduction of the conditional response in the novelty-facilitated extinction condition relative to the standard extinction condition. In a human sample, Lucas et al. (2018) assessed reinstatement immediately after extinction training and confirmed this finding of reduced reinstatement after novelty-facilitated extinction training relative to standard extinction. The procedure used in the human studies involves training in a differential conditioning procedure in which two angry face CSs were presented, one followed by the electrocutaneous US on a portion of the trials (33% in Dunsmoor et al., 2015, 2019, and 50% in Lucas et al., 2018) and the second presented alone. Extinction comprised 16 presentations of both CSs in the absence of the US in the standard extinction control, whereas in the novelty-facilitated extinction group, the CS+ co-terminated with a novel event – an innocuous tone. Dunsmoor et al. (2015) assessed spontaneous recovery and

reinstatement (following the presentation of three unpaired electrocutaneous stimuli) in a test session 24 hours after extinction, whereas Lucas et al. (2018) assessed reinstatement immediately after extinction training. Both studies provided evidence that extinction retention was enhanced after novelty facilitated extinction training, a finding replicated by Dunsmoor et al. (2019) who used the same basic design as Dunsmoor et al. (2015) and also collected brain-imaging data. Moreover, both Dunsmoor et al. (2015) and Lucas et al. (2018) found that the extent of extinction retention seen after standard extinction (without novel stimuli) was correlated with the extent of self-reported Intolerance of Uncertainty (Carleton, Norton, & Asmundson, 2007), whereas no such relation was found after novelty facilitated extinction. This finding may provide some information as to the mechanism that mediates the effect of novelty-facilitated extinction training.

Kryptos and Engelhard (2018) failed to find evidence for novelty-facilitated extinction in a conditioned avoidance paradigm. After a brief differential conditioning training (2 CS+-US and 2 CS- alone) participants were presented with an avoidance conditioning phase during which the occurrence of the US could be prevented by pressing the space bar. This was followed by extinction training (12 trials per CS), which was novelty-facilitated for one group and standard for another, the presentation of three USs to induce reinstatement, and a reinstatement test. US expectancy ratings, self-reported fear of the CSs, and avoidance behaviour provided evidence for reinstatement in both groups and, contrary to expectations, self-reported fear of the CSs was *enhanced* in the novelty-facilitated extinction group. It should be noted, however, that Kryptos and Engelhard's (2018) procedure differed from Dunsmoor et al's (2015) in a number of ways, including the use of different CSs (spiders vs. angry faces), different dependent measures (electrodermal activity was not measured), a very brief fear conditioning phase, and the collection of self-report measures in-

between the experimental phases. . Nevertheless, the study suggests that novelty-facilitated extinction may be subject to boundary conditions that are as yet to be determined.

Potential underlying mechanisms

Currently, putative mediating mechanisms of novelty-facilitated extinction training are not well understood. Dunsmoor et al. (2015) discussed an explanation in terms of counter conditioning, but conceded that this may be unlikely as traditional theories of counter conditioning assume a competition between opposing motivational tendencies, appetitive and aversive, for instance (for a detailed review of counter conditioning see Keller et al., this volume). In a study with rodents, Escobar, Arcediano, and Miller (2001) demonstrated that interference between outcomes does not require the outcomes to be emotionally salient. Here, training CSa with outcome C interfered with a previously acquired association between CSa and outcome B (in comparison to a control that was trained with CSa and C explicitly unpaired). This retroactive interference between outcomes resembles novelty-facilitated extinction, however, CSs and outcomes were neutral stimuli employed in a sensory preconditioning procedure and it is not clear whether retroactive interference between outcomes would be observed if the outcomes differed in salience, as is the case in novelty-facilitated extinction training.

Although novelty per se has been shown to be rewarding (Berlyne, 1960), it seems difficult to conceptualize how the novel, but rather neutral tone that replaces the US during extinction training, can be evaluated as sufficiently motivating to counteract conditioning with an aversive electrotactile stimulus. Alternatively, it may be that the novel stimulus, which is presented after the CS that was paired with the US previously, results in an enhanced prediction error, which in turn strengthens the acquisition of the inhibitory CS-noUS association. This interpretation is consistent with recent neuroimaging data suggesting different patterns of neutral activation between novelty-facilitated and standard extinction

(Dunsmoor et al., 2019). During standard extinction, differential activation during CS+ and CS- in areas associated with threat inhibition (ventromedial prefrontal cortex; dorsolateral prefrontal cortex) and threat appraisal and expression (insula and dorsal anterior cingulate cortex) was larger than during novelty-facilitated extinction. While promising, the assumption that the presentation of an unexpected novel stimulus during extinction will enhance prediction error across extinction training is in need of further exploration. The ‘novel’ stimulus used by Dunsmoor et al. (2015, 2019) and Lucas et al. (2018) was a pure tone of 440 or 800 Hz presented at a low volume of 60 or 80 dB, respectively. Although unexpected and surprising when encountered initially, orienting to this novel stimulus, a reliable indicator of stimulus novelty and processing, should have habituated within a rather small number of trials (Siddle & Lipp, 1997). One means of counteracting this reduction in orienting would be to vary the nature of the tone that replaces the aversive electrocutaneous stimulus. Such a manipulation should result in an even stronger reduction of relapse than the repeated presentation of the same tone.

Lucas et al. (2018) suggested a third explanation by pointing out that the introduction of the novel tone stimulus was not the only change occurring at the transition from acquisition to extinction. Rather, the reinforcement schedule employed changed from partial (33 or 50% respectively) during acquisition to continuous during extinction as the novel tone was presented after every presentation of the CS+ during extinction training. Thus, the CS+ was followed by an uncertain outcome during acquisition, but a certain outcome during novelty facilitated extinction training. The suggestion that the reduction in fear relapse observed after novelty facilitated extinction training was mediated by a reduction in outcome uncertainty is supported by the finding that the extent of fear recovery varied as a function of self-reported Intolerance of Uncertainty in the standard extinction conditions of both past studies (Dunsmoor et al., 2015; Lucas et al., 2018). Thus, novelty-facilitated extinction may

be mediated by a reduction in outcome uncertainty due to the replacement of the US with a novel stimulus during extinction, but further studies are required to test this possibility relative to competing hypotheses. These studies would need to hold the reinforcement schedules used during acquisition and novelty-facilitated extinction constant to assess the relative contribution of outcome novelty and change in outcome (un-)certainty.

Applied clinical work may provide a fourth potential explanation for the effectiveness of novelty-facilitated extinction training. Moritz, Jelinek, Klinge, and Naber (2007) developed a technique called association splitting to reduce obsessive thoughts, based on the fan effect (Anderson, 1974). The fan effect as described in the memory literature refers to the observation that associating a particular stimulus with multiple concepts may weaken individual associations, rendering their retrieval slower and more error prone. Moritz et al. (2007) reasoned that this effect could be utilized in the treatment of obsessive thoughts, as obsessive thoughts are believed to be based on associative networks that converge on one particular obsessive cognition. Adding additional, meaningful associations to the network will reduce the strength of the associations with the obsessive cognition and divert attention away from them. This approach has been shown to be successful in the treatment of OCD when used in addition to standard CBT (Jelinek, Hauschildt, Hottenrott, Kellner, & Moritz, 2018). Taken with the results seen for novelty-facilitated extinction training, these findings suggest that this approach may be usefully extended to other anxiety disorders like phobias or generalized anxiety disorder by broadening the associative networks that the feared stimuli or situations are embedded in.

To date, novelty facilitated extinction training has been shown to prevent the return of fear indicated by electrodermal responses due to spontaneous recovery (Dunsmoor et al., 2015, 2019) and reinstatement (Lucas et al., 2018). Whether novelty facilitated extinction training will also be robust against other relapse inductions (renewal or reacquisition) or will

reduce relapse assessed, not only by electrodermal responding, but also a wider range of conditioning indices is yet to be assessed. The latter issue is particularly pertinent given the results reported by Krypotos and Engelhard (2018) and the failure of Lucas et al. (2018) to find effects of novelty-facilitated extinction on online evaluations of CS valence. Moreover, the generality of the phenomenon remains to be tested as the three successful studies in humans to date all employed angry faces as conditional stimuli and pure tones as novel stimuli. Replicating the very promising initial findings using CSs that are not regarded as phylogenetically fear relevant (Öhman, 1986; Mallan, Lipp & Corchrane, 2013) and presenting ‘novel’ stimuli during extinction, that are more varied or drawn from a wider range (such as natural sounds for instance) to retain their novelty, will enable the identification of boundary conditions of the phenomenon and enhance the potential for meaningful translation into clinical practice.

Additional USs during extinction

Rationale for approach

The idea that presenting unpaired USs or occasional CS-US pairings during extinction training would lead to stronger extinction and a reduction in relapse appears counterintuitive. On the other hand, however, this has been shown by several studies in the animal learning literature (for unpaired US presentations see: Frey & Butler, 1977; Rauhut, Thomas, & Ayres, 2001; for occasionally paired CS-US presentations see Bouton, Woods, & Pineño, 2004; Gershman, Jones, Norman, Monfils, & Niv, 2013) even if, as expected, the reduction in conditional responding during extinction training proceeded slower in the presence of paired or unpaired USs. Different mechanisms have been discussed for this observation but, to date, no consensus has been achieved.

Overview of findings to date

Bouton et al. (2004) assessed the effect of additional CS-US pairings and of unpaired presentations of the US during extinction training in an appetitive conditioning paradigm with rodents, using the speed of reacquisition as an index of the strength of extinction. In their second experiment, three groups of rats received acquisition training with a 100% reinforcement schedule followed by either standard extinction, extinction with occasional CS-US pairings or extinction with unpaired presentations of the US. The number of CS-US trials and US alone presentations during extinction was reduced gradually from 1:8 to 1:12 and 1:24 over successive days of extinction training. An additional control group received US only presentations during acquisition and exposure to the context only during 'extinction'. Occasional reinforcement slowed extinction of conditional responding, however, relative to standard extinction, reacquisition was also slowed. Most remarkably, reacquisition was even slower in the group that had received unpaired presentations of the US during extinction. In this group, reacquisition did not differ from that seen in the control group that had never been exposed to paired presentations of CS and US (see Woods & Bouton, 2007; Mickley et al., 2009; Thomas, Longo, & Ayres, 2005, for similar results).

Vervliet, Vansteenwegen, and Hermans (2010) were the first to assess the effect of unpaired USs during extinction training in human differential fear conditioning. Two groups of 16 participants were trained in an ABA renewal design with habituation, acquisition and test in context A and extinction in context B. The groups differed in that 6 unpaired USs were presented during the intertrial intervals of extinction in group Unpaired. US expectancy was higher during CS+ than CS- on the first test trial in the standard extinction group, but not in group Unpaired, providing some evidence for a difference between the groups. However, this result needs to be interpreted with caution, as the Group \times CS interaction was not significant. Electrodermal responses yielded the desired Group \times CS interaction, with electrodermal renewal evident after standard extinction, but not after extinction training with unpaired USs.

Shiban, Wittmann, Weißinger, and Mühlberger (2015) compared the effects of gradual and standard extinction training on human fear conditioning in a virtual reality setting. In gradual extinction training the abrupt transition between acquisition and extinction, from a high percentage of CS-US pairings to none, is replaced by a gradual increase in the percentage of CS alone presentations and a decrease in the percentage of CS-US pairings (see Bouton et al., 2004; Gershman et al., 2013). All participants were presented with 18 CS+ and 18 CS- trials with an aversive air puff US presented 2 s after the onset of the 8 s CS+ on 80% of the trials. Extinction training commenced after a 10 minute break and consisted of 22 presentations of the CS+ and CS- alone during standard extinction or 17 CS+ alone, 5 CS+-US and 22 CS- trials during gradual extinction training (CS+-US pairings were presented as the 1st, 3rd, 6th, 10th and 15th extinction trials). Reinstatement was assessed 24 hours later following two unpaired presentations of the air puff US. During the last part of extinction and during reinstatement, blink startle during CS+ was larger than during CS- in the standard extinction group, but not in the gradual extinction group, a result taken to indicate enhanced reinstatement after standard extinction. No differences were observed in electrodermal responses or US expectancy ratings. While interesting, these results need to be interpreted with caution. CS+ and CS- were not counterbalanced (CS+ was always a spider, CS- a scorpion) which led to differential US expectancies prior to acquisition. Startle was elicited 6 s after CS onset and may have been affected by the air puff US that was presented 2 s after CS onset. Finally, the absence of a three way interaction involving the factor trial in the analysis of startle reinstatement leaves it open whether the between group difference in startle modulation was due to between group differences during extinction or in response to the reinstatement manipulation.

Culver et al. (2018) assessed the effects of occasional CS-US pairings during extinction on the spontaneous recovery, re-acquisition and re-extinction of electrodermal

responses, US expectancy and CS valence ratings. Two groups of participants were exposed to habituation, acquisition, and extinction phases on day 1, and tests for spontaneous recovery, reacquisition, and re-extinction on day 2. One group received standard extinction training whereas the second received 6 CS+-US pairings evenly spaced among the extinction trials. The pattern of results was complex, with some evidence for slower reacquisition of electrodermal responses after partial reinforcement during extinction. After occasionally reinforced extinction training, spontaneous recovery of US expectancy was reduced, but there was no clear evidence that CS valence ratings were differentially affected by occasionally reinforced extinction training.

Two further studies provide a direct comparison of the effects of unpaired USs during extinction and occasionally reinforced extinction training (Van den Akker et al., 2015; Thompson et al. 2018). Van den Akker et al. employed an appetitive learning paradigm in which the CSs, two distinct boxes containing either a small cup with chocolate mousse (US) or no US, were placed in front of the participants while they rated their expectancy of and desire for chocolate mousse. Participants then opened the boxes and consumed the content – if there was one. The experiment comprised acquisition, extinction, and re-acquisition phases. After the last re-acquisition trial, saliva flow in response to CS+ and CS- was also measured. The three groups of 30 participants differed in their treatment during extinction. During standard extinction, no USs were presented. During occasionally reinforced extinction, two reinforced CS+ were presented as the second and sixth of the 40 extinction trials in accordance with the notion of gradual extinction training. Finally, in the unpaired group, USs (small cups of chocolate mousse) were presented during the intertrial intervals following one CS+ and one CS- presented on trials 2 and 6 of extinction respectively. Reacquisition of chocolate mousse expectancy was slower after unpaired and occasionally reinforced

extinction in comparison to standard extinction, but there was no differential effect on the desire for chocolate mousse, saliva flow, or valence ratings.

Thompson et al. (2018) employed fear conditioning to assess the effects of unpaired US presentations and of occasional CS-US pairings during extinction training in comparison to a standard extinction control. Three groups of 24 participants were presented with habituation and acquisition training followed by extinction training after a break of 10 minutes. After a further 10-minute break, spontaneous recovery was assessed and 3 USs were presented to induce reinstatement. Reinstatement was tested after a 2-minute break and followed by a reacquisition phase with a 50% reinforcement schedule. One group received standard extinction training, a second received 5 CS+-US pairings during extinction (2 each during the first 2 blocks of 16 extinction trials and one early during the last block), and a third group received 5 additional USs during intertrial intervals at positions equivalent to the paired trials in the occasionally reinforced group. Counterbalanced across participants, equal numbers of the unpaired USs were presented after CS+ and after CS- trials. Electrodermal responses were measured continuously and CS valence ratings were taken after each phase of the experiment. Spontaneous recovery of electrodermal responses was evident after standard extinction, but not after occasionally reinforced or unpaired extinction training. Extinction training did not differentially affect reinstatement, but reacquisition was absent after extinction with unpaired US presentations whereas it was significant after standard and occasionally reinforced extinction, which did not differ from each other. Valence ratings of the CSs did not differ across the groups.

The available literature seems to provide an inconsistent pattern of results, which may be due to the use of different assessments of the strength of extinction (i.e., renewal, spontaneous recovery, reinstatement or reacquisition), and the use of different measures (e.g., skin conductance, US expectancy ratings, and ratings of CS valence) across studies.

However, it seems that ratings of CS valence taken after the experimental phases are not affected by the presentation of USs during extinction, be they unpaired or as part of occasionally reinforced extinction. US expectancy yielded somewhat inconsistent results with no clear effect when measured during tests of spontaneous recovery or during renewal, but a significant result when assessed during reacquisition. Fear learning as assessed by skin conductance seems to yield the most consistent evidence for strengthened extinction as a result of additional US presentations, more so, if assessed during spontaneous recovery, or, in the case of unpaired US presentations, reacquisition. Clearly, more empirical work is needed that systematically varies the method of assessing the strength of extinction and the dependent measures used. This work should also assess the manner in which the additional USs are distributed across extinction given the work on gradual extinction (Gershman et al., 2013). Overall, the pattern of results looks very promising, given the range of measures and assessments used and that the results originate from different laboratories.

Potential underlying mechanisms

Studies presenting unpaired USs during extinction training frequently cite Rescorla and Skucy's (1969) analysis of extinction learning as their foundation. In brief, Rescorla and Skucy explored whether withdrawal of food reward during extinction training had non-associative motivational effects that may affect conditional and other responding above the effects due to the removal of the response-outcome contingency. Rescorla and Skucy confirmed this proposal finding slower extinction when random presentations of food reward were made during extinction training. Frey and Butler (1977) extended the analysis of presenting the US during extinction training to a Pavlovian paradigm – eye lid conditioning in the rabbit – and found that explicitly unpaired presentations of the US during extinction training slowed extinction, but also retarded subsequent reacquisition. The authors attributed this finding to non-associative effects, but discussed the possibility of an inhibitory learning

mechanism. Rauhut et al. (2001) replicated the finding in a conditional emotional response paradigm, showing both reduced renewal and retarded reacquisition. They proposed habituation to the US as a potential mechanism for the effect of additional US presentations. This explanation was rejected by Thomas, Longo, and Ayres (2005) who found a reduction of renewal only for the CS that had been presented during unpaired extinction training, but not for other CSs. The authors also found no effect of US alone presentations after acquisition or of blocked presentations of CSs and USs – manipulations that should have led to US habituation. They also rejected the possibility that the CS had become a Pavlovian conditioned inhibitor after extinction with explicitly unpaired USs, as the CS did not inhibit responses to a second CS paired with the US. Thomas et al. (2005) proposed a hybrid theory combining Bouton's (2002) notion that an inhibitory association is formed between CS and US during extinction which competes with the CS-US association acquired during acquisition, with Rescorla-Wagner's (1972) proposal that the CS-US association formed during acquisition is weakened. This weakening of the CS-US association is stronger if responding is evoked in the presence of the CS during extinction (for instance by unpaired USs) and affected by the presence of other stimuli presented with the CS (see protection from extinction; Lovibond, Davis, & O'Flaherty, 2000). After a context change, the newly acquired inhibitory association suffers a larger generalization decrement than does the residual CS-US association, which depending on their relative strengths will either lead to renewal or not. Thus, unpaired USs during extinction training exert their influence not by strengthening the inhibitory association formed between CS and US, but by weakening the CS-US association formed during acquisition.

Bouton et al.'s (2004) study of the effects of occasional CS-US pairings during extinction training was motivated by Ricker and Bouton's (1996) research on reacquisition after successful extinction. To explain their findings, Ricker and Bouton (1996) had proposed

a ‘trial signalling’ mechanism that was derived from sequential theories such as Capaldi’s (1994). During acquisition, participants not only learn an association between CS and US, but they also acquire information about the trial sequence and learn that reinforced trials follow other reinforced trials. Conversely, during standard extinction, participants learn a CS-noUS association and that non-reinforced trials follow non-reinforced trials, which leads to faster reacquisition if reinforcement is presented again. Occasionally reinforced extinction training not only enables the formation of a CS-noUS association, but also supports learning that non-reinforced trials can follow reinforced trials, thus slowing down reacquisition and providing potential protection against relapse.

Dunsmoor, Kroes, Moscatelli, Evans, Davachi, and Phelps (2018) suggested a mechanism derived from research on episodic memory – event segmentation – that may account for the effects of additional US presentations during extinction (for a similar account see Matute, Lipp, Vadillo, & Humphreys, 2011). Event segmentation has been proposed as a mechanism that protects memory items against interference from other items encountered in close temporal proximity by chunking items within one episode together and including boundaries that separate successive episodes (Ezzyat, & Davachi, 2011). Segmentation may occur explicitly as part of the experimental design when different experimental phases are conducted across different days or separated by assessments of ratings or implicitly by the omission or addition of significant stimuli like the US. Interventions that reduce event segmentation such as for instance gradual extinction and that render acquisition and extinction phases more similar, may reduce the likelihood that learning from acquisition and extinction is encoded in separate memory episodes and hence increase the likelihood that new learning interferes with prior learning.

In addition to the accounts put forward above, others seem feasible. Presenting unpaired USs during extinction training may exert a non-associative influence by increasing

the participant's arousal, which may facilitate learning of the inhibitory CS-noUS relationship. It also may protect against relapse by serving as a form of inoculation. As reinstatement is typically induced by presenting unpaired USs after successful extinction training (usually two or three in quick succession), one might argue that presenting several USs spaced during intertrial intervals repeatedly exposes the participant to a reinstatement manipulation followed by CSs without the US. Thus, the repeated exposure to a reinstatement like manipulation may result in enhanced resistance to relapse. A similar 'learning to learn' explanation may account for the finding that occasional CS-US pairings during extinction training reduce relapse. One might argue that presenting repeated sequences of CS-US and CS alone trials is akin to repeated (weak) acquisition and extinction training, which may strengthen extinction by leading to stronger inhibitory learning.

More empirical research is needed to elucidate the mechanism(s) that lead to a strengthening of extinction after training with unpaired US presentations or occasional CS-US pairings. There is, for instance, no clear demonstration as to whether a graded reduction of the frequency of unpaired US presentations/additional CS-US pairings during extinction training as used by Bouton et al. (2004) or van den Akker et al. (2015) is superior to extinction training in which these trials are more evenly spaced (Vervliet et al., 2010; Culver et al., 2018).

Extinction with additional stimuli similar to the CSs

Rationale for approach

A third approach to strengthen the extinction of human fear conditioning is to include additional stimuli during extinction training that are perceptually and/or conceptually similar to the CS+, but were not presented during acquisition. This approach aims to strengthen the generalization of extinction learning. Although the underlying mechanisms are unclear,

current models of Pavlovian learning can offer an explanation¹. The Rescorla-Wagner model (Rescorla & Wagner, 1972), for instance, predicts generalization of associative learning from a CS+ to generalized conditional stimuli (GCS) as a function of the degree of shared elements between the CS and the GCS. A GCS will include elements that overlap with the CS+ as well as unique elements that are distinct from the CS+. Therefore, the Rescorla-Wagner model predicts that learning to the CS+ will only partially transfer to the GCS based only on overlapping elements and thus, result in a decrement in responding to the GCS compared to the CS+. Thus, at extinction test, fear responding to the CS+ is predicted to be greater if extinction occurred to a GCS compared to the CS+ due to the unextinguished unique elements of the CS+. Furthermore, a novel GCS at extinction test, which has some overlapping and some unique elements, will elicit fear to the extent that the overlapping elements have remaining excitatory value after extinction.

Overview of findings to date

Empirical evidence to date suggests that exposure to different elements of a CS+ via extinction with a GCS, without the CS+ itself, is not sufficient to extinguish fear of the CS+ (Vervliet & Greens, 2014). For example, in an early study, Vervliet et al. (2005), assigned fear conditioned participants to either a standard extinction group, which received four presentations of the CS+ and CS-, or to a generalization group, which received four presentations of two GCS. The generalization stimuli were perceptually and categorically similar to the CSs (i.e., different shapes with similar features). No significant differences were found between the groups during extinction. However, when the original CSs were presented during test, the generalization group exhibited larger responses to the CS+ than to

¹ Although the current discussion is couched in terms of the Rescorla-Wagner model alternative approaches such as Pearce and Hall (1980) or the more recent latent cause models by Courville, Daw, and Touretzky (2006) or Gershman, Monfils, Norman, and Niv (2017) will also account for the findings. The Rescorla-Wagner model was chosen for its relative simplicity and given that it is the first formal model of Pavlovian conditioning.

the CS- which was not observed in the standard extinction group. Extinction with the GCSs resulted in extinction retention for the GCS but not the CS+. Moreover, Vervliet, Vansteenwegen, and Eelen (2004), found that extinction using the CS+ reduced fear to the CS+ and the GCS more so than extinction with the GCS only, presumably because extinction to the CS+ extinguished all overlapping elements shared by the CS+ and the GCS. Hence, extinction to the CS+ appears to be more beneficial than extinction to the GCS in extinguishing fear to the CS+ and generalizing fear extinction learning to other stimuli.

More recently, Vervoort, Vervleit, Bennett and Baeyens (2014) extended this work based primarily on perceptual similarity of GCS to include conceptually related GCS. After an experimental procedure in which four categories of abstract figures were formed and a member from one category was paired with an aversive electrical stimulus while the other was not, it was found that conditional fear responses generalized to the other members of the conditioned category as indicated by skin conductance and online shock-expectancy. Furthermore, subsequent extinction of the conditional stimulus generalized to the other category members. However, extinguishing a non-conditioned member failed to reduce fear of the conditioned member itself. Thus, similar to the prior findings with perceptually similar GCSs, fears generalized readily across conceptually related stimuli, but the strength of extinction generalization was stronger for the conditional stimulus than for conceptually similar stimuli.

Xu et al. (2018) further extended this work by examining extinction with multiple stimuli that varied elements of the CS by targeting topological features (e.g., a grey shape with a smaller white shape inside it). Group 1 completed extinction with the CSs only, group 2 with one topological GCS only, and group 3 with three topological GCSs. No group differences were observed during extinction, however at retest, partially consistent with the earlier studies (Vervliet et al., 2004; 2005; Vervoort et al., 2014), online US expectancy

ratings revealed more return of fear for groups extinguished with GCSs only, however, skin conductance responses suggested that extinction retention occurred for all groups.

Similar findings showing greater benefit following extinction with the CS+ than GCS have been observed even when varying the number of different GCS used during extinction. For example, after differential fear conditioning, Zbozinek and Craske (2018) divided participants into three conditions; one in which participants completed the extinction phase with nine presentations of the CS+, another with nine presentations of one GCS, and the third with one trial each of nine different GCS. A spontaneous recovery test was conducted one week later, with a selection of the GCS and CSs. Similar to prior studies, spontaneous recovery was not observed after exposure to the CS+ compared to exposure to a single or multiple GCS during extinction. However, extinction with multiple GCS reduced fear of those GCSs more than extinction with a single GCS. These results add to prior studies and suggest that extinction learning may be stronger and more generalizable when extinction training includes the original CS+.

In accord, Waters, Kershaw and Lipp (2018) assessed whether generalization of extinction training, using fear relevant stimuli (i.e., dogs), could be enhanced if, the original CSs presented during acquisition and several additional stimuli that were categorically similar to the CS+ were also presented during extinction training (i.e., two other dogs). The additional stimuli contained unique elements whereby one was perceptually more similar and the other perceptually more different to the CS+ dog. Participants were trained with two CSs (images of dogs), one of which was followed by an aversive acoustic US. During extinction, one group was presented with the CSs used during acquisition only whereas a second group was presented with the CSs as well as the two additional GCS dogs. This study also included a generalization test phase after extinction (in which two novel test dogs were presented) and a spontaneous recovery test phase with the original CSs. Participants in the multiple stimulus

condition exhibited larger SCRs during extinction, suggestive of greater physiological arousal, but reduced SCRs to the novel test dogs and less spontaneous recovery of SCRs to the CSs compared to the standard extinction control group. Findings are partially consistent with the Rescorla-Wagner model that would predict that extinction with both the CS+ and GCS would extinguish overlapping and unique elements of the CS+ and GCS and thus, limit responding to the novel test dogs to novel stimulus elements not shared with CS+ or GCS.

Struyf, Hermans, and Vervliet (2018) took a different approach to studying the effects of extinction training with stimuli that are not the CS+. They reasoned that the GCSs used in past studies did not only differ from the CS+ perceptually, but also in their ability to elicit a fear response. Thus, Struyf et al. (2018) compared extinction with the CS+ to extinction with a GCS that elicited little fear (weak GCS) and a GCS that elicited the same amount of fear as did the CS+ (peak GCS). This was achieved by employing pictures of five faces morphed on a continuum from neutral to fearful (neutral, low fear, medium fear, high fear, and extreme fear). The neutral face served as CS- and the medium fear face served as CS+. Across three groups of participants, extinction comprised 6 presentations of either the low fear face (weak GCS), the CS+, or the extreme fear face (peak GCS). It should be noted that the weak GCS was perceptually more similar to the CS+ than the peak GCS. The peak GCS expressed more fear than did the weak GCS and the CS+ and at the beginning of extinction training motivated a higher expectancy of the unconditional stimulus than the weak GCS. Replicating past findings, extinction with the weak GCS resulted in less generalization of extinction (larger unconditioned stimulus expectancy to CS+ and peak GCS during a generalization test) than did extinction with the peak GCS or the CS+. Extinction with the peak GCS, however, reduced US expectancy and electrodermal responding to the CS+ in a manner similar to extinction with the CS+ itself. Moreover, it was superior in reducing US expectancy and electrodermal responding to the peak GCS relative to extinction with the CS+. These results

suggest that exposure to a stimulus that elicits a stronger emotional response and strong expectations of a negative outcome may result in more generalization than exposure to the trained stimulus itself. These results are important as they emphasise the role of emotional arousal over the perceptual similarity between CS and GCS. It should be noted, however, that the current results may hinge on the use of a peak GCS that was similar to the CS+ (same facial identity) but signalled a larger threat level due to a stronger facial expression, a stimulus situation that may be difficult to recreate in the clinic.

Potential underlying mechanisms

Taken together, the findings to date suggest that extinction is strongest when the CS+ is present in addition to other generalization stimuli during extinction compared to extinction with either the CS+ alone or GCS alone – unless the GCS can elicit fear to a similar extent to the CS+ (Struyf et al., 2018). From a Rescorla-Wagner (1972) perspective, this may be because the CS+ and the GCS undergo extinction of both their common and unique elements, thus widening the range of stimulus elements that are now associated with absence of the US (noUS). Alternatively, the presence of the GCS and the CS- during the learning of the CS+-noUS association during extinction may enhance prediction error. That is, the absence of the US on GCS and CS- trials might enhance the processing of US omission on CS+ trials during extinction (Pearce & Hall, 1980).

Exposure to multiple other stimuli not associated with the US could also weaken or eliminate the original CS-US fear association (Kindt, 2018). Although the studies to date cannot rule this out, one would expect that weakening of the original CS-US association would result in reduced rather than increased fear responding during extinction with multiple stimuli as found in Waters et al. (2018). It is also possible that the presentation of additional GCS and the CS+ during extinction enhances engagement and arousal and the experience of greater emotional variability during extinction training (e.g., Rowe & Craske, 1998; Waters et

al., 2015) in turn, makes learning about the new CS+-noUS association more salient. This may enhance the likelihood of recalling extinction memories rather than the original fear memories during subsequent re-exposure to fear-related stimuli.

Although multiple stimuli have been utilized during treatment-outcome studies of exposure-based therapies with anxious patients, the CS+ has rarely, if ever, been presented in these studies, and the therapeutic procedures have relied entirely on the presentation of GCS (e.g., Waters, et al., 2014; see Öst & Ollendick, 2018). Moreover, being clinical trials focused on treatment outcomes rather than on elucidating underlying mechanisms of action, these studies rarely include suitable comparison conditions to determine whether presentation of multiple stimuli produces superior treatment-outcomes or to determine the mechanisms that mediate these effects. However, some clinical analogue studies have compared fear responding to novel stimuli and the original CSs after exposure to either multiple feared stimuli (akin to multiple GCS) or a single feared stimulus (akin to a single GCS, given the original CS+ was not included in either condition). Similar to our experimental study (Waters et al., 2018), these studies have found evidence of increased physiological and emotional reactivity during extinction, enhanced extinction generalization (i.e., less fear responding to novel stimuli post-extinction), and enhanced extinction retention (i.e., less fear responding to the original test stimulus) in the multiple stimulus group. For example, Rowe and Craske (1998) found more fear across exposure trials and a trend towards higher anxiety post-treatment in response to the original test spider, but also less fear to a novel spider in spider phobic adults exposed to four different spider stimuli during extinction compared to repeated exposure to the same spider. In another study (Shiban, Schelhorn, Pauli, & Mühlberger, 2015), spider-phobic patients were randomly allocated virtual reality (VR) exposure treatment in either one or four different contexts and were exposed to either one or four different spiders. All participants completed both a VR test with a novel spider in a novel

context and an in vivo behavioral avoidance test (BAT) pre-, post-treatment and at follow-up. Short- and long-term effects of fear attenuation and behavioral avoidance were observed only in the multiple stimulus condition, and there was no evidence for superiority of the combined multiple stimuli and contexts condition. Thus, the implementation of multiple stimuli during exposure therapy seems to have both short-term and long-term beneficial effects on treatment outcomes compared to single stimuli alone. However, these studies have not included the original CS+ during exposure, which should further strengthen effects, and further studies that test the robustness of the effects against mechanisms of relapse such as reinstatement and renewal are required.

Clearly, the work on multiple stimuli during extinction is in its infancy and further research is required to determine the mechanisms underlying the beneficial effects of including both the original CS+ and GCS that differ in the extent to which they elicit fear during extinction training. Moreover, as the original CS+ is typically unavailable to work with in exposure therapy, further research is required to determine the stimulus elements of GCSs that might make them as potent as the CS+ in strengthening extinction learning, as well as other novel approaches to help recreate the CS+ during extinction training (e.g., use of VR technology to recreate the CS+ as closely as possible to its original form or imagery of the CS+ in addition to exposure therapy with GCS).

Conclusions

In summary, novelty-facilitated extinction, presenting the US during extinction, either unpaired or in occasional CS-US pairings, and including multiple exemplars of the CS+ stimulus category during extinction have been shown to be promising techniques to strengthen extinction and prevent relapse in human fear conditioning and extinction research. More work is needed to generalize effects to other stimuli, test robustness across different manipulations of inducing relapse and measures of fear, elucidate the mechanism(s)

underlying each manipulation, and develop novel procedures for translating laboratory procedures into clinical practice. We expect that elaboration of these novel approaches will render the next 100 years of human fear conditioning and extinction research as informative and exciting as the first 100.

References

- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6(4), 451-474. doi:10.1016/0010-0285(74)90021-8
- Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York, NY: McGraw-Hill.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, 52, 976-986. doi:10.1016/S0006-3223(02)01546-9
- Bouton, M. E., Woods, A. M., & Pineño, O. (2004). Occasional reinforced trials during extinction can slow the rate of rapid reacquisition. *Learning and Motivation*, 35(4), 371-390. doi:10.1016/j.lmot.2004.05.001
- Capaldi, E. J. (1994). The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. *Psychonomic Bulletin & Review*, 1(2), 156-181. doi:10.3758/BF03200771
- Carleton, R. N., Norton, M. A. P. J., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders*, 21(1), 105-117. doi:10.1016/j.janxdis.2006.03.014
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, 10(7), 294-300. doi:https://doi.org/10.1016/j.tics.2006.05.004
- Craske, M., Treanor, M., Conway, C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10-23. doi:10.1016/j.brat.2014.04.006
- Dunsmoor, J. E., Campese, V. D., Ceceli, A. O., LeDoux, J. E., & Phelps, E. A. (2015). Novelty-Facilitated Extinction: Providing a Novel Outcome in Place of an Expected

Threat Diminishes Recovery of Defensive Responses. *Biological Psychiatry*, 78(3), 203-209.
doi:10.1016/j.biopsych.2014.12.008

Dunsmoor, J. E., Kroes, M. C. W., Li, J., Daw, N. D., Simpson, H. B., & Phelps, E. A. (2019). Role of Human Ventromedial Prefrontal Cortex in Learning and Recall of Enhanced Extinction. *The Journal of Neuroscience*, 39(17), 3264.
doi:10.1523/JNEUROSCI.2713-18.2019.

Dunsmoor, J. E., Kroes, M. C. W., Moscatelli, C. M., Evans, M. D., Davachi, L., & Phelps, E. A. (2018). Event segmentation protects emotional memories from competing experiences encoded close in time. *Nature Human Behaviour*, 2(4), 291-299.
doi:10.1038/s41562-018-0317-4

Else, J. W. B., Van Ast, V. A., & Kindt, M. (2018). Human memory reconsolidation: A guiding framework and critical review of the evidence. *Psychological Bulletin*, 144(8), 797-848. doi:10.1037/bul0000152

Escobar, M., Arcediano, F., & Miller, R. R. (2001). Conditions favoring retroactive interference between antecedent events (cue competition) and between subsequent events (outcome competition). *Psychonomic Bulletin & Review*, 8(4), 691-697.
doi:10.3758/BF03196205

Ezzyat, Y., & Davachi, L. (2011). What Constitutes an Episode in Episodic Memory? *Psychological Science*, 22(2), 243-252. doi:10.1177/0956797610393742

Gershman, S. J., Monfils, M.-H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *eLife*, 6, e23763. doi:10.7554/eLife.23763

Gershman, S., Jones, C., Norman, K., Monfils, M.-H., & Niv, Y. (2013). Gradual extinction prevents the return of fear: implications for the discovery of state. *Frontiers in Behavioral Neuroscience*, 7(164). doi:10.3389/fnbeh.2013.00164

- Jelinek, L., Hauschildt, M., Hottenrott, B., Kellner, M., & Moritz, S. (2018). "Association splitting" versus cognitive remediation in obsessive-compulsive disorder: A randomized controlled trial. *Journal of Anxiety Disorders*, *56*, 17-25.
doi:10.1016/j.janxdis.2018.03.012
- Jones, M. C. (1924a). The Elimination of Children's Fears. *Journal of Experimental Psychology*, *7*(5), 382-390. doi:10.1037/h0072283
- Jones, M. C. (1924b). A laboratory study of fear: The case of Peter. *The Pedagogical Seminary and Journal of Genetic Psychology*, *31*(4), 308-315.
doi:10.1080/08856559.1924.9944851
- Kindt, M. (2018). The surprising subtleties of changing fear memory: a challenge for translational science. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1742), 20170033. doi:10.1098/rstb.2017.0033
- Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature Neuroscience*, *12*(3), 256-258.
doi:10.1038/nn.2271
- Kryptos, A.-M., & Engelhard, I. M. (2018). Testing a novelty-based extinction procedure for the reduction of conditioned avoidance. *Journal of Behavior Therapy and Experimental Psychiatry*, *60*, 22-28. doi:10.1016/j.jbtep.2018.02.006
- Lipp, O. V., Sheridan, J., & Siddle, D. A. T. (1994). Human Blink Startle During Aversive and Nonaversive Pavlovian Conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(4), 380-389. doi:10.1037/0097-7403.20.4.380
- Lovibond, P. F., Davis, N. R., & O'Flaherty, A. S. (2000). Protection from extinction in human fear conditioning. *Behaviour Research and Therapy*, *38*(10), 967-983.
doi:10.1016/S0005-7967(99)00121-7

Lubow, R. E. (1989). *Latent inhibition and conditioned attention theory*. Cambridge: Cambridge University Press.

Lucas, K., Luck, C. C., & Lipp, O. V. (2018). Novelty-facilitated extinction and the reinstatement of conditional human fear. *Behaviour Research and Therapy*, *109*, 68-74.
doi:10.1016/j.brat.2018.08.002

Mallan, K. M., Lipp, O. V., & Cochrane, B. (2013). Slithering snakes, angry men and out-group members: What and whom are we evolved to fear? *Cognition and Emotion*, *27*(7), 1168-1180. doi:10.1080/02699931.2013.778195

Matute, H., Lipp, O. V., Vadillo, M. A., & Humphreys, M. S. (2011). Temporal contexts: Filling the gap between episodic memory and associative learning. *Journal of Experimental Psychology: General*, *140*(4), 660-673. doi:10.1037/a0023862

McEvoy, P. M., Moulds, M. L., Grisham, J. R., Holmes, E. A., Moscovitch, D. A., Hendrie, D., . . . Erceg-Hurn, D. M. (2017). Assessing the efficacy of imagery-enhanced cognitive behavioral group therapy for social anxiety disorder: Study protocol for a randomized controlled trial. *Contemporary Clinical Trials*, *60*, 34-41.
doi:10.1016/j.cct.2017.06.010

Mickley, A. G., DiSorbo, A., Wilson, G. N., Huffman, J., Bacik, S., Hoxha, Z., . . . Kim, Y.-H. (2009). Explicit disassociation of a conditioned stimulus and unconditioned stimulus during extinction training reduces both time to asymptotic extinction and spontaneous recovery of a conditioned taste aversion. *Learning and Motivation*, *40*(2), 209-220. doi:10.1016/j.lmot.2009.01.001

Moritz, S., Jelinek, L., Klinge, R., & Naber, D. (2007). Fight Fire with Fireflies! Association Splitting: A Novel Cognitive Technique to Reduce Obsessive Thoughts. *Behavioural and Cognitive Psychotherapy*, *35*(5), 631-635.
doi:10.1017/S1352465807003931

Neumann, D. L., Lipp, O. V., & Cory, S. E. (2007). Conducting extinction in multiple contexts does not necessarily attenuate the renewal of shock expectancy in a fear-conditioning procedure with humans. *Behaviour Research and Therapy*, *45*(2), 385-394. doi:10.1016/j.brat.2006.02.001

Öhman, A. (1986). Face the beast and fear the face: Animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology*, *23*, 123-145. doi:10.1111/j.1469-8986.1986.tb00608.x

Öst, L-G., & Ollendick, T. H. (2017). Brief, intensive and concentrated cognitive behavioural treatments for anxiety disorders in children: A systematic review and meta-analysis. *Behaviour Research and Therapy*, *97*, 134-145. doi: 10.1016/j.brat.2017.07.008. Epub 2017 Jul 17.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552. doi:10.1037/0033-295X.87.6.532

Rauhut, A. S., Thomas, B. L., & Ayres, J. J. B. (2001). Treatments that weaken Pavlovian conditioned fear and thwart its renewal in rats: Implications for treating human phobias. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*(2), 99-114. doi:10.1037/0097-7403.27.2.99

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*, 77-94. doi:10.1037/h0027760

Rescorla, R. A., & Skucy, J. C. (1969). Effect of response-independent reinforcers during extinction. *Journal of Comparative and Physiological Psychology*, *67*(3), 381-389. doi:10.1037/h0026793

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. Prohazy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appleton Century Crofts.

Ricker, S. T., & Bouton, M. E. (1996). Reacquisition following extinction in appetitive conditioning. *Animal Learning & Behavior*, *24*(4), 423-436.
doi:10.3758/BF03199014

Rowe, M. K., & Craske, M. G. (1998). Effects of varied-stimulus exposure training on fear reduction and return of fear. *Behaviour Research and Therapy*, *36*(7), 719-734.
doi:10.1016/S0005-7967(97)10017-1

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*, 49-53. doi:10.1038/nature08637

Shiban, Y., Schelhorn, I., Pauli, P., & Mühlberger, A. (2015). Effect of combined multiple contexts and multiple stimuli exposure in spider phobia: A randomized clinical trial in virtual reality. *Behaviour Research and Therapy*, *71*, 45-53.
doi:10.1016/j.brat.2015.05.014

Shiban, Y., Wittmann, J., Weißinger, M., & Mühlberger, A. (2015). Gradual extinction reduces reinstatement. *Frontiers in Behavioral Neuroscience*, *9*(254).
doi:10.3389/fnbeh.2015.00254

Siddle, D. A. T., & Lipp, O. V. (1997). Orienting, habituation and information processing: The effects of omission, the role of expectancy, and the problem of dishabituation. In P. J. Lang, R. F. Simons, & M. Balaban (Eds.), *Attention and orienting: Sensory and motivation processes* (pp. 23-40). London: Erlbaum.

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

Thomas, B. L., Longo, C. L., & Ayres, J. J. B. (2005). Thwarting the renewal (relapse) of conditioned fear with the explicitly unpaired procedure: Possible interpretations and implications for treating human fears and phobias. *Learning and Motivation*, 36(4), 374-407. doi:10.1016/j.lmot.2004.11.005

Thompson, A., & Lipp, O. V. (2017). Extinction during reconsolidation eliminates recovery of fear conditioned to fear-irrelevant and fear-relevant stimuli. *Behaviour Research and Therapy*, 92, 1-10. doi:10.1016/j.brat.2017.01.017

Thompson, A., McEvoy, P. M., & Lipp, O. V. (2018). Enhancing extinction learning: Occasional presentations of the unconditioned stimulus during extinction eliminate spontaneous recovery, but not necessarily reacquisition of fear. *Behaviour Research and Therapy*, 108, 29-39. doi:10.1016/j.brat.2018.07.001

Treanor, M., Brown, L. A., Rissman, J., & Craske, M. G. (2017). Can memories of traumatic experiences or addiction be erased or modified? A critical review of research on the disruption of memory reconsolidation and its applications. *Perspectives on Psychological Science*, 12(2), 290-305. doi:10.1177/1745691616664725.

van den Akker, K., Havermans, R. C., & Jansen, A. (2015). Effects of occasional reinforced trials during extinction on the reacquisition of conditioned responses to food cues. *Journal of Behavior Therapy and Experimental Psychiatry*, 48, 50-58. doi:10.1016/j.jbtep.2015.02.001

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

Vervliet, B., Vansteenwegen, D., Baeyens, F., Hermans, D., & Eelen, P. (2005). Return of fear in a human differential conditioning paradigm caused by a stimulus change after extinction. *Behaviour research and therapy*, 43(3), 357-371.

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory*, *11*(5), 555-558.

Vervliet, B., Vansteenwegen, D., & Hermans, D. (2010). Unpaired shocks during extinction weaken the contextual renewal of a conditioned discrimination. *Learning and Motivation*, *41*(1), 22-31. doi:10.1016/j.lmot.2009.08.001

Vervoort, E., Vervliet, B., Bennett, M., & Baeyens, F. (2014). Generalization of Human Fear Acquisition and Extinction within a Novel Arbitrary Stimulus Category. *PLoS ONE*, *9*(5), e96569. doi:10.1371/journal.pone.0096569

Waters, A. M., Farrell, L. J., Zimmer-Gembeck, M. J., Milliner, E., Tiralongo, E., Donovan, C. L., McConnell, H., Bradley, B. P., Mogg, K., & Ollendick, T. H. (2014). Augmenting one session treatment of children's specific phobias with attention training towards positive stimuli. *Behaviour Research and Therapy*. *62*, 107-119.

Waters, A. M., Kershaw, R., & Lipp, O. V. (2018). Multiple fear-related stimuli enhance physiological arousal during extinction and reduce physiological arousal to novel stimuli and the threat conditioned stimulus. *Behaviour Research and Therapy*, *106*, 28-36. doi:10.1016/j.brat.2018.04.005

Waters, A. M., & Le Beau, R. & Craske, M. G. (2017). Experimental psychopathology and clinical psychology: An integrative model to guide clinical science and practice. *Psychopathology Review*. *4*, 112-128.

Waters, A. M., Potter, A., Jameson, L., Bradley, B. P., & Mogg, K. (2015). Predictors of treatment outcome in anxious children receiving group-based cognitive behavioural therapy: Pre-treatment attention bias for threat and emotional variability during exposure tasks. *Behaviour Change*. *32*, 143-158.

Woods, A. M., & Bouton, M. E. (2007). Occasional reinforced responses during extinction can slow the rate of reacquisition of an operant response. *Learning and Motivation*, 38(1), 56-74. doi:10.1016/j.lmot.2006.07.003

Xu, L., Su, H., Xie, X., Yan, P., Li, J., Xifu, Z. (2018). The topological properties of stimuli influence fear generalization and extinction in humans. *Frontiers of Psychology*, 9, 409. doi: [10.3389/fpsyg.2018.00409](https://doi.org/10.3389/fpsyg.2018.00409)

Zbozinek, T. D., & Craske, M. G. (2018). Pavlovian extinction of fear with the original conditional stimulus, a generalization stimulus, or multiple generalization stimuli. *Behaviour Research and Therapy*, 107, 64-75.