Check for updates

Empirical Paper

JOURNAL OF EXPERIMENTAL PSYCHOPATHOLOGY

Journal of Experimental Psychopathology
April-June 2023: 1–16
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20438087231175147
journals.sagepub.com/home/jepp
Sage

Reducing avoidance of learnt fear: Extinction of an imminent threat signal partly decreases costly avoidance to a distal threat signal

Alex H. K. Wong
Department of Psychology, Educational Sciences, and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands

Department of Psychology (Biological Psychology, Clinical Psychology, and Psychotherapy), University of Würzburg, Würzburg, Germany

Andre Pittig
Translational Psychotherapy, Institute of Psychology, Georg-August-Universität Göttingen, Göttingen, Germany

Abstract

In the interplay of fear and avoidance, not only imminent threat signals that directly predict potential threat evoke avoidance, but also distal threat signals that predict these imminent threat signals. Avoidance of learnt fear refers to avoidance to a distal threat signal that prevents the occurrence of an imminent threat signal. In clinical anxiety, it is often pathological given its persistence in the absence of threat and the impairments it inflicts. The current study examined whether fear extinction to an imminent threat signal would effectively reduce avoidance of learnt fear in a sensory preconditioning procedure. Three neutral preconditioning stimuli (PSs), serving as distal threat or safety signals, were paired with three neutral to-be conditioned stimuli (CSs), serving as imminent threat or safety signals. After assessing baseline levels of costly avoidance to the PSs, two CSs were paired with threat. One of these CSs then ceased to predict threat during extinction training. In test, participants showed limited avoidance to the PS that signaled the extinguished CS, however, the level of avoidance was still stronger compared to a PS that signaled a safety CS. Results suggest that exposure to an imminent threat signal partly reduces avoidance to a distal threat.

Keywords
Avoidance; Fear conditioning; High-order conditioning; Extinction; Anxiety

Introduction

Fear-related avoidance generally refers to behavioral responses to a threat signal that prevent expected imminent harm (i.e., safety behavior). Moreover, not only these imminent threat signals are avoided, but also stimuli that predict these signals (distal threat signals). Unlike escape, which terminates the presentation of these signals, avoidance reduces or prevents these signals. We previously referred to this avoidance in response to distal threat signals as "avoidance of learnt fear" (Wong et al., 2022). For instance, an individual with dog phobia may take a lengthy detour around the park (distal threat signal). While this action does not terminate the distal threat signal (i.e., seeing the park during the detour), it prevents encountering people walking

Corresponding author:
Andre Pittig, Translational Psychotherapy, Institute of Psychology, Georg-August-Universität Göttingen, Kurze-Geismar-Straße 1, Göttingen 37073, Germany.
Email: andre.pittig@uni-goettingen.de

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

dogs (imminent threat signal) in the park, which is directly associated with the perceived aversive outcome (dog attack). In fact, empirical studies have shown that individuals with clinical anxiety often engage in avoidance of learnt fear. For instance, individuals with specific phobias reported avoiding places or stimuli that signal the presence of phobic-related stimuli and hence the perceived aversive outcome (e.g., Katz, 1974; Kleinknecht & Lenz, 1989; Sawchuk et al., 2000; Walz et al., 2016). Avoidance of learnt fear in clinical anxiety is often pathological, especially when it is costly and inflicts severe impairments. For example, clinically anxious individuals take a lengthy detour to avoid places where fear-related stimuli or trauma reminders are expected to be encountered (e.g., Corrigan et al., 2007). Furthermore, pathological avoidance of learnt fear is unnecessary particularly in the absence of realistic threat. Thus, it is of clinical importance to further the understanding of avoidance of learnt fear and its reduction.

Avoidance of learnt fear can be systematically examined in highly controlled laboratory paradigms. We have recently proposed a variety of laboratory procedures for examining avoidance of learnt fear (Wong et al., 2022), including higher-order conditioning. This procedure builds on a typical fear conditioning procedure, in which a neutral conditioned stimulus (CS) is repeatedly paired with an aversive unconditioned stimulus (US). On top of this, a so-called higher-order CS is either paired with the CS before (sensory preconditioning) or after (second-order conditioning) CS-US pairings. It is typically found that this higher-order CS evoked high threat expectancy or conditioned fear despite it has never been directly paired with a US (e.g., Davey & Arulampalam, 1982; Dunsmoor et al., 2011; Vansteenwegen et al., 2000; White & Davey, 1989). Importantly, responding to higher-order CS cannot be explained by stimulus generalization, given that a higher-order CS evoked little to no responses when it was not paired with a CS (e.g., Prewitt, 1967; Rizley & Rescorla, 1972).

When comparing second-order and sensory preconditioning, second-order conditioning comes with a methodological limitation: In typical second-order conditioning procedures, a higher-order CS is paired with a CS after CS-US pairings already took place; however, pairings between the higher-order CS and the CS are typically not reinforced by a US, thus rendering this procedure similar to conditioned inhibition (e.g., A+/AB-). This may lead to limited responses to the higher-order CS in second-order conditioning. This limitation is not inherent in sensory preconditioning, given that the pairings between higher-order CS (often termed preconditioning stimulus; PS) and the CS occurred before CS-US pairings. Thereby, it provides a useful testbed for examining avoidance of learnt fear. Furthermore, this procedure allows repeated assessment of avoidance to the distal threat signal before and after the CS (imminent threat signal) acquires threat value. This enables the investigation of how avoidance of learnt fear increases after CS-US pairings.

Despite the clinical importance of reducing avoidance of learnt fear, there are only few empirical studies. Based on the assumption that fear motivates avoidance (e.g., Krypotos et al., 2015; Pittig et al., 2020), it is valuable to first review sensory preconditioning studies examining the reduction of conditioned fear to a PS. Some preliminary studies examined whether extinguishing the PS – CS association effectively reduce conditioned fear to the PS (Coppock, 1958; Vansteenwegen et al., 2000). These studies first paired a PS with a neutral CS, but then immediately extinguished this association before CS-US pairings. This procedure, so-called pre-extinction, effectively reduced conditioned fear to the PS in test. However, pre-extinction provides limited practicality in a clinical context. First, a PS – CS association may reflect a realistic association. For instance, it is realistic to encounter people walking their dogs in a park (e.g., park-dog association), and thereby it is impractical to extinguish this association. Second, pre-extinction was carried out before fear acquisition, thus it could not be practically translated to anxiety-related treatments. Third, perhaps most importantly, the CS – US association is not targeted and thus remains intact, preserving conditioned fear to the CS (e.g., Debiec et al., 2006; Holmes et al., 2014).

Despite the limited clinical implication of pre-extinction, it corroborates with theoretical accounts on how a PS evokes conditioned fear. Pre-extinction supports the notion of a chain-like stimulus–stimulus structure, in other words, PS presentation activates CS representation, which then activates US representation (Gewritz & Davis, 2000). Weakening the PS-CS association cuts off the activation of the following CS – US association and thus reducing fear to the PS (Rescorla & Wagner, 1972; Wagner & Brandon, 1989). Similarly, extinguishing the CS – US association reduces conditioned fear to the PS (Archer & Sjödén, 1982; Rizley & Rescorla, 1972). Returning to the notion that fear motivates avoidance, it is expected that extinguishing the CS-US association would also decrease avoidance responses to the PS.

The current study thus sought to examine whether extinguishing the CS – US association would effectively reduce avoidance of learnt fear. Using a sensory preconditioning procedure, we paired one of three PSs with one of three CSs, respectively. After assessing baseline levels of avoidance to the PSs, two of these CSs then signaled a US, whereas the remaining CS signaled safety. In a following extinction procedure, one of the threat-related CS was no longer reinforced by a US, whereas the CS – US contingencies for the remaining CSs remained unchanged. At test, we assessed avoidance responses to the PSs. Of note, financial rewards were introduced to compete with avoidance responses. That is, engaging in avoidance responses led to a loss of reward, thus rendering avoidance of learnt fear costly. This manipulation more closely models costly, pathological avoidance in clinical anxiety (Pittig et al., 2020). It is expected that participants would show a reduction in costly avoidance of learnt fear to the PS that signaled an extinguished CS. The current study also explored whether individual difference factors that contribute to the development of clinical anxiety, such as trait anxiety or intolerance of uncertainty (Chambers et al., 2004; Fetzner et al., 2013; Jorm et al., 2000), would have any impact on avoidance of learnt fear. Trait anxiety is a predisposition tendency to experience negative emotional responses to situations in general, whereas intolerance of uncertainty is characterized by negative emotional responses to uncertainty. Preliminary evidence suggests these risk factors are associated with pathological avoidance (Andreatta et al., 2017; Flores et al., 2018; Pittig et al., 2014; Zuj et al., 2020; see also a systematic review from Wong et al., under review). Therefore, we exploratorily examined whether these risk factors affect the retention of avoidance of learnt fear after Pavlovian extinction.

Method

Participants

Undergraduates or residents from Würzburg, Germany, were recruited and were compensated by either partial course credit or 10€. Additionally, participants received extra financial reward depending on their overall avoidance performance throughout the entire experiment. According to a stimulation-based power analysis (Kumle et al., 2021), a sample size of 50 achieves 93% power to detect an expected effect size of b = 18.93 (see https://osf.io/yznaw for the preregistration). We recruited a total of 60 participants to account for attrition rates due to exclusion criteria or technical difficulties. The Ethics Committee of the Institute of Psychology at the University of Würzburg (GZ 2018-25) in accordance with the Declaration of Helsinki.

Apparatus and Materials

Three standardized 2D black and white drawings (apple, car, and dog) from Snodgrass and Vanderwart (1980) and three colored geometric shapes (orange triangle, purple hexagon, and red circle) served as visual stimuli presented in the experiment.

Presentation software (Neurobehavioral Systems Inc., Berkeley, CA, Version 20.1) was employed to present all experimental instructions, visual stimuli, and recorded all self-reported ratings. BrainVision Recorder (Brain Products, GmbH, Gliching, Germany) was used to measure skin conductance via two Ag/AgCl electrodes at a sampling rate of 1000 Hz. A DS7A Digitimer stimulator was used to generate an electric US which consisted of 125 pulses separated by 5 ms.

Procedure

After providing written informed consent, participants filled in the German version of the Intolerance of Uncertainty scale (UI-18; Freeston et al., 1994; Gerlach et al., 2008) and the German version of the Depression Anxiety Stress Scale-21 (DASS-21; Lovibond & Lovibond, 1995; Nilges & Essau, 2015). The UI-18 measures cognitive and behavioral responses to uncertainty (see Carleton et al., 2007), whereas the DASS-21 measures and distinguishes between depression, anxiety, and stress. Next, we attached US electrodes to the wrist of participants' non-dominant hand. Skin conductance electrodes filled with isotonic gel were also attached to the hypothenar muscles on the same hand.

A US workup procedure was carried out, in which a US intensity of 0.2 mA was increased gradually until a level that was individually perceived as "definitely unpleasant but not painful." Immediately after US calibration, participants were led through a reward-matching procedure. This procedure was identical to our recent study (Wong & Pittig, 2022a). In brief, participants were presented with a series of questions assessing how much monetary reward was sufficient for them to tolerate an electric stimulation. Participants had to answer either "yes" or "no" to each of these questions. The amount of competing reward between the highest amount that received a "No" and the lowest amount that received a "Yes" was chosen as the maximum reward available per trial. This calibrated level of competing reward was assumed to promote sufficient avoidance-approach conflict per trial (c.f. Schlund et al., 2016). The stimuli to be presented in the experiment were then presented individually on screen, and valence ratings to each of them were collected. The valence visual analog scale (−50 = unpleasant, 0 = neutral, +50 = pleasant) was located below the visual stimuli; participants were asked to indicate their valence ratings to each stimulus along the scale. After participants had indicated their valence ratings, the conditioning task was carried out. The conditioning task consisted of five phases: *Preconditioning stage, Baseline costly avoidance, Pavlovian fear acquisition training, Pavlovian extinction,* and *Post-extinction test* (see Table 1).

Preconditioning Stage

The black and white drawings (apple, car, and dog) served as PSext, PSsafe, and PSthreat, respectively. The color geometric shapes (orange triangle, purple hexagon, and red circle) served as CSext, CSsafe, and CSthreat, respectively. The PSs and CSs were counterbalanced across all participants. PSext always signaled the onset of CSext, PSsafe

Table 1. PS indicates preconditioning stimuli; CS indicates conditioned stimuli; + indicates US presentation; - indicates US omission; * indicates avoidance availability; CS and € in brackets indicate the presentation of a CS and a competing reward, respectively, depending on avoidance; Number in parentheses indicates the number of trials per trial type.

Preconditioning	Baseline Costly Avoidance	Pavlovian Fear Acquisition Training	Pavlovian Extinction	Post-extinction Test
PSext → CSext (8)	PSext* [CSext, €] (4)	CSext + (4)	CSext - (6)	PSext*- [€] (4)
PSsafe → CSsafe (8)	PSsafe* [CSsafe, €] (4)	CSsafe - (4)	CSsafe - (6)	PSsafe*- [€] (4)
PSthreat → CSthreat (8)	PSthreat* [CSthreat, €] (4)	CSthreat + (2)	CSthreat + (2)	PSthreat*- [€] (4)

always signaled the onset of CSsafe, whereas PSthreat always signaled the onset of CSthreat. The PSs were presented individually on screen alongside a question "Which of the following shape does this picture predict?". The PS remained on screen until response. Participants could respond by pressing 3 designated keys: pressing "A" indicated that the presented PS signaled CSext; pressing "G" indicated that the presented PS signaled CSsafe; pressing "L" indicated that the presented PS signaled CSthreat. After a response had been recorded, the CS that followed the PS was presented on screen for 4 s, followed by a 2 s feedback informing whether the participant's response was correct or not. Each PS was presented for 8 trials each in a pseudo-randomized order so that the same trial type would not be presented more than two times in a row. This pseudo-randomization of presentation order was applied to all the following phases. The intertrial intervals (ITIs) in this phase were 4 s.

Baseline Costly Avoidance:

Participants were instructed that they could avoid the outcome predicted by the PSs. This could be done by indicating the degree of avoidance on a dimensional avoidance scale presented alongside the PS. The avoidance scale ranged from 0% (Certainly not avoid) to 100% (Certainly avoid) with an interval of 1%. The avoidance ratings were negatively proportional to the chance of CS presentation. For example, an avoidance rating of 65% would lead to a 35% chance of CS presentation. In addition, participants were instructed that whenever avoidance was available, a competing reward would be presented on that trial. The amount of competing reward that could be received was, however, inversely proportional to the indicated avoidance. For example, an avoidance response of 65% would result in a loss of 65% of the maximum competing reward. Participants were informed that all rewards gained throughout the task would be paid off at the end of the task. On each trial, the PS and the avoidance scale were presented on screen until an avoidance rating had been indicated. The PS then remained on screen for 8 s. Of note, the avoidance response did not terminate PS presentation, so that it was not confounded with an escape response. Depending on the indicated avoidance ratings, either the corresponding CS or a white blank screen was presented for 4 s. Reward feedback was then presented for 2 s. The ITI was randomized between 11 and 15 s, and the same was applied to all the following phases.

Pavlovian Fear Acquisition Training:

Before this phase began, participants were instructed that avoidance responses and competing rewards were temporarily removed. CSext and CSsafe were presented 4 times each, whereas CSthreat was presented 2 times. CSext and CSthreat were fully reinforced by an electric US, whereas CSsafe was never reinforced. Each CS was presented along with a US expectancy scale with 100 intervals, ranging from 0% (Certain no electric stimulation), 50% (Uncertain) to 100% (Certain electric stimulation). After US expectancy ratings had been indicated, the CS remained on screen for 8 s, which was followed by a US or not depending on the CS type.

Pavlovian Extinction

This phase continued without any break. CSext and CSsafe were presented 6 times, and none of them were reinforced by a US. CSthreat was presented 2 times, in which it continued to be fully reinforced by a US. The rationale for the continuous reinforcement of CSthreat was to ensure that CSext and CSthreat had acquired the same excitatory strength, given they had the same number of reinforced trials. Similar to the previous phase, the CS was presented on screen alongside a US expectancy scale until response. The CS then stayed on screen for 8 s. A US was delivered immediately after CS offset, but only on CSthreat trials.

Test

Prior to this phase, participants were informed that avoidance and competing reward were made available again. The three PSs were presented 4 times each, along with an avoidance scale. After an avoidance rating had been indicated, a 1 s fixation cross appeared, and participants were then prompted to indicate their US expectancy. Once a US expectancy rating had been indicated, the PS stayed on screen for 8 s. None of these stimuli were reinforced by a CS

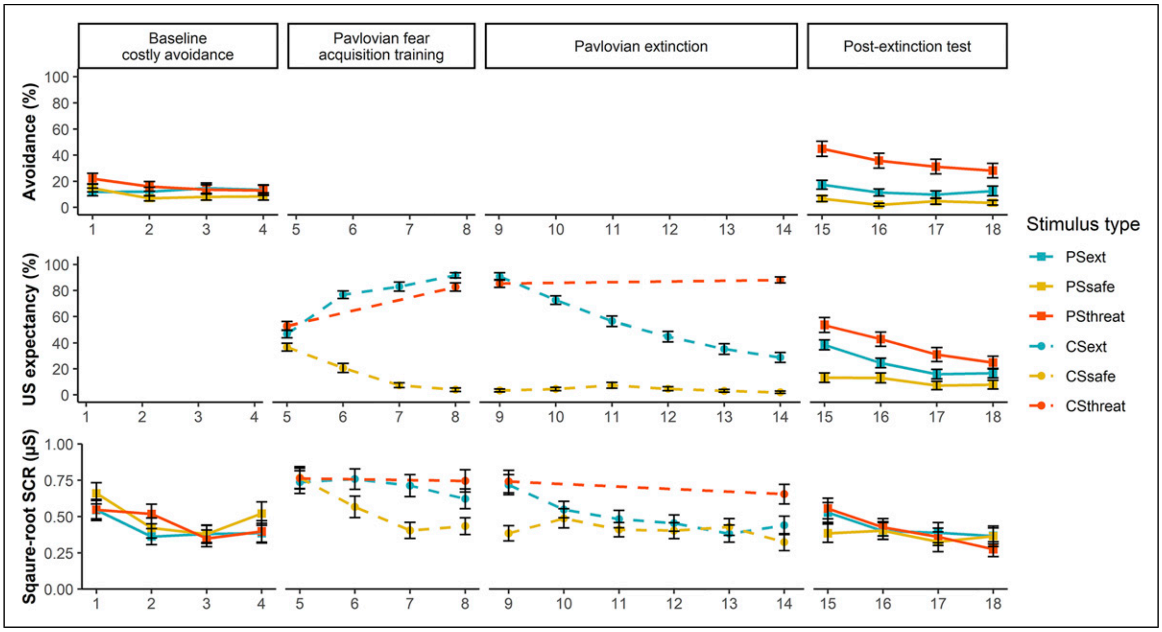

Figure 1. Avoidance (A), US expectancy (B), and square-root SCRs (C). For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

nor a US. Instead, a blank screen of 4 s appeared immediately after PS offset, followed by a reward feedback of 2 s.

After the experiment, participants were asked to indicate their valence ratings to each of the PSs and the CSs. They were then debriefed and thanked for their participation.

Scoring and Analysis

Skin conductance measured before *Pavlovian fear acquisition training* was not analysed due to the lack of anticipatory fear. Only skin conductance recorded during the 8 s of stimulus presentations were analysed (i.e., CSs in *Pavlovian fear acquisition training* and *Pavlovian extinction*, and PSs in *Post-extinction test*). However, to remain transparent, SCRs in all phases except the Preconditioning stage are presented in Figure 1. We applied a 1 Hz filter and a 50 Hz notch filter to the SCR data. Next, the difference between the peak response and the corresponding trough in the interval of 1 s after stimulus onset to stimulus offset was calculated. This procedure was first done by an automatic detection of the peak and minimum responses in each stimulus interval. We then checked if the minimum response was located at the trough corresponding to the maximum response; if not, we manually adjusted it accordingly. If no responses were detected, it would be scored as zero. At last, the processed SCR was then square root transformed to reduce skewness (Boucsein et al., 2012). The processing of SCR data was done by an experimenter blinded to the trial types.

Most data were analyzed within a linear mixed model framework. The analyses were separated into three parts: manipulation check, main hypotheses, and exploratory analyses. All analyses were preregistered (https://osf.io/yznaw).

Manipulation Check

We first checked whether participants acquired the PS-CS contingencies in the *Preconditioning* phase with an ANOVA test. The proportion of accurate responses served as dependent variable, whereas PS type (PSext, PSsafe, and PSthreat) and a linear trend repeated measures across trials (Trial) served as fixed effects.

We then checked whether the levels of baseline avoidance responses to the PSs differed from each other in the *Baseline Costly avoidance* phase with an ANOVA test. Avoidance responses served as dependent variable, whereas PS type and Trial served as fixed effects. Given that we expected null differences in baseline avoidance between PSs, we ran a Bayesian model to confirm an absence of an effect if the frequentist models failed to reject the null hypothesis (Kruschke, 2015). This Bayesian model calculated the posterior distribution of the 95% highest density intervals that fall under the area around the null value (Kruschke, 2015; Kruschke & Liddell, 2018).

Moreover, we checked whether participants acquired differential conditioned fear to the CSs in *Pavlovian fear acquisition training*. To this end, US expectancy ratings or SCRs served as dependent variable, whereas CS type

(CSext, CSsafe, and CSthreat) and Trial served as fixed effects. Noted that given there were only two CSthreat trials, the trial number of CSthreat was mapped onto the first and last trial of *Pavlovian fear acquisition training* in the linear mixed models. Two orthogonal contrasts were applied to this model. The first contrast entailed the comparison between responding averaged across CSext and CSthreat with responding to CSsafe, examining the differential responding to reinforced CSs and to a non-reinforced CS. The second contrast entailed the comparison between responding to CSext and CSthreat, to examine whether conditioned fear acquired to the reinforced CSs differed from each other. The interactions of these contrasts were of primary interest to assess the development of differential responding to the reinforced and non-reinforced CSs.

Main Hypotheses

We first examined changes in responding to the different CSs during *Pavlovian extinction*. US expectancy ratings or SCRs served as dependent variable, whereas CS type and Trial served as fixed effects. Similar to *Pavlovian fear acquisition training*, the two CSthreat trials in this phase was mapped onto the first and last trial of *Pavlovian extinction*. Noted that we deviated from our pre-registered analyses for one contrast, given that the current contrast was more sensitive in detecting whether threat expectancy or conditioned fear to CSext was effectively decreased when compared to CSthreat. Analyses from the pre-registration were reported in the Supplementary Materials. We established two non-orthogonal contrasts. First, we compared responding to CSext with responding to CSthreat. This contrast assessed whether responding to CSext decreased across extinction, whereas responding maintained at a strong level to a reinforced CSthreat. Second, we compared responding to CSext with responding to CSsafe. This contrast evaluated whether responding to CSext decreased across extinction trial compared to a safety CSsafe. These two non-orthogonal contrasts were Bonferroni-corrected.

Our primary interest was to examine whether extinguishing the CSext-US association would lead to a decrease in avoidance to PSext in *Post-extinction test*. To this end, we included two contrasts in this model. We again deviated from our pre-registered analyses for one contrast, given that the current contrast allowed us to focus more on the difference in avoidance to PSext and PSthreat. Analyses from the pre-registration were reported in the Supplementary Materials. Overall, the pre-registered analyses and the current analyses yielded highly similar results. First, we assessed whether avoidance to PSthreat was stronger to PSext. This contrast evaluated whether a PS that signaled a non-extinguished threat-related CS evoked stronger avoidance to a PS that signaled an extinguished threat-related CS. The second contrast assessed whether avoidance to PSext differed from avoidance to PSsafe. This evaluated whether CSext extinction completely extinguished avoidance to PSext. These two non-orthogonal contrasts were Bonferroni-corrected.

Furthermore, to examine whether changes in avoidance to PSs occurred before and after acquiring threat or safety value of the CSs, we employed a cross-phase analysis. Specifically, we compared avoidance responses on the first trial of *Post-extinction test* to the last trial of *Baseline costly avoidance*. The remaining trials in *Test* were excluded to preclude the ongoing effect of extinction learning. In this model, avoidance served as dependent variable, whereas PS type and Trial (last trial of *Baseline costly avoidance* compared to first trial of *Post-extinction test*) served as fixed effects. The same non-orthogonal contrasts employed in *Post-extinction test* were applied to this model (Contrast 1: PSext vs PSthreat; Contrast 2: PSext vs PSsafe).

To examine the change in valence ratings to the PSs and CSs before and after the experiment, valence ratings served as dependent variable, whereas Phase (pre- and post-experiment), Trial type (PSs vs CSs), and Threat type (PSext, PSthreat, CSext, CSthreat vs PSsafe and CSsafe) served as fixed effects.

We also assessed whether higher avoidance responses in *Post-extinction test* would be associated with a decrease in post-extinction threat expectancy or conditioned fear, especially to PSthreat. This set of analyses assessed whether avoidance of learnt fear was at least partly driven by threat prevention. Thus, US expectancy ratings or SCRs served as dependent variable, whereas Avoidance and Stimulus type served as fixed effects. Similarly, to examine whether the change in valence ratings would be associated with the extent of avoidance to the PSs in *Post-extinction test*, avoidance served as dependent variable whereas PS type served as a fixed effect. In these two sets of analyses, only the first trial in *Post-extinction test* was included to minimize the confounding effect of ongoing extinction. Two orthogonal contrasts were applied. The first one compared responding to PSthreat with responding averaged across PSext and PSsafe. The second contrast compared responding to PSext with responding to PSsafe.

Exploratory Analyses

Finally, we exploratorily examined whether trait anxiety or intolerance of uncertainty had any effect on avoidance to the PSs in *Post-extinction test*. The same models and contrasts employed in *Post-extinction test* were employed, with the addition of trait anxiety or intolerance of uncertainty as a continuous fixed effect. That is, avoidance served as dependent variable, whereas PS type, Trial, and Anxiety or Intolerance of uncertainty served as fixed effects, with one contrast comparing avoidance to PSext with PSthreat, and a second contrast comparing avoidance to PSext with PSsafe.

For all the aforementioned linear mixed models, participants served as a random effect. All main effects and higher-order interactions were analyzed in separate models (Hayes et al., 2012). The degree of significance was reported with Satterthwaite approximation for degrees of freedom (Satterthwaite, 1941).

Results

Analyses were restricted to participants who had 1) acquired the correct PS-CS contingencies and 2) acquired differential US expectancy ratings to the CSs. The first criterion was defined by scoring correct responses to the last two trials for each PS in *Preconditioning*, whereas the second criterion was defined by indicating higher US expectancy ratings to the last two trials of CSext and CSthreat compared to the last two trials of CSsafe in *Pavlovian fear acquisition training*. A total of 7 participants were excluded based on these criteria, leaving a total of 53 participants in the final sample (see Table 2). All exclusion criteria were preregistered. Data are available at https://osf.io/8fuqs/.

Table 2. Demographic data for the final sample.

	Mean (Standard Deviation)
Age	24.17 (4.03)
Sex – Females	37 (69.81%)
US intensity	0.84 mA (0.26)
DASS21-anxiety (0–42)	4.34 (5.18)
DASS21-depression (0–42)	5.85 (5.86)
DASS21-stress (0–42)	8.94 (7.85)
UI-18 (0–90)	38.96 (11.88)

Manipulation Check

Preconditioning and Baseline Costly Avoidance. For the *Preconditioning* phase, all participants acquired the correct PS – CS contingencies (see Supplementary Materials for detailed analyses). Figure 1 shows avoidance responses to the PSs. Overall, participants exhibited low levels of avoidance to the PSs during *Baseline costly avoidance*. Unexpectedly, we observed a significant interaction between Trial and PS type, $F(1, 583) = 3.30$, $p = .037$. Follow-up analyses revealed that avoidance to PSthreat decreased significantly sharper across trials compared to PSext, $b_{\text{PSext vs PSthreat}} = -104.75$, SE = 41.86, Bonferroni-corrected $p = .038$, whereas there was no evidence for any differences in avoidance between other PSs (smallest Bonferroni-corrected $p = .234$). As the CSs predicted by the PS had not yet acquired threat or safety value, in addition to all PSs were counterbalanced across participants, this difference was largely due to random variation.

Pavlovian Fear Acquisition Training

Figure 1 shows the US expectancy ratings across the different phases of the experiment. The first contrast examined whether differential responding was acquired to the reinforced CSs and the non-reinforced CS. Participants showed an increase in US expectancy ratings to CSext and CSthreat across trials, whereas a decrease in US expectancy ratings to CSsafe across trials, resulting in a significant interaction between CS type (CSext&CSthreat vs CSsafe) and Trial, $b_{\text{CS type(CSext\&CSthreat vs CSsafe)*Trial}} = -214.09$, SE = 14.20, $p < .001$. The second contrast examined whether conditioned fear or threat expectancy to the two reinforced CSs differed from each other. Unexpectedly, US expectancy to CSext increased faster than those to CSthreat, supported by a significant interaction between CS type (CSext vs CSthreat) and Trial, $b_{\text{CS type(CSext vs CSthreat)*Trial}} = 57.01$, SE = 25.04, $p = .023$. This pattern was presumably due to CSext acquiring stronger excitatory strength than CSthreat because of more CSext than CSthreat trials during *Pavlovian fear acquisition training*.

Figure 1 shows the SCRs across the different phases of the experiment. The same contrasts mentioned above were used. Participants developed stronger SCRs to CSext and CSthreat across trials and weaker SCRs to CSsafe across trials, confirmed by a significant interaction between CS type (CSext&CSthreat vs CSsafe) and Trial, $b_{\text{CS type(CSext\&CSthreat vs CSsafe)*Trial}} = -0.87$, SE = 0.26, $p < .001$. Unlike the US expectancy ratings, there was no evidence that the acquisition of conditioned fear to CSext differed from that to CSthreat, $b_{\text{CS type(CSext vs CSthreat)*Trial}} = 0.4$, SE = 0.45, $p = .311$.

In sum, participants successfully acquired the correct PS-CS and CS-US contingencies and exhibited low levels of baseline costly avoidance to the PSs.

Main Hypotheses

Pavlovian Extinction. Two non-orthogonal contrasts were analysed in this phase. The first contrast examined whether threat expectancy or SCRs was maintained to a reinforced CS compared to the extinction stimulus during *Pavlovian extinction*. US expectancy ratings to CSext decreased quickly across extinction trials, whereas US expectancies to CSthreat remained similar across extinction trials. This pattern was supported by a significant interaction between CS type (CSext vs CSthreat) and Trial, bCS type(CSext vs CSthreat)*Trial = 515.45, SE = 37.95, Bonferroni-corrected $p < .001$. The second contrast examined whether responding to CSext decreased across trials, compared to CSsafe that had never been associated with threat. US expectancy ratings to CSext rapidly across trials, whereas US expectancies to CSsafe remained stable at a low level across trials, bCS type (CSext vs CSsafe)*Trial = 516.02, SE = 34.18, Bonferroni-corrected $p < .001$. Given that US expectancy to CSext remained seemingly higher than to CSsafe on the last extinction trial, we conducted an additional analysis to check whether this difference was significant or not, despite this analysis not being pre-registered. Results confirmed that this difference was significant, bCS type(CSext vs CSsafe) = 26.81, SE = 3.67, Bonferroni-corrected $p < .001$, suggesting that extinction of US expectancy to CSext was incomplete.

With regard to the SCR data, responding to CSext decreased across extinction trials, whereas responding to CSthreat remained relatively stable across trials, bCS type(CSext vs CSthreat)*Trial = 1.50, SE = 0.60, Bonferroni-corrected $p = .026$. The second contrast confirmed that conditioned fear to CSext decreased more quickly than to CSsafe across extinction trials, bCS type(CSext vs CSsafe)*Trial = −1.76, SE = 0.59, Bonferroni-corrected $p = .006$. Similar to the US expectancy data, we carried out an additional non-preregistered analysis to check whether SCRs to CSext were successfully extinguished. There was no evidence that SCRs to CSext differed from CSsafe on the last extinction trial, bCS type(CSext vs CSsafe) = −0.12, SE = 0.068, Bonferroni-corrected $p = .186$. In sum, participants maintained strong threat expectancy and SCRs to the reinforced CSthreat compared to the (now) non-reinforced CSext. Importantly, participants exhibited a decrease in both threat expectancy and SCRs to CSext, but only SCRs to it were successfully extinguished.

Post-Extinction Test. Figure 1 shows the avoidance responses to the PSs in Post-extinction test. The first contrast examined whether avoidance responses to PSthreat would be greater than to PSext. Participants showed stronger avoidance responses to PSthreat compared to PSext averaged across test trials, $b_{\text{PS type(PSext vs PSthreat)}} = 22.21$, SE = 2.22, Bonferroni-corrected $p < .001$, suggesting that avoidance to a PS that signaled an extinguished threat-related CS was weaker than to a PS that signaled an non-extinguished threat-related CS . Although avoidance to PSthreat seemingly decreased more rapidly across trials compared to PSext, this interaction did not reach significance, $b_{\text{PS type(PSext vs PSthreat)*Trial}} = -89.26$, SE = 45.38, Bonferroni-corrected $p = .100$. The second contrast examined whether avoidance responses to PSext differed from PSsafe. Averaged across trials, avoidance responses to PSext were stronger than to PSsafe, $b_{\text{PS type(PSext vs PSsafe)}} = 8.49$, SE = 1.37, Bonferroni-corrected $p < .001$. No other effects reached significance (smallest $p = .131$).

Cross-Phase Analysis. The first contrast examined whether avoidance responses selectively increased to PSthreat compared to PSext from *Costly Baseline avoidance* to *Post-extinction test*. Results showed that avoidance responses increased more strongly to PSthreat compared to PSext across the transition of the two phases, $b_{\text{PS type(PSext vs PSthreat)*Trial}} = 27.94$, SE = 6.62, Bonferroni-corrected $p < .001$. The second contrast examined whether changes in avoidance responses to PSext differed from responses to PSsafe between the two phases. Although there was a descriptively stronger increase in avoidance responses to PSext compared to PSsafe during the transition between *Costly baseline avoidance* to *Post-extinction test*, this difference did not reach significance, $b_{\text{PS type(PSext vs PSsafe)*Trial}} = 5.42$, SE = 4.95, Bonferroni-corrected $p = .552$. On face value, this contradicted with the pattern of persistence avoidance to PSext during *Post-extinction test*. This null difference could be due to baseline avoidance to PSext being descriptively stronger than to PSsafe, leading to a weaker increase in avoidance in *Post-extinction test*.

Avoidance Predicting Conditioned Fear. On the first trial of *Post-extinction test*, an increase in avoidance to PSthreat led to a decrease in US expectancy ratings, whereas avoidance averaged across PSext and PSsafe led to an increase in US expectancy ratings (Figure 2(a)), $b_{\text{Avoidance*PS type(-PSext\&PSsafe vs PSthreat)}} = -0.23$, SE = 0.06, $p < .001$. Surprisingly, while avoidance to PSext had limited predictive value to subsequent US expectancy ratings, an increase in avoidance to PSsafe led to an increase in US expectancy ratings, $b_{\text{Avoidance*PS type(PSext vs PSsafe)}} = 0.31$, SE = 0.15, $p = .038$. For the SCRs data (Figure 2(b)), avoidance responses were not associated with subsequent SCRs to the PSs (smallest $p = .087$).

Valence Ratings and Their Correlation to Avoidance. Valence ratings to stimuli associated with an extinguished CS-US association (PSext and CSext) and those that did not (PSthreat and CSthreat) decreased from pre- to post-experiment, whereas valence of stimuli associated with safety throughout the entire experiment (PSsafe and CSsafe) increased (Figure 2(c)). This pattern was supported by a significant interaction between Phase and Threat type, $b_{\text{Phase*Threat type}} = -36.59$, SE = 3.61, $p < .001$. There was no evidence that valence changes differed between PSs and CSs (smallest $p = .078$).

In regard to whether a change in valence was associated with the magnitude of avoidance of learnt fear (Figure 2(d)), although the negative change in valence averaged across PSext and PSthreat was seemingly associated with an increase in avoidance, this association did not reach significance, $b_{\text{Valence change*PS type(PSext\&PSthreat vs PSsafe)}} = 0.086$, SE = 0.07, $p = .192$. Similarly, the association between valence change and avoidance to PSext did not differ from that to PSthreat, $b_{\text{Valence change*PS type(PSext vs PSthreat)}} = -0.14$, SE = 0.09, $p = .132$.

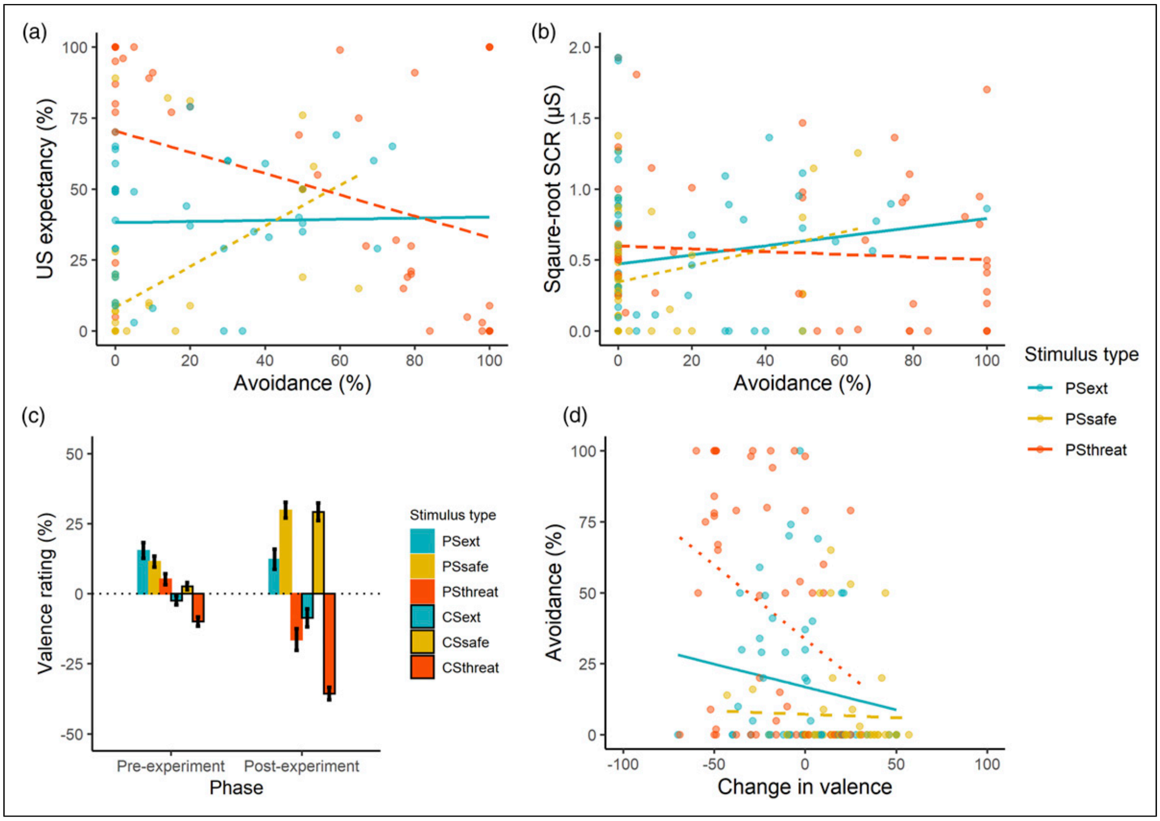

Figure 2. Top panel. Relationship between avoidance of learnt fear and the subsequent (A) US-expectancy ratings and (B) SCRs on the first trial of Test. Bottom Panel. Association of the change in valence and avoidance of learnt fear. (C) The comparison of valence ratings to the PSs and CSs pre- and post-experiment. (D) The relationship between the change in valence to the PSs and avoidance of learnt fear on the first trial of Test. Negative value on the x-axis indicates a negative change in valence, whereas a positive value indicates an opposite pattern. Darker color dots indicate more overlapping data points. The lines represent the line of best fit for each PS for visual aid. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

Exploratory Analyses

Risk Factors and Avoidance of Learnt Fear. We exploratorily examined the effect of trait anxiety and intolerance of uncertainty on avoidance of learnt fear during *Post-extinction test* (Figure 3). The first contrast compared avoidance responses to PSthreat to PSext. No interactions involving trait anxiety nor intolerance of uncertainty reached significance (smallest Bonferroni-corrected p = .326). That is, there was no evidence that either risk factor affected differential avoidance responses to PSext and PSthreat. The second contrast compared avoidance responses to PSext and PSsafe. With regard to trait anxiety, an increase in trait anxiety was associated with less decrease in avoidance responses to PSext compared to PSsafe, $b_{\text{PS type(PSext vs PSsafe)*Trial*Anxiety}} = -14.85$, SE = 5.42, Bonferroni-corrected p = .013. No other interactions involving trait anxiety reached significance (smallest Bonferroni-corrected p = .204). With regard to intolerance of uncertainty, no interactions involving it reached significance (smallest Bonferroni-corrected p = .212).

US Expectancy Ratings and SCRs in Post-Extinction Test. Although not pre-registered, we examined the US expectancy ratings and SCRs to the PSs in *Post-extinction test* with identical contrasts that we used for the avoidance data (Contrast 1: PSext vs PSthreat; Contrast 2: PSext vs PSsafe). These responses were assessed after participants performed their avoidance response. Participants showed higher US expectancy ratings to PSthreat than to PSext averaged across trials, $b_{\text{PS type(PSext vs PSthreat)}} = 14.15$, SE = 2.39, Bonferroni-corrected p < .001. However, there were no differences in the decrease in US expectancies across trials between the two PSs, $b_{\text{PS type(PSext vs PSthreat)*Trial}} = -55.86$, SE = 49.08, Bonferroni-corrected p = .512. Furthermore, the second contrast showed that the decrease in US expectancies to PSext were stronger than to

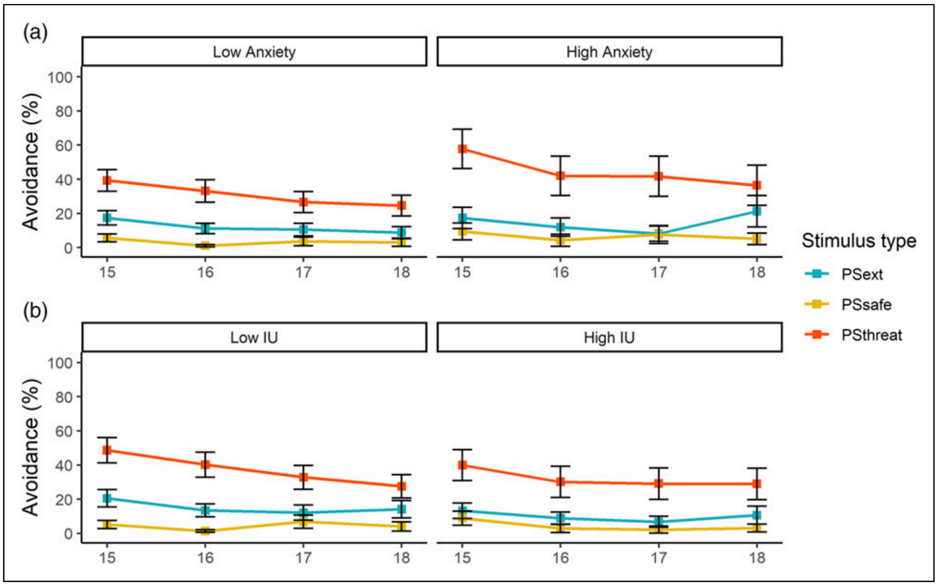

Figure 3. The effect of trait anxiety (A) or intolerance of uncertainty (B) on avoidance of learnt fear during Post-extinction test. Trait anxiety/Intolerance of uncertainty was divided into high and low values (via median split) for descriptive purpose. Error bars indicate standard error of the mean. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

PSsafe; avoidance to the former decreased more strongly than to the latter across test trials, $b_{\text{PS type(PSext vs PSsafe)*Trial}} = 119.84$, SE = 37.47, Bonferroni-corrected $p = .003$.

With regard to SCRs, the first contrast showed that there was a general decrease in SCRs across trials averaged across PSext and PSthreat, $b_{\text{Trial}} = -1.62$, SE = 0.33, Bonferroni-corrected $p < .001$. No other effects involving PS type reached significance (smallest Bonferroni-corrected $p = .306$). The second contrast showed that there were no differences concerning SCRs to PSext and PSsafe (smallest Bonferroni-corrected $p = .092$).

Discussion

Using a sensory preconditioning procedure, the main goal of the current study was to examine whether extinguishing the CS-US association would effectively reduce costly avoidance of learnt fear. Main findings showed i) replication of costly avoidance of learnt fear acquisition, ii) a reduction of costly avoidance of learnt fear following CS extinction training, and iii) a link between the degree of avoidance of learnt fear and threat expectancies following avoidance decisions.

The current study found the acquisition of costly avoidance of learnt fear, replicating studies examining avoidance of learnt fear via higher-order conditioning procedures (Klein et al., 2021; Wong & Pittig, 2022b) and generally the acquisition of avoidance of learnt fear (e.g., Mennella et al., 2022; Pittig et al., 2014; Rinck et al., 2016). This was evident in two patterns. First, during *Post-extinction test*, avoidance responses to PSthreat was stronger than to PSext. Second, compared to *Baseline costly avoidance*, avoidance responses to PSthreat in *Post-extinction test* increased more than to PSext. These patterns suggest that avoidance of learnt fear was selectively acquired to a stimulus that predicted a non-extinguished threat-related CS (CSthreat). However, participants exhibited prior differences in baseline avoidance of learnt fear between the PSs. Specifically, baseline avoidance to PSthreat decreased more rapidly across trials compared to that to PSext. However, this difference in baseline avoidance did not necessarily confound the results, if anything, participants still exhibited stronger avoidance to PSthreat compared to PSext after CSthreat acquired threat value.

A key finding was that after CS extinction, participants showed lower avoidance to PSext. This highlights that extinguishing the threat value of an imminent threat signal reduces avoidance to a distal threat signal that predicts it. This pattern is consistent with the notion that fear and avoidance to a higher-order CS operates via a chain-like structure. Targeting parts temporally closer to threat affects responding to parts more distal to threat. The current study also provided support that avoidance of learnt fear is driven by threat prevention, thereby accounting for the reduction of avoidance to a distal signal that predicted an extinguished imminent threat signal. Specifically, an increase in avoidance to PSthreat that was linked to a decrease in post-avoidance US expectancy, whereas there was no significant association between avoidance and post-avoidance US expectancy to PSext. Therefore, this pattern suggests that extinguishing threat expectancy is effective in reducing avoidance of learnt fear.

One interesting finding was that despite avoidance to PSext was at a low level, it remained significantly stronger than to PSsafe. This apparently suggested a persistence in avoidance to PSext, even when conditioned fear to CSext, as indicated by SCRs, had been successfully extinguished during *Pavlovian extinction*. However, threat expectancy to CSext remained incomplete, as indicated by significant differential US expectancies to CSext and CSsafe on the last extinction trial. This incomplete extinction of threat expectancy to CSext might have contributed to the persistent avoidance to PSext. Indeed, a moderate level of post-avoidance US expectancies to PSext on early test trials despite low level of avoidance reflected residual US expectancies from incomplete extinction. Interestingly, the nearly zero correlation between avoidance to PSext and post-avoidance US expectancy ratings in the regression analyses suggested that factors other than threat expectancy may drive avoidance of learnt fear, which will be discussed below.

One factor that potentially drives avoidance of learnt fear is negative valence. Negative valence per se has been suggested to be sufficient to motivate avoidance (Chen & Bargh, 1999; Hans Phaf et al., 2014; Krieglmeyer et al., 2010), even when a threat is not expected anymore (Baeyens et al., 1995; De Houwer et al., 2001). The current study showed that when compared to baseline valence, valence to PSs that was or had been associated with a threat (PSext and PSthreat) decreased, consistent with studies that a higher-order CS that indirectly predicted a US acquired negative valence (Wong & Pittig, 2022a; Yu et al., 2014). However, we did not find any evidence that a negative change in valence was associated with an increase in avoidance of learnt fear in the current study (cf. Wong & Pittig, 2022b). One potential explanation for this null result is that we did not assess the change in valence immediately after Pavlovian fear acquisition; given that none of the stimuli were reinforced in *Post-extinction test*, negative valence to PSs associated with threat might have been reduced. Another factor that potentially reinforces avoidance of learnt fear is to prevent negative emotions evoked by the imminent threat signal (e.g., conditioned fear evoked by the CS+). However, the current findings provided limited evidence for this potential explanation given that avoidance of learnt fear to PSext persisted even after conditioned fear to CSext was extinguished as indexed by SCRs. Nonetheless, SCRs are characterized by its high variability (Lykken & Venables, 1971), in addition to not being a sensitive measure for fear (Boucsein et al., 2012), thus not necessarily exclude the possibility that avoidance of learnt fear is driven by the prevention of negative emotions evoked by the CS+. Future studies can include measurements that are arguably more sensitive to measure emotion distress such as self-reported fear or distress ratings, or startle eyeblink responses. Exploratory analyses revealed that trait anxiety and intolerance of uncertainty had different impacts on costly avoidance of learnt fear. Trait anxiety was associated with limited reduction of avoidance to PSext compared to PSsafe across test trials, suggesting that trait anxiety was linked to a persistence in avoidance of learnt fear to a PS that signaled an extinguished CS. This pattern expanded on findings that trait anxiety was associated with impaired reduction in safety behaviors to a stimulus that no longer signaled threat (e.g., Andreatta et al., 2017; Nishi et al., 2019; Xia et al., 2017). In contrast, there was no evidence that intolerance of uncertainty had any effect on avoidance of learnt fear (cf. Flores et al., 2018). However, it should be noted that preliminary evidence for the effect of risk factors on avoidance of learnt fear was mixed; while some preliminary evidence showed a positive link between risk factors and avoidance of learnt fear (Pittig et al., 2014), others found null risk factors effects on avoidance of learnt fear (Klein et al., 2021; Wong & Pittig, 2022b). Therefore, the effect of risk factors on avoidance of learnt fear required replication in well powered samples (Morris, Zuj, & Mertens, 2021; Ney et al., 2018; Wong et al., under review).

With regard to clinical implications, CS extinction reduces avoidance of learnt fear, providing empirical support that exposure-sessions can reduce pathological avoidance of learnt fear. Specifically, the current findings support the notion that expectancy violation (i.e., mismatch between an expected outcome and the actual absence of the outcome) plays a major role in the reduction in pathological avoidance (Craske et al., 2014). However, avoidance was still stronger to PSext than to PSsafe, suggesting the persistence of avoidance of learnt fear after CS extinction. This suggests that exposure-sessions alone may not be sufficient to completely extinguish avoidance of learnt fear. Future studies can examine whether combining other methods that optimize extinction learning, for example, presenting multiple fear stimuli in multiple contexts (e.g., Shiban, Schelhorn, Pauli, & Mühlberger, 2015; Waters, Kershaw, & Lipp, 2018), employing counterconditioning techniques (see Keller et al., 2020), or inducing positive affect during extinction (Meulders et al., 2014; Zbozinek et al., 2015). However, one caveat is that the apparent persistence in avoidance to PSext might have been due to incomplete threat expectancy extinction to CSext. Thereby, suggesting sufficient exposure sessions may be able to completely extinguish pathological avoidance of learnt fear. Interestingly, an increase in avoidance to a distal safety stimulus predicts an increase in threat expectancy to it. This preliminary evidence suggests that avoidance to distal safety stimuli may induce pathological threat beliefs (cf. Engelhard et al., 2015; van Dis et al., 2022; van Uijen et al., 2018), thus measures should be made to prevent clinically anxious individuals to engage in avoidance of learnt fear to distal safety stimuli.

The current study had some limitations. First, we did not assess valence ratings to all stimuli immediately after Pavlovian fear acquisition, thus negative valence to certain stimuli might have been attenuated at the end of the experiment. Second, despite conditioned fear to CSext, as indexed by SCRs, was completely extinguished on the last extinction trial, extinction of US expectancy to it was incomplete. We used relatively few extinction trials (c.f. Pittig & Wong, 2021; Vervliet & Indekeu, 2015) given that CSext was fully reinforced, thus we expected rapid extinction learning to it (see partial reinforcement extinction effect; Chan & Harris, 2019; Humphreys, 1939). Future studies can use more extinction trials to ensure a complete extinction in responding to the CS+. Third, there seemed to have pre-existing biases to baseline valence to the PSs and CSs, as the PSs were generally rated more positively than the CSs before the conditioning task. This bias in baseline valence might have confound with the change in valence to the stimuli after the conditioning task.

In conclusion, using a sensory preconditioning procedure, the current study found that CS extinction reduced costly avoidance of learnt fear. Threat expectancy was found to be one major factor that motivates avoidance of learnt fear, thus extinguishing threat expectancy to the CS reduces avoidance to a distal threat signal. Trait anxiety was found to be associated with persistent avoidance of learnt fear to a distal threat signal that predicted an extinguished CS. Future studies can examine how effectively CS extinction prevents return of avoidance of learnt fear in multiple-day paradigms. Additionally, future studies can examine whether factors other than threat prevention, such as relief or safety signals generated by behavioral avoidance (Papalini et al., 2021; Papalini et al., 2022; Vervliet et al., 2017; see also Wong et al., 2022), play a role in interfering the reduction in avoidance of learnt fear.

Acknowledgments

AHKW was supported in part by funds of the Bavarian State Ministry of Science and the Arts and the University of Würzburg to the Graduate School of Life Sciences (GSLS), University of Würzburg. The authors would like to thank Ulkar Huseynzade for her help in data collection and processing.

Author Contributions

Alex H. K. Wong: Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Software; Project administration; Resources; Supervision; Visualization; Writing—original draft, review and editing

Andre Pittig: Conceptualization; Methodology; Writing—Review and editing.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: AHKW was supported in part by funds of the Bavarian State Ministry of Science and the Arts and the University of Würzburg to the Graduate School of Life Sciences (GSLS), University of Würzburg. Open access funding was supported by the Open Access Publication Funds of the University of Göttingen.

Data Availability

The datasets generated and analysed during the current study are available in the Open Science Framework repository, https://osf.io/8fuqs/

ORCID iDs

Alex H. K. Wong https://orcid.org/0000-0003-2227-0231
Andre Pittig https://orcid.org/0000-0003-3787-9576

Supplemental Material

Supplemental material for this article is available online.

References

Andreatta, M., Michelmann, S., Pauli, P., & Hewig, J. (2017). Learning processes underlying avoidance of negative outcomes. *Psychophysiology*, *54*(4), 578–590. https://doi.org/10.1111/psyp.12822

Archer, T., & Sjödén, P-O. (1982). Higher-order conditioning and sensory preconditioning of a taste aversion with an exteroceptive CSi. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, *34 (Pt 1)*(1), 1–17. https://doi.org/10.1080/14640748208400886

Baeyens, F., Eelen, P., & Crombez, G. (1995). Pavlovian associations are forever: on classical conditioning and extinction. *Journal of Psychophysiology*, *9*(1), 127–141

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., roth, W. T., Dawson, M. E., & Filion, D. L., Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

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. https://doi.org/10.1016/j.janxdis.2006.03.014

Chambers, J. A., Power, K. G., & Durham, R. C. (2004). The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of generalized anxiety disorder. *Journal of Anxiety Disorders*, *18*(5), 587–607. https://doi.org/10.1016/j.janxdis.2003.09.001

Chan, C. K. J., & Harris, J. A. (2019). The partial reinforcement extinction effect: The proportion of trials reinforced during conditioning predicts the number of trials to extinction. *Journal of experimental psychology. Animal learning and cognition*, *45*(1), 43–58. https://doi.org/10.1037/xan0000190

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*(2), 215–224. https://doi.org/10.1177/0146167299025002007

Coppock, W. J. (1958). Pre-extinction in sensory preconditioning. *Journal of Experimental Psychology*, *55*(3), 213–219. https://doi.org/10.1037/h0046379

Corrigan, I., Samuelson, K. A. M., Fridlund, B., & Thomé, B. (2007). The meaning of posttraumatic stress-reactions following critical illness or injury and intensive care treatment. *Intensive & critical care nursing*, *23*(4), 206–215. https://doi.org/10.1016/j.iccn.2007.01.004

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Davey, G. C. L., & Arulampalam, T. (1982). Second-order 'fear' conditioning in humans: Persistence of CR2 following extinction of CR1. *Behaviour Research and Therapy*, *20*(4), 391–396. https://doi.org/10.1016/0005-7967(82)90099-7

Debiec, J., Doyere, V., Nader, K., & LeDoux, J. E. (2006). Directly reactivated, but not indirectly reactivated, memories undergo reconsolidation in the amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3428–3433. https://doi.org/10.1073/pnas.0507168103

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: a review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853–869. https://doi.org/10.1037/0033-2909.127.6.853

Dunsmoor, J. E., White, A. J., & LaBar, K. S. (2011). Conceptual similarity promotes generalization of higher order fear learning. *Learning & Memory*, *18*(3), 156–160. https://doi.org/10.1101/lm.2016411

Engelhard, I. M., van Uijen, S. L., van Seters, N., & Velu, N. (2015). The effects of safety behavior directed towards a safety cue on perceptions of threat. *Behavior Therapy*, *46*(5), 604–610. https://doi.org/10.1016/j.beth.2014.12.006

Fetzner, M. G., Horswill, S. C., Boelen, P. A., & Carleton, R. N. (2013). Intolerance of Uncertainty and PTSD Symptoms: Exploring the construct relationship in a community sample with a heterogeneous trauma history. *Cognitive Therapy and Research*, *37*(4), 725–734. https://doi.org/10.1007/s10608-013-9531-6

Flores, A., Lopez, F. J., Vervliet, B., & Cobos, P. L. (2018). Intolerance of uncertainty as a vulnerability factor for excessive and inflexible avoidance behavior. *Behaviour Research and Therapy*, *104*, 34–43. https://doi.org/10.1016/j.brat.2018.02.008

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why dopeople worry? *Personality and Individual Differences*, *17*(6), 791–802. https://doi.org/10.1016/0191-8869(94)90048-5

Gerlach, A. L., Andor, T., & Patzelt, J. (2008). Die bedeutung von unsicherheitsintoleranz für die generalisierte angststörung modellüberlegungen und entwicklung einer deutschen version der unsicherheitsintoleranz-skala. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *37*(3), 190–199. https://doi.org/10.1026/1616-3443.37.3.190

Gewirtz, J. C., & Davis, M. (2000). Using pavlovian higher-order conditioning paradigms to investigate the neuralsubstrates of emotional learning and memory. *Learning & Memory*, *7*(5), 257–266. https://doi.org/10.1101/lm.35200

Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance, and affect: a meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Frontiers in Psychology*, *5*, 378. https://doi.org/10.3389/fpsyg.2014.00378

Hayes, A. F., Glynn, C. J., & Huge, M. E. (2012). Cautions regarding the interpretation of regression coefficients and hypothesis tests in linear models with interactions. *Communication Methods and Measures*, *6*(1), 1–11. https://doi.org/10.1080/19312458.2012.651415

Holmes, N. M., Cai, S. Y., Lay, B. P. P., Watts, N. R., & Westbrook, R. F. (2014). Extinguished second-order conditioned fear responses are renewed but not reinstated. *Journal of experimental psychology. Animal learning and cognition*, *40*(4), 440–456. http://dx.doi.org/10.1037/xan0000036

Humphreys, L. G. (1939). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *Journal of Experimental Psychology*, *25*(2), 141–158. https://doi.org/10.1037/h0058138

Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Korten, A. E., & Rodgers, B. (2000). Predicting anxiety and depression from personality: Is there a synergistic effect of neuroticism and extraversion? *Journal of Abnormal Psychology*, *109*(1), 145–149. https://doi.org/10.1037/0021-843X.109.1.145

Katz, R. C. (1974). Single session recovery from a hemodialysis phobia: A case study. *Journal of Behavior Therapy and Experimental Psychiatry*, *5*(2), 205–206. https://doi.org/10.1016/0005-7916(74)90116-5

Keller, N. E., Hennings, A. C., & Dunsmoor, J. E. (2020). Behavioral and neural processes in counterconditioning: Past and future directions. *Behaviour Research and Therapy*, *125*(1), 103532. https://doi.org/10.1016/j.brat.2019.103532

Klein, Z., Berger, S., Vervliet, B., & Shechner, T. (2021). High avoidance despite low fear of a second-order conditional stimulus. *Behaviour Research and Therapy*, *136*, 103765. https://doi.org/10.1016/j.brat.2020.103765

Kleinknecht, R. A., & Lenz, J. (1989). Blood/injury fear, fainting and avoidance of medically-related situations: A family correspondence study. *Behaviour Research and Therapy*, *27*(5), 537–547. https://doi.org/10.1016/0005-7967(89)90088-0

Krieglmeyer, R., Deutsch, R., De Houwer, J., & De Raedt, R. (2010). Being moved: Valence activates approach-avoidance behavior independently of evaluation and approach- avoidance intentions. *Psychological Science*, *21*(4), 607–613. https://doi.org/10.1177/0956797610365131

Kruschke, J. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (second ed.). Academic Press.

Kruschke, J. K., & Liddell, T. M. (2018). The bayesian new statistics: hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, *25*(1), 178–206. https://doi.org/10.3758/s13423-016-1221-4

Krypotos, A-M., Effting, M., Kindt, M, & Beckers, T. (2015). Avoidance learning: A review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, *9*(1), 189. https://doi.org/10.3389/fnbeh.2015.00189

Kumle, L., Võ, M. L-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior research methods*, *53*(6), 2528–2543. https://doi.org/10.3758/s13428-021-01546-0

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (second ed.). Sydney Psychology Foundation.

Lykken, D.T., & Venables, P.H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, *8*(5), 656–672. http://.doi.org/10.1111/j.1469-8986.1971.tb00501.x

Mennella, R., Bavard, S., Mentec, I., & Grèzes, J. (2022). Spontaneous instrumental avoidance learning in social contexts. *Scientific Reports*, *12*(1), 17528. https://doi.org/10.1038/s41598-022-22334-6

Meulders, A., Meulders, M., & Vlaeyen, J. W. S. (2014). Positive affect protects against deficient safety learning during extinction of fear of movement-related pain in healthy individuals scoring relatively high on trait anxiety. *The Journal of Pain*, *15*(6), 632–644. https://doi.org/10.1016/j.jpain.2014.02.009

Morriss, J., Zuj, D. V., & Mertens, G. (2021). The role of intolerance of uncertainty in classical threat conditioning: Recent developments and directions for future research. *International journal of psychophysiology: official journal of the International Organization of Psychophysiology*, *166*(1), 116–126. https://doi.org/10.1016/j.ijpsycho.2021.05.011

Ney, L., Wade, M., Reynolds, A., Zuj, D., Dymond, S., Matthews, A., & Felmingham, K. (2018). Critical evaluation of current data analysis strategies for psychophysiological measures of fear conditioning and extinction in humans. *International journal of psychophysiology: official journal of the International Organization of Psychophysiology*, *134*(1), 95–107. https://doi.org/10.1016/j.ijpsycho.2018.10.010

Nilges, P., & Essau, C. (2015). Die depressions-angst-stress-skalen: Der DASS– einscreeningverfahren nicht nur fur schmerzpatienten [depression, anxiety and stress scales: DASS – a screening procedure not only for pain patients]. *Der Schmerz*, *29*(6), 649–657. https://doi.org/10.1007/s00482-015-0019-z

Nishi, Y., Osumi, M., Nobusako, S., Takeda, K., & Morioka, S. (2019). Avoidance behavioral difference in acquisition and extinction of pain-related fear. *Frontiers in Behavioral Neuroscience*, *13*(1), 236. https://doi.org/10.3389/fnbeh.2019.00236

Papalini, S., Ashoori, M., Zaman, J., Beckers, T., & Vervliet, B. (2021). The role of context in persistent avoidance and the predictive value of relief. *Behaviour Research and Therapy*, *138*(1), 103816. https://doi.org/10.1016/j.brat.2021.103816

Papalini, S., Neefs, L., Beckers, T., Oudenhove, L. V., & Vervliet, B. (2022). Overnight fasting affects avoidance learning and relief. *Nutritional Neuroscience*, Advanced Online Publication. 1–14. https://doi.org/10.1080/1028415X.2022.2103068

Pittig, A., Schulz, A. R., Craske, M. G., & Alpers, G. W. (2014). Acquisition of behavioral avoidance: Task-irrelevant conditioned stimuli trigger costly decisions. *Journal of Abnormal Psychology*, *123*(2), 314–329. https://doi.org/10.1037/a0036136

Pittig, A., & Wong, A. H. K. (2021). Incentive-based, instructed, and social observational extinction of avoidance: Fear opposite actions and their influence on fear extinction. *Behaviour Research and Therapy*, *137*, 103797. https://doi.org/10.1016/j.brat.2020.103797

Pittig, A., Wong, A. H. K., Glück, V. M., & Boschet, J. M. (2020). Avoidance and its bidirectional relationship with conditioned fear: Mechanisms, moderators, and clinical implications. *Behaviour Research and Therapy*, *126*(2), 103550. https://doi.org/10.1016/j.brat.2020.103550

Prewitt, E. P. (1967). Number of preconditioning trials in sensory preconditioning using cer training. *Journal of Comparative and Physiological Psychology*, *64*(2), 360–362. https://doi.org/10.1037/h0088033

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

Rinck, M., Koene, M., Telli, S., Moerman-van den Brink, W., Verhoeven, B., & Becker, E. S. (2016). The time course of location-avoidance learning in fear of spiders. *Cognition & Emotion*, *30*(3), 430–443. https://doi.org/10.1080/02699931.2015.1009873

Rizley, R. C., & Rescorla, R. A. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative and Physiological Psychology*, *81*(1), 1–11. https://doi.org/10.1037/h0033333

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika*, *6*(5), 309–316. https://doi.org/10.1007/BF02288586

Sawchuk, C. N., Lohr, J. M., Tolin, D. F., Lee, T. C., & Kleinknecht, R. A. (2000). Disgust sensitivity and contamination fears in spider and blood-injection-injury phobias. *Behaviour Research and Therapy*, *38*(8), 753–762. https://doi.org/10.1016/S0005-7967(99)00093-5

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M., & Dymond, S. (2016). The tipping point: value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage*, *136*(1), 94–105. https://doi.org/10.1016/j.neuroimage.2016.04.070

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. https://doi.org/10.1016/j.brat.2015.05.014

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology. Human learning and memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

van Dis, E. A. M., Krypotos, A-M., Zondervan-Zwijnenburg, M. A. J., Tinga, A. M., & Engelhard, I. M. (2022). Safety behaviors toward innocuous stimuli can maintain or increase threat beliefs. *Behaviour Research and Therapy*, *156*(2), 104142. https://doi.org/10.1016/j.brat.2022.104142

Vansteenwegen, D, Crombez, G, Baeyens, F, Hermans, D, & Eelen, P (2000). Pre-extinction of sensory preconditioned electrodermal activity. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*, *53*(4), 359–371. https://doi.org/10.1080/713932734

van Uijen, S. L., Leer, A., & Engelhard, I. M. (2018). Safety behavior after extinction triggers a return of threat expectancy. *Behavior Therapy*, *49*(3), 450–458. https://doi.org/10.1016/j.beth.2017.08.005

Vervliet, B., & Indekeu, E. (2015). Low-Cost avoidance behaviors are resistant to fear extinction in humans. *Frontiers in Behavioral Neuroscience*, *9*(2), 351. https://doi.org/10.3389/fnbeh.2015.00351

Vervliet, B., Lange, I., & Milad, M. R. (2017). Temporal dynamics of relief in avoidance conditioning and fear extinction: Experimental validation and clinical relevance. *Behaviour Research and Therapy*, *96*(1), 66–78. http://doi.org/10.1016/j.brat.2017.04.011

Wagner, A.R., & Brandon, S.E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S.B. Klein, & R.R. Mowrer (Eds), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 149–190). Lawrence Erlbaum Associates, Inc.

Walz, N., Mühlberger, A., & Pauli, P. (2016). A human open field test reveals thigmotaxis related to agoraphobic fear. *Biological Psychiatry*, *80*(5), 390–397. https://doi.org/10.1016/j.biopsych.2015.12.016

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. https://doi.org/10.1016/j.brat.2018.04.005

White, K., & Davey, G. C. (1989). Sensory preconditioning and UCS inflation in human "fear" conditioning. *Behaviour Research and Therapy*, *27*(2), 161–166. https://doi.org/10.1016/0005-7967(89)90074-0

Wong, A. H. K., Aslanidou, A., Malbec, M., Pittig, A., Wieser, M. J., & Andreatta, M. *A systematic review of the inter-individual differences in avoidance learning*. (under review) https://psyarxiv.com/2xbwd/

Wong, A. H. K., & Pittig, A. (2022a). A dimensional measure of safety behavior: A non- dichotomous assessment of costly avoidance in human fear conditioning. *Psychological Research*, *86*(1), 312–330. https://doi.org/10.1007/s00426-021-01490-w

Wong, A. H. K., & Pittig, A. (2022b). Avoiding a feared stimulus: Modelling costly avoidance of learnt fear in a sensory preconditioning paradigm. *Biological Psychology*, *168*(1), 108249. https://doi.org/10.1016/j.biopsycho.2021.108249

Wong, A. H. K., Wirth, F. M., & Pittig, A. (2022). Avoidance of learnt fear: Models, potential mechanisms, and future directions. *Behaviour Research and Therapy*, *151*(1), 104056. https://doi.org/10.1016/j.brat.2022.104056

Xia, W., Dymond, S., Lloyd, K., & Vervliet, B. (2017). Partial reinforcement of avoidance and resistance to extinction in humans. *Behaviour Research and Therapy*, *96*(2), 79–89. https://doi.org/10.1016/j.brat.2017.04.002

Yu, T., Lang, S., Birbaumer, N., & Kotchoubey, B. (2014). Neural correlates of sensory preconditioning: a preliminary fMRI investigation. *Human Brain Mapping*, *35*(4), 1297–1304. https://doi.org/10.1002/hbm.22253

Zbozinek, T. D., Holmes, E. A., & Craske, M. G. (2015). The effect of positive mood induction on reducing reinstatement fear:

Relevance for long term outcomes of exposure therapy. *Behaviour Research and Therapy*, *71*(2), 65–75. https://doi.org/10.1016/j.brat.2015.05.016

Zuj, D. V., Xia, W., Lloyd, K., Vervliet, B., & Dymond, S. (2020). Negative reinforcement rate and persistent avoidance following response-prevention extinction. *Behaviour Research and Therapy*, *133*(133), 103711. https://doi.org/10.1016/j.brat.2020.103711

Author Biographies

Alex H. K. Wong is an assistant professor at the Clinical Psychology section at Erasmus University Rotterdam. His research focuses on the different learning processes in anxiety-related psychopathology in humans. He is particularly interested in the bio-behavioral processes that underlie the acquisition, generalization, and extinction of fear and avoidance.

Andre Pittig is a Professor for Translational Psychotherapy at the Georg Elias Mueller Institute of Psychology at the Georg August Universität Göttingen, Germany. His research is in the intersection of experimental and clinical psychology and psychopathology and focuses on emotional learning and approach-avoidance conflicts, their bio-behavioral mechanisms, and their relevance for the development, maintenance, and exposure-based treatment of anxiety and related disorders.