

# Between Scylla and Charybdis: Fixed-ratio avoidance response effort and unavoidable shock extinction in humans

Simon Dymond<sup>a,b,\*</sup>, Weiike Xia<sup>a</sup>, Daniel V. Zuj<sup>c</sup>, Martyn Quigley<sup>a</sup>

<sup>a</sup> School of Psychology, Swansea University, Singleton Campus, Swansea SA2 8PP, United Kingdom

<sup>b</sup> Department of Psychology, Reykjavík University, Menntavegur 1, Nauthólsvík, Reykjavík 101, Iceland

<sup>c</sup> School of Psychological Sciences, University of Tasmania, Locked Bag 1340, Launceston, TAS 7250, Australia

## ARTICLE INFO

### Keywords:

Avoidance  
Extinction  
Noneliminable  
Shock  
Punishment

## ABSTRACT

Avoidance of potential threat may become maladaptive when it is indiscriminate and resistant to change. Here, we investigated the resistance to change of high and low avoidance response effort when avoidance extinction involved unavoidable presentations of the aversive event (shock) in humans. Following fear conditioning, participants prevented upcoming shock delivery by responding on high (i.e., fixed ratio, FR-20) and low (FR-5) negative reinforcement schedules. Next, noneliminable shock was used for an avoidance extinction procedure whereby responding was followed by, rather than prevented, shock. During a subsequent standard extinction and response prevention test phase, we found that High effort (FR-20) avoidance would be more readily extinguished than Low effort (FR-5) avoidance. It was also predicted that fear, threat expectancy, and psychophysiological (skin conductance) responses would decrease on avoidable trials and increase on unavoidable trials before extinguishing to low levels. It was found that in the final extinction re-test phase when avoidance was possible, responding increased, particularly for low effort cues. Both fear and expectancy remained high. Individual differences on clinically relevant measures of trait anxiety, intolerance of uncertainty and experiential avoidance were associated with greater levels of fear and threat expectancy. Overall, unavoidable shock extinction may hold promise for further translational investigations of avoidance learning, extinction, and clinical treatment development.

## 1. Introduction

Avoidance of actual threat is adaptive, but a tendency to avoid perceived threatening situations can readily become maladaptive and lead to diagnosis of an anxiety-related disorder [1]. The effort or amount of behaviour expended to counteract perceived threat makes it more or less likely that maladaptive avoidance becomes a default way of coping. Unfettered, it is possible that the effectiveness of treatment applied to any maladaptive avoidance is likely to be impacted by one's prior history with response effort in controlling potential threat. Here, we report the findings of a laboratory-based treatment experiment designed to investigate the impact of high and low response effort on an understudied form of avoidance unlearning or extinction in humans.

Avoidance learning in humans is commonly studied by a combination of Pavlovian fear conditioning and operant learning paradigms [2–7]. In Pavlovian human fear conditioning, an initially neutral stimulus (conditioned stimulus or CS+) such as a tone or an abstract figure, is

paired with an aversive stimulus (the unconditioned stimulus or US) such as electric shock. Another neutral stimulus, the CS-, may also be paired with the absence of the US. As a result, presentations of the CS+ alone, but not the CS-, become sufficient to elicit conditioned fear responses (CR) typically assessed using physiological arousal (e.g., skin conductance), ratings (e.g., US expectancy), or response tendencies (e.g., freezing). To study operant avoidance conditioning requires the addition of a simple response requirement such that responses (e.g., pressing a key on a keyboard) during CS+ presentations now prevent the upcoming US. Responding during CS- presentations is unnecessary since shock never followed that cue. As a result, the probability of what is often termed *US-avoidance* responses made in the presence of the CS+ increases [8]. Once acquired, avoidance learning may be subject to extinction by withholding the aversive event; avoidance responses are now expected to decrease since they are both ineffective and unnecessary, yet threat beliefs may remain intact. This is assumed to be the result of continued responding under such extinction procedures not

\* Corresponding author at: School of Psychology, Swansea University, Singleton Campus, Swansea SA2 8PP, United Kingdom.

E-mail address: [s.o.dymond@swansea.ac.uk](mailto:s.o.dymond@swansea.ac.uk) (S. Dymond).

<https://doi.org/10.1016/j.bbr.2024.115299>

Received 17 June 2024; Received in revised form 25 September 2024; Accepted 17 October 2024

Available online 20 October 2024

0166-4328/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

allowing for opportunities to disconfirm the hypothesis that the US is in fact withheld. Extinction procedures like this, however, rarely if ever produce extended reductions in avoidance, tend to be studied in discrete trial blocks, to at least partially rely on task instructions, and tend to focus on variants of response prevention in tandem with extinction [9].

Research conducted using these paradigms described above has tended to focus on acquisition and extinction of low-cost avoidance (i.e., that is where avoidance evolves minimal costs or response effort). Clinically relevant investigations where avoidance requires costs or effort such as varying probabilities of the presentation of an aversive stimulus or acquisition of an appetitive reward, may lead to resistance to extinction and the maintenance of threat beliefs [10–17,8]. Here, we sought to further investigate an operant manipulation of avoidance response effort to discern its impact on the extinction of avoidance. Dymond et al. [18] examined the impact of high and low response effort operationalised as responding on reinforcement schedules on the extinction and return of avoidance. In Experiment 1, participants first completed a *threat conditioning phase* where they were presented with shapes that signalled the delivery (CS+) or absence (CS-) of a 250 ms electric shock. An *avoidance conditioning phase* then followed where participants could avoid the delivery of the shock by pressing a spacebar several times. During this phase, completion of high (FR-20) and low (FR-5) fixed-ratio (FR) reinforcement schedules were used to differentiate between *high* and *low* avoidance response efforts (i.e., pressing of the spacebar). Participants were able to judge if they had made enough avoidance responses, as each press of the spacebar resulted in an incremental increase of an avoidance bar that was presented on each trial. Next, an *extinction phase* began where the shock was never delivered but avoidance responses could be made (and the avoidance bar was present), and an *“extinction-feedback removal” phase* where the avoidance bar was not presented. A final re-extinction phase then followed where the avoidance bar was visible once more. Proportion of avoidance responses and threat expectancy and fear ratings were collected throughout the experiment. Critically, during the re-extinction phase there was no difference in participants' avoidance responses and threat expectancy and fear ratings based on response effort, however, ratings did differ as a function of CS type (CS+, CS-). In Experiment 2 similar findings were reported during a reinstatement testing phase.

Despite the acknowledged influence of Pavlovian and operant processes on avoidance, most of the extant research has tended to employ extinction procedures based largely on the former over the latter. For instance, withholding the US during extinction effectively mimics the purpose of avoidance conditioning; that is, to prevent upcoming US delivery. Persistent avoidance responding during extinction is hence considered maladaptive precisely because the US no longer occurs (unknownst to participants as they do not receive opportunities to disconfirm this contingency except when response prevention is added to the extinction procedure). Thus, avoidance is expected to extinguish when responding is deemed unnecessary to prevent the US. This Pavlovian arrangement assumes that US omission can both maintain avoidance conditioning and foster its unlearning (extinction; [4,19]).

One promising operant method with translational potential involves eliminating the response-aversive event contingency altogether [20,21,4,22]; that is, the aversive event (shock) occurs regardless of responding (i.e., it is unavoidable). The assumption is that following a history whereby avoidance reliably cancelled upcoming shock, the onset of extinction where that history of responding is now no longer effective at preventing shock should lead to a sustained reduction in avoidance. As a real-world example, unavoidable shock procedures may mimic processes involved in safety learning during exposure therapy. For instance, an agoraphobic client exposed to the outdoors is confronted with fearful consequences despite avoidance. Learning new ways of adapting to this changed environment necessitate non-avoidance-based methods since avoidance is now no longer effective. As a result, unavoidable shock extinction procedures may prompt adaptive forms of coping. Unavoidable shock is deemed to be an effective extinction procedure because “it

engenders maximal change between training and extinction conditions” ([22], p. 258). Findings demonstrate that presentations of unavoidable shock can extinguish free-operant avoidance responding in rats [23,24] and humans [25,26]. However, to date, little is known about the impact of unavoidable shock extinction on response-effort avoidance in humans [4].

The clinical relevance of translational work on avoidance and avoidance extinction should at least partially be reflected in individual differences in anxiety-relevant personality traits. Although research on individual differences in avoidance is relatively nascent [27,28], increased rates of avoidance are seen in non-clinical samples high in trait anxiety ([29]; see [14]), intolerance of uncertainty ([30]; see [31]), and neuroticism ([32]; see [33]). Here, we sought to extend this literature by administering assessments of trait anxiety and intolerance of uncertainty as well as experiential avoidance, which has been relatively understudied to date in the experimental psychopathology literature. Experiential avoidance is a concept originating from the acceptance and mindfulness literature and refers to an individual's attempts to reduce, prevent or minimise contact with negative or upsetting emotions and their willingness to commit to valued living [34,35]. To date, little is known about its impact on avoidance or avoidance extinction.

The aim of the present experiment was to test the impact of response-effort on avoidance extinction under an unavoidable shock paradigm in humans. It was expected that High effort (FR-20) avoidance would be more readily extinguished than Low effort (FR-5) avoidance both during an unavoidable shock presentations phase and in a subsequent extinction and response prevention test phase. We also predicted that fear, expectancy of shock, and psychophysiological (skin conductance) responses would decrease on avoidable trials and increase on unavoidable trials before extinguishing to low levels. Finally, we expected there would be associations between scores on clinically relevant questionnaires assessing trait anxiety, intolerance of uncertainty, and psychological flexibility on avoidance responding, fear, and expectancy ratings during a standard re-extinction phase.

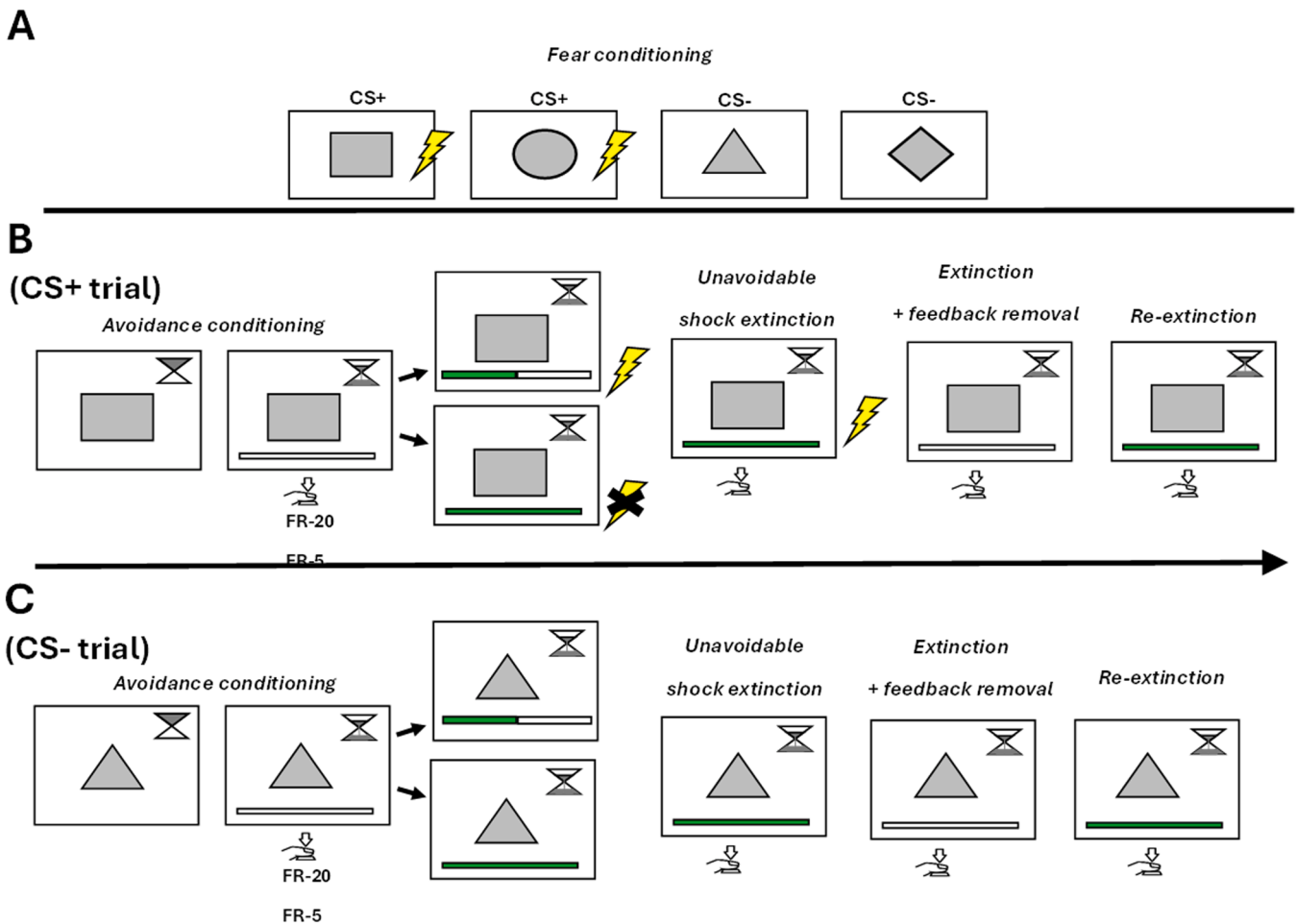
## 2. Materials and method

### 2.1. Participants

Based on our previous research [18,36] we recruited 42 participants (23 female,  $M_{age} = 21.76$ ,  $SD = 3.91$ , range 18–35 years). However, six participants failed to achieve a learning criterion of a minimum 80 % response rate (i.e., responding on 10 out of 12 trials) for at least one of the CS+s. One participant was excluded due to technical problems, resulting in  $n=35$  eligible for analysis. A sensitivity analysis [37] with  $\alpha = .05$ , power  $(1 - \beta) = 0.8$ , 1 group (i.e., a within-subjects design), and 2 measurements indicated that a sample size of 35 could detect an effect size equal to or greater than Cohen's  $f = 0.24$ . The study was approved by the School of Psychology Research Ethics Committee, Swansea University and all participants gave prior informed consent.

### 2.2. Apparatus and stimuli

Stimuli were presented on a 17" computer screen with a 60 Hz refresh rate and the task was programmed in *OpenSesame* [38]. The CSs were a grey coloured square, diamond, circle, and triangle shapes (counter-balanced across participants) presented in the middle of the screen. CSs were allocated by type (CS+ or CS-) and by response effort (high or low), resulting in 4 CSs: CS+ High, CS+ Low, CS- High, and CS- Low. The avoidance bar appeared beneath the CSs and consisted of a black rectangular box. Participants could press the spacebar which filled up the avoidance bar incrementally as indicated by a green progress bar. Each response was either 1/20th (i.e., FR-20) or 1/5th (i.e., FR-5) of the bar's total length representing high and low response effort, respectively. A black hourglass filled with grey sand was presented on the top right corner of the screen and was used to indicate trial duration (Fig. 1). For



**Fig. 1.** Overview of phases. (a) Demonstrated fear conditioning. Two CSs were followed by shock (CS+) and two were not (CS-). (b) Summarises the timeline of a CS+ trial (reading from left to right). A CS is first presented with an hourglass signalling the time until CS offset. During avoidance conditioning, pressing the spacebar filled the horizontal avoidance bar according to the relevant fixed-ratio schedules. During extinction responses filled the avoidance bar. All shocks were withheld in this phase. In the extinction and feedback prevention phase, responses did not fill up the avoidance bar and all shocks continued to be withheld. (c) Displays the timeline of a typical CS- trial. All details were the same as CS+ trials except the absence of shock.

all trials in all phases, a fixation cross was presented for 2 s, CSs for 10 s, and intertrial intervals (ITI) ranged from 4 to 8 s.

Every 12 trial-block (3 trials per CS), participants provided threat expectancy and fear ratings. Threat expectancy was measured using a visual analogue scale (VAS) shown underneath each CS, and given the instructions, “during each geometrical figure, a rating scale will also appear at the bottom of this screen. The scale runs from 0 to 100, where ‘0’ means “I expect certainly no shock”, ‘50’ means “I expect maybe a shock”, and ‘100’ means “I expect certainly a shock”. Fear ratings were measured using a VAS that ranged from 0 (“Not at all fearful”) to 100 (“Very fearful”). Participants were instructed, “after a few trials, you will also be asked to make fear ratings for each stimulus. Indicate your fearfulness of the stimulus by moving the mouse cursor along the slider bar and click the left mouse button to enter your rating.”

The US was a 250 ms electric shock generated using a STM200 stimulator (BIOPAC Systems) and administered through a surface electrode (MLADDF30 bar electrode with two 9 mm contacts spaced 30 mm apart). Electrode gel was applied to the right forearm and the electrode held in place with a band. Shock was individually calibrated at the beginning of the session. The current was initially set at 35 mV and increased or decreased in steps of 2.5 mV (the maximum was 100 mV). Participants were asked to report how uncomfortable they found the shock. When a level was deemed “uncomfortable but not painful” twice consecutively, it was used for that participant.

Skin conductance response amplitude was recorded with two

reusable Ag/AgCl electrodes filled with a non-hydrating gel and attached to the middle phalange of the first and second fingers. The electrodermal signal was sampled at 1000 Hz with a notch filter of 10 Hz and recorded in micro-siemens ( $\mu$ S) via the BIOPAC MP150.

### 2.3. Questionnaires

#### 2.3.1. State-trait anxiety inventory (STAI)

Trait anxiety was measured using the STAI [29], a 20-item questionnaire consisting of anxiety symptoms that are rated for frequency of occurrence using a 4-point scale ranging from 1 (“Almost never”) to 4 (“Almost always”). Reversed scoring is used for positive items and total scores range from 20 to 80 with higher scores indicating higher trait anxiety.

#### 2.3.2. Intolerance of uncertainty scale (IUS) – short form

The IUS-short form [30] has 12-items assessing intolerance of uncertainty on a 5-point scale ranging from 1 (“Not at all characteristic of me”) to 5 (“Entirely characteristic of me”).

#### 2.3.3. Brief experiential avoidance questionnaire (BEAQ)

The BEAQ is a 15-item scale assessing the tendency to reduce or prevent unpleasant emotions [34] and is scored on a 6 point scale ranging from 1 (strongly disagree) to 6 (strongly agree).

In the present study, internal consistency was good, with Cronbach’s

$\alpha = .93$  for STAI, .96 for IUS-short form, and .80 for BEAQ, respectively.

## 2.4. Procedure

First, participants completed the consent form and questionnaires on arrival and then underwent shock calibration and attachment of skin conductance electrodes. Participants received the following instructions:

*Welcome. In this experiment, you will be presented with geometrical shapes: Triangles, Squares, Circles and Diamonds. DIAMONDS and CIRCLES will be followed by an electrical stimulation, whilst SQUARES and TRIANGLES will not. There will also be a small hourglass in the upper right corner to let you know for how much longer the image will be shown. During the task you will have the opportunity to avoid the electrical stimulation by pressing the SPACEBAR several times. You will receive instructions and practice trials on this in a moment. First, you will see all the shapes and the hourglass twice to get used to them. During this phase, you will NOT get any electrical shock. Press SPACE to start.*

Participants then received the following avoidance practice instructions:

*Now you will practice how to avoid an upcoming electrical stimulation. During the experiment, you will see an empty bar appear underneath the figure. Each time you press the SPACEBAR, you will see that the bar will 'fill' up. Once the bar is filled, the response is complete and an upcoming shock is cancelled. During the practice, you will see ONLY the bar. Fill up the bar 4 times to complete the practice. Press SPACEBAR to start.*

The final set of instructions consisted of the following:  
*min. Press SPACEBAR to begin the experiment.*

The experiment consisted of 5 phases: *fear conditioning, avoidance conditioning, unavoidable shock extinction, extinction and feedback removal, and re-extinction test* (Fig. 1). All phases consisted of 12 trials (3 of each CS), except for *avoidance conditioning* which consisted of 48 trials or 4 blocks of 12 trials of each CS (see Table 1).

During the *fear conditioning* phase, stimuli were presented in a pseudorandom order such that no CS was presented twice consecutively. All CS+ trials (CS+ High and CS+ Low) were coupled, at stimulus offset, with the US at a 100 % reinforcement rate. CS- trials were never followed by shock.

In the *avoidance conditioning* phase, 5 s after CS onset the avoidance bar appeared underneath for the remainder of the trial and participants had 10 s to respond. With each press of the spacebar (i.e., each response as part of the FR response requirement), the avoidance bar would fill up incrementally: for 'High' trials, 20 spacebar presses were required to fill up the bar (conjunctive FI 10 FR-20), whereas for 'Low' trials, 5 spacebar presses were required (conjunctive FI 10 FR-5). During CS+ trials, filling up the bar cancelled the upcoming US on CS+ offset. If no responding occurred or the FR response requirements were not met, then the US still occurred on CS offset. During CS- trials, shock never followed regardless of spacebar presses. To ensure avoidance was acquired, a predetermined learning criterion was adopted whereby a minimum 80 % response rate (i.e., responding on 10 out of 12 trials) for at least one of the CS+s was required during *avoidance conditioning*.

During *unavoidable shock extinction*, participants could no longer avoid shock by filling up the avoidance bar. The avoidance bar appeared

5 s after CS presentation and filled up according to condition (high or low); however, the US always followed CS+ trials, regardless of responding, while CS- trials were never followed by the US. This phase consisted of 2 blocks of 12 trials (3 trials of each CS).

Next, during *extinction and feedback removal*, all shocks were withheld and pressing the spacebar no longer resulted in any visual change to the avoidance bar (i.e., response-contingent feedback was prevented).

Finally, during *re-extinction*, pressing the spacebar again filled the bar according to either FR 5 or FR 20 schedule requirements but shock was withheld on all trials.

After the experiment ended, participants were debriefed and compensated.

## 2.5. Data analysis

Skin conductance data was processed using *AcqKnowledge* (BIOPAC Systems, Santa Barbara, USA) and SCRs were calculated as the first peak (amplitude of the response) to occur within 0.5–5 s after CS onset. Prior to analysis, SCRs were range-corrected per participant to account for individual differences [39] and square root transformed across all phases to normalize the data [40]. Response criterion was set to 0.2  $\mu$ S and values less than it were scored as zero. Participants who had more than 90 % zero responses in all trials were excluded from analysis [41]. Seven participants were excluded from SCR analyses as they were identified as 'non-responders'. Furthermore, three more participants were excluded due to missing data resulting in SCR analysis of 25 participants' data. A sensitivity analysis [37] indicated that with  $\alpha = .05$ , power  $(1 - \beta) = 0.8$ , 1 group (i.e., a within-subjects design), and 2 measurements that a sample size of 25 could detect an effect size equal to or greater than Cohen's  $f = 0.29$ .

Fear and threat expectancy ratings were collected after each block of trials in each phase. Avoidance responses were recorded, and successful avoidance responses were scored '1' if participants had responded to criterion (i.e., 20 for high effort trials; 5 for low effort trials). Threat expectancy and fear ratings were analysed using 2 (CS)  $\times$  2 (response effort) RM-ANOVAs in the *fear conditioning* and *re-extinction* phases, a 2 (CS)  $\times$  2 (response effort)  $\times$  4 (block) RM-ANOVA in the *avoidance conditioning* phase, and 2 (CS)  $\times$  2 (response effort)  $\times$  3 (block) RM-ANOVAs in the *extinction* and *extinction and feedback removal* phases.

Follow-up analyses were Bonferroni-corrected. Where sphericity was not met, a Greenhouse-Geisser correction was used. Effect sizes were reported where results were significant. For both mixed ANOVAs and RM-ANOVAs, partial eta-squared was reported and for the independent t-test, Cohen's  $d$  was reported. Cohen's  $d$  is interpreted according to the criteria of 0.2, 0.5 and 0.8 as small, medium, and large effects, respectively [42]. Alpha was  $\alpha = 0.05$ . Analyses were performed in SPSS version 25 for Mac (IBM Corp., Armonk, New York, USA).

## 3. Results

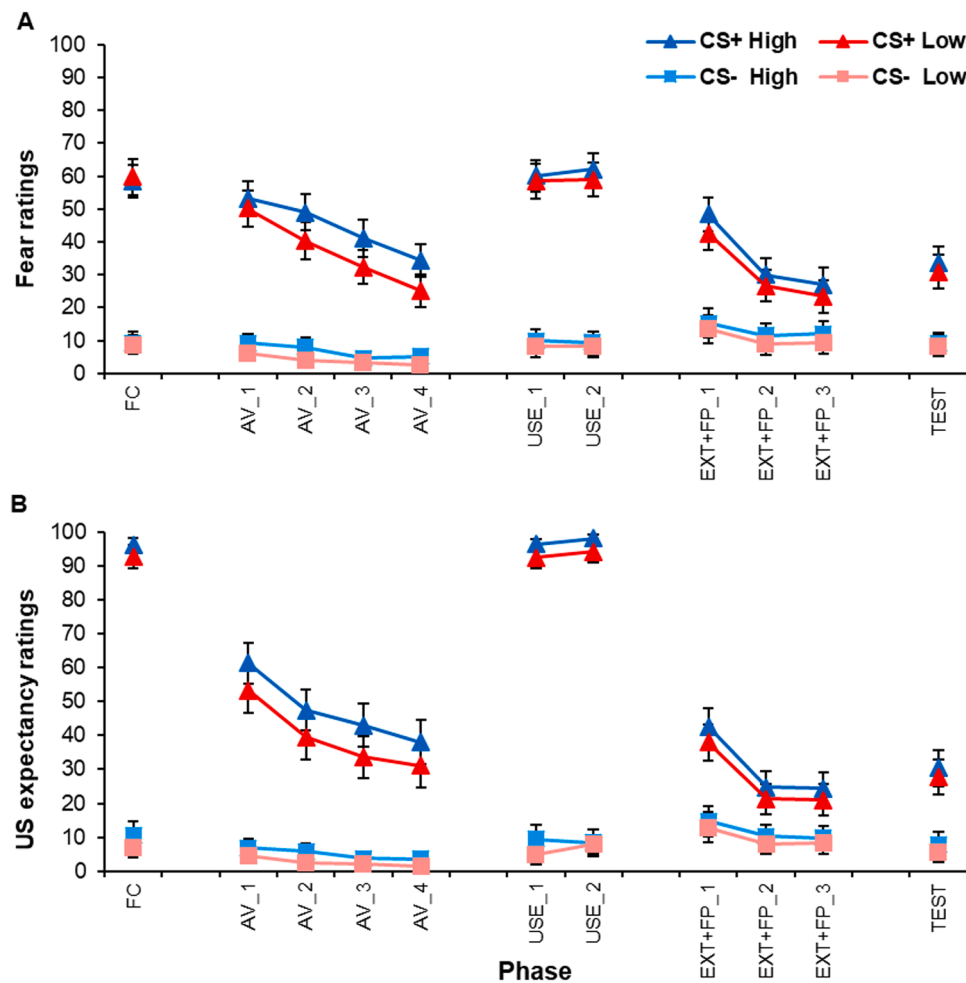
### 3.1. Fear conditioning

Differential Pavlovian fear conditioning was demonstrated (Fig. 2). Mixed ANOVAs showed significant main effects of CS in fear ratings ( $F(1,$

**Table 1**  
Experimental design.

|     |      | Phase             |                        |                              |                                 |               |
|-----|------|-------------------|------------------------|------------------------------|---------------------------------|---------------|
| CS  |      | Fear conditioning | Avoidance conditioning | Unavoidable shock extinction | Extinction and feedback removal | Re-extinction |
| CS+ | High | 3                 | 12                     | 6                            | 9                               | 3             |
|     | Low  | 3                 | 12                     | 6                            | 9                               | 3             |
| CS- | High | 3                 | 12                     | 6                            | 9                               | 3             |
|     | Low  | 3                 | 12                     | 6                            | 9                               | 3             |

*Note.* During the *avoidance*, *unavoidable shock extinction*, *extinction and feedback removal*, and *re-extinction* phases, trials were presented in blocks of 12 trials (3 trials per CS). High and low represent the number of spacebar presses required according to the fixed ratio schedule, where high refers to 20 presses and low refers to 5 presses.



**Fig. 2.** Behavioural ratings: (A) shows fear ratings and (B) shows threat/US expectancy ratings. FC = fear conditioning; AV = avoidance conditioning; USE = unavoidable shock extinction; EXT+FP = extinction and feedback prevention/removal; TEST refers to the re-extinction test phase. Numbers represent blocks per phase. Error bars represent SEM.

$_{34}) = 75.85, p < .001, \eta_p^2 = .69$ ), expectancy ratings ( $F_{(1, 34)} = 274.5, p < .001, \eta_p^2 = .89$ ) and SCR ( $F_{(1, 24)} = 7.31, p < .05, \eta_p^2 = .23$ ). As CS+ and CS- contingencies were explicitly stated, we did not check for awareness.

### 3.2. Avoidance conditioning

#### 3.2.1. Avoidance behaviour

Fig. 2A suggests a clear difference in avoidance behaviour between CS+ and CS- trials, and no effect of Load. The 2 (CS)  $\times$  2 (Load)  $\times$  12 (Trial) mixed ANOVA confirmed this as there was a significant main effect of CS ( $F_{(1, 34)} = 88.00, p < .001, \eta_p^2 = .72$ ), but not Load ( $F_{(1, 34)} = 4.10, p > .05$ ). There was also a significant interaction effect of CS  $\times$  Trial ( $F_{(3.45, 117.1)} = 4.27, p < .01, \eta_p^2 = .11$ ), but no three-way interaction ( $F_{(5.17, 175.79)} = 1.17, p > .05$ ) or CS  $\times$  Load interaction ( $F_{(1, 34)} = 3.69, p = .06$ ). To explore the CS  $\times$  Trial interaction, follow-up analyses were split by CS. The ensuing repeated ANOVAs showed a significant main effect of Trial for CS- ( $F_{(11, 374)} = 3.96, p < .001, \eta_p^2 = .10$ ), but not CS+ ( $F_{(11, 374)} = 0.96, p > .05$ ), indicating a general decrease in avoidance responses over time for CS- trials only.

#### 3.2.2. Fear ratings

Fig. 2A suggests a stronger decline in fear ratings for CS+ Low compared to CS+ High over time, and a general difference between CS+ and CS- fear ratings. The 2 (CS)  $\times$  2 (Load)  $\times$  4 (Block) mixed ANOVA failed to confirm this pattern ( $F_{(2.47, 83.84)} = 2.63, p = .07$ ), although a significant CS  $\times$  Load interaction was detected ( $F_{(1, 34)} = 6.59, p < .05, \eta_p^2 = .16$ ).

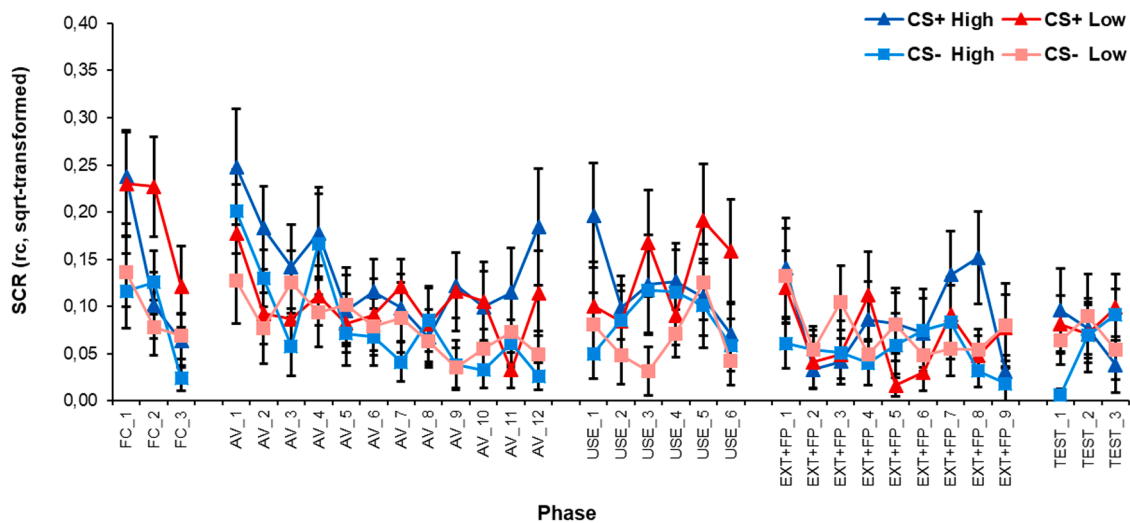
Follow-up analyses were split by CS and compared fear ratings between Load conditions per block. Bonferroni-corrected pairwise comparisons showed that from Block 2 onwards fear ratings to CS+ High were significantly higher than CS+ Low ( $p$ 's  $< .01$ ). There were no differences between CS- High and CS- Low ( $p$ 's  $> .05$ ).

#### 3.2.3. Expectancy

Fig. 2B shows a similar pattern for expectancy ratings, however the mixed ANOVA revealed only significant interaction effects of CS  $\times$  Load ( $F_{(1, 34)} = 10.03, p < .01, \eta_p^2 = .23$ ) and CS  $\times$  Block ( $F_{(1.43, 48.48)} = 7.92, p < .01, \eta_p^2 = .19$ ). Follow-up analyses showed that participants expected shock significantly more for CS+ High than CS+ Low ( $p < .05$ ), but equally for CS- High and CS- Low ( $p > .05$ ). Furthermore, there was a gradual decrease in shock expectancy ratings over time for CS+ trials ( $p$ 's  $< .05$ ). For CS- trials, however, there was no difference in shock expectancy ratings between blocks ( $p$ 's  $> .05$ ).

#### 3.2.4. SCR

Fig. 3 suggests a general decline of skin conductance responses throughout the avoidance conditioning phase. The 2 (CS)  $\times$  2 (Load)  $\times$  2 (Trial) mixed ANOVA revealed significant main effects of CS ( $F_{(1, 24)} = 12.79, p < .01, \eta_p^2 = .35$ ) and Trial ( $F_{(5.87, 140.8)} = 3.54, p < .01, \eta_p^2 = .13$ ), and a non-significant effect of Load ( $F_{(1, 24)} = 4.25, p = .05$ ). There were no significant interactions ( $F$ 's  $\leq 3.10, p$ 's  $> .05$ ). Together, this indicates that participants responded more strongly to CS+ than CS-, and SCR decreased in general over time.



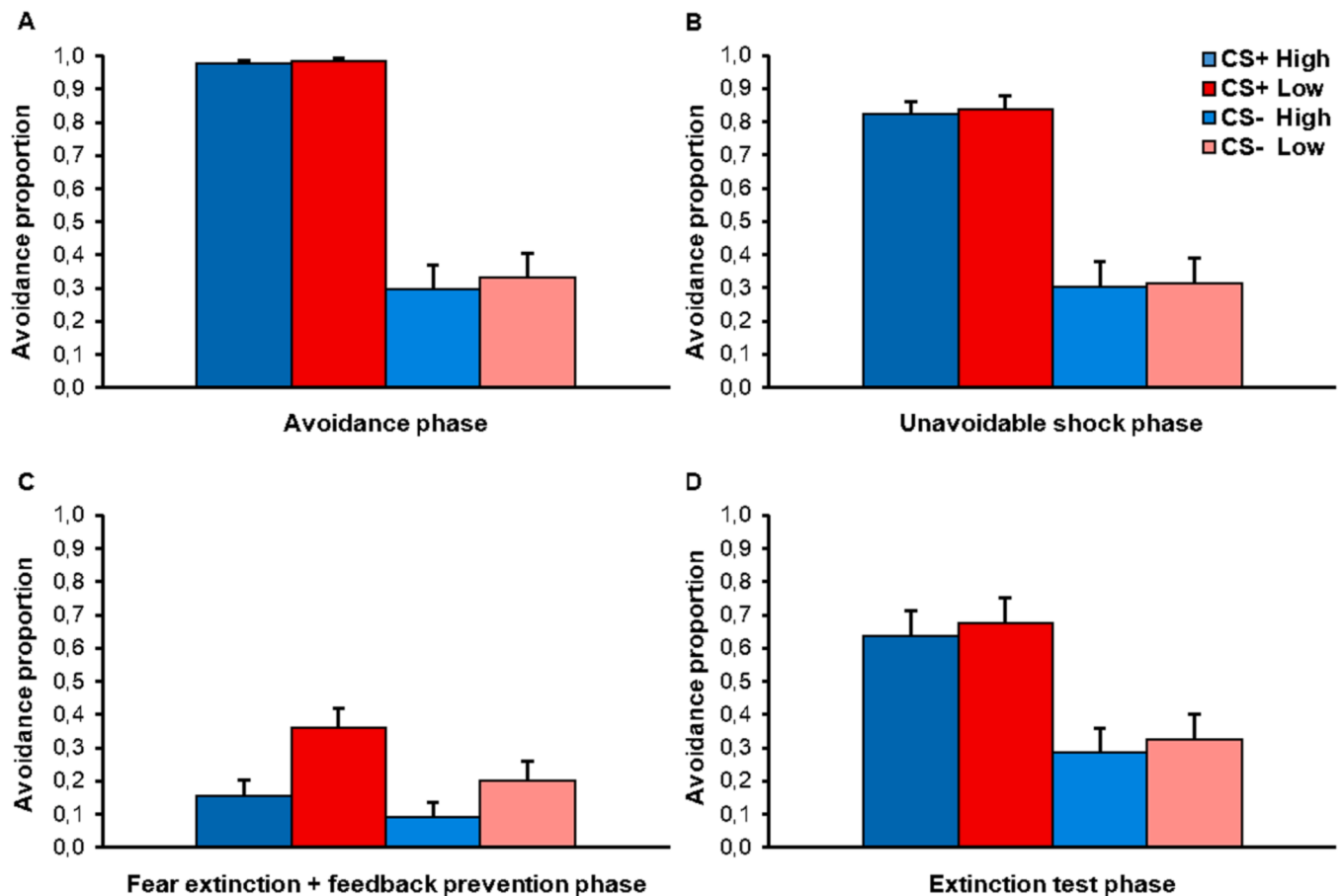
**Fig. 3.** Skin conductance responses - range-corrected and square root transformed. FC = fear conditioning; AV = avoidance conditioning; USE = unavoidable shock extinction; EXT+FP = extinction and feedback prevention/removal; TEST refers to the re-extinction test phase. Numbers represent blocks per phase. Error bars represent SEM.

### 3.3. Unavoidable shock extinction

#### 3.3.1. Avoidance behaviour

During *unavoidable shock extinction*, participants received shocks at the end of all CS+ High and CS+ Low presentations regardless of

avoidance. The 2 (CS) x 2 (Load) x 6 (Trial) mixed ANOVA showed only a significant CS x Trial interaction effect ( $F_{(3,54, 120.45)} = 7.11, p < .001, \eta_p^2 = .17$ ). There were no further interaction effects ( $F$ 's  $\leq 2.15, p$ 's  $> .05$ ). Follow-up analyses were split by CS. RM-ANOVAs showed that avoidance responses decreased linearly for CS+ trials ( $F_{\text{Trial}(5, 170)} = 11.64, p$



**Fig. 4.** Proportion of successful avoidances made per CS during (A) avoidance conditioning, (B) unavoidable shock extinction, (C) extinction and feedback removal, and (D) re-extinction test. Successful avoidance was achieved by pressing the spacebar either 5 times (low) or 20 times (high). During unavoidable shock extinction, participants received shock following CS+ high and CS+ low regardless of avoidance behaviour. Error bars represent SEM.

<.01,  $\eta_p^2 = .26$ ), but did not change for CS- trials ( $F_{\text{Trial}(5, 170)} = 1.23, p >.05$ ). This suggests that avoidance behaviour towards the CS+ decreased regardless of effort (Fig. 4B).

### 3.3.2. Fear ratings

Fig. 2A shows an increase in fear ratings following participants' exposure to unavoidable CS+ trials. A 2 (Phase) x 2 (CS) x 2 (Load) mixed ANOVA showed a significant three-way interaction ( $F_{(1, 34)} = 4.24, p <.05, \eta_p^2 = .11$ ). Follow-up pairwise comparisons showed that following unavoidable shock trials, fear ratings to both CS+ High and CS+ Low significantly increased ( $p$ 's <.001). Fear ratings for the CS- Low ( $p = .06$ ) and for the CS- High ( $p >.05$ ) did not differ. Fear ratings remained higher for the CS+ than CS- ( $F_{\text{CS}(1, 34)} = 92.12, p <.001, \eta_p^2 = .73$ ) but did not change over time ( $F_{\text{Block}(1, 34)} = 0.08, p >.05$ ) and nor were they effected by Load ( $F_{(1, 34)} = 0.93, p >.05$ ).

### 3.3.3. Expectancy

Fig. 2B suggests an immediate increase in expectancy ratings to both CS+ conditions and the CS- High following the *unavoidable shock extinction* phase. This was confirmed with a significant three-way interaction ( $F_{\text{Phase} \times \text{CS} \times \text{Load}(1, 34)} = 5.69, p <.05, \eta_p^2 = .14$ ) and follow-up comparisons showing significant increases in expectancy ratings to CS+ High and CS+ Low trials ( $p$ 's <.001) only. Similar to fear ratings, the 2 (CS) x 2 (Load) x 2 (Block) mixed ANOVA showed that expectancy ratings remained higher to CS+ trials than CS- trials ( $F_{\text{CS}(1, 34)} = 4563.7, p <.001, \eta_p^2 = .93$ ). There were no other effects ( $F$ 's  $\leq 1.34, p$ 's >.05).

### 3.3.4. SCR

A 2 (CS) x 2 (Load) x 6 (Trial) mixed ANOVA was conducted to research the effects of unavoidable shocks on skin conductance response. There was a significant main effect of CS ( $F_{(1, 24)} = 4.79, p <.05, \eta_p^2 = .16$ ) indicating a higher SCR to CS+ trials than CS- trials. There were no other effects ( $F$ 's  $\leq 1.61, p$ 's >.05).

## 3.4. Extinction and feedback removal

### 3.4.1. Avoidance behaviour

Avoidance presses were monitored throughout the *extinction and feedback removal* phase even as the avoidance bar no longer showed visual progress. Successful avoidances were scored if the participant had reached trial criterion. To assess the immediate impact of combined extinction and feedback removal, a 2 (Phase) x 2 (CS) x 2 (Load) RM-ANOVA was used comparing the last trial of *avoidance* to the first trial of *extinction and feedback removal*. Following a significant three-way interaction ( $F_{(1, 34)} = 7.03, p <.05, \eta_p^2 = .17$ ), Bonferroni-corrected pairwise comparisons revealed significant decreases in avoidance responses for the CS+ high ( $p <.001$ ) and CS- high ( $p <.05$ ) between *unavoidable shock extinction* and this phase. There were no differences, however, for the CS+ Low or CS- Low ( $p$ 's >.05). This indicates an immediate decrease in avoidance responses in the high effort condition only.

Fig. 4 suggests that these decreases lasted throughout *extinction and feedback removal* and this was confirmed with a significant three-way interaction ( $F_{\text{CS} \times \text{Load} \times \text{Trial}(5.55, 188.5)} = 2.70, p <.05, \eta_p^2 = .07$ ). To analyze this interaction, follow-up Load x Trial mixed ANOVAs were performed per CS. For the CS+, there was a significant Load x Trial interaction ( $F_{(5.23, 177.7)} = 5.37, p <.001, \eta_p^2 = .14$ ) and follow-up contrasts revealed this interaction to be linear ( $p <.05$ ) indicating that avoidance responses for the CS+ High decreased faster over time than for the CS+ Low. For the CS-, there was only a significant main effect of Load ( $F_{(1, 34)} = 7.64, p <.01, \eta_p^2 = .18$ ) indicating a higher response rate for the CS- Low in general.

Furthermore, follow-up pairwise comparisons on the last block of *extinction and feedback removal*, showed that there were no significant differences in avoidance responding between the CS+ High and both CS- ( $p$ 's <.05) indicating successful extinction. Avoidance responding to

the CS+ Low, however, remained higher than both CS- ( $p$ 's >.05).

### 3.4.2. Fear ratings

To investigate the effects of response feedback between phases, a 2 (Phase) x 2 (CS) x 2 (Load) mixed ANOVA revealed significant interaction of Phase x CS ( $F_{(1, 34)} = 16.70, p <.01, \eta_p^2 = .33$ ), but no effects of Load ( $F_{(1, 34)} = 1.95, p >.05$ ). Follow-up analyses showed that fear ratings to the CS+ were lower in the first block of *extinction + feedback removal* than during *unavoidable shock extinction* ( $p <.01$ ). Fear ratings to the CS- were only marginally higher ( $p = .05$ ).

Fig. 2A suggests a faster decline in fear ratings for the CS+ High than CS+ Low, while fear ratings to the CS- remain relatively equal. The significant interaction of CS x Block ( $F_{(1.33, 45.21)} = 11.97, p <.001, \eta_p^2 = .26$ ) indicates a faster decline of fear ratings for all CS+ trials compared to CS-. There were, however, no main effects or interaction effects of Load ( $F$ 's  $\leq 2.65, p$ 's >.05) suggesting that this decline was equal between high and low effort. Finally, follow-up analyses between CSs over time revealed that participants remained more fearful of the CS+ than the CS- throughout the phase ( $p$ 's >.05).

### 3.4.3. Expectancy

Fig. 2B suggests an immediate decrease in CS+ expectancy ratings and increase in CS- expectancy ratings following response prevention regardless of effort. The 2 (Phase) x 2 (CS) x 2 (Load) mixed ANOVA confirmed this through a significant interaction effect of Phase x CS ( $F_{(1, 34)} = 140.3, p <.001, \eta_p^2 = .81$ ) and significant follow-up pairwise comparisons ( $p$ 's <.05).

Similar to fear ratings, further analyses revealed a significant CS x Block interaction effect ( $F_{(1.39, 47.15)} = 9.21, p <.01, \eta_p^2 = .21$ ) indicating a faster decline of expectancy ratings towards CS+ trials. Follow-up analyses between CSs over time also revealed that participants remained more expectant of shock towards the CS+ than the CS- ( $p$ 's >.05).

### 3.4.4. SCR

Analyses were performed to investigate the immediate effect of response prevention between phases as well as throughout the *extinction and feedback removal* phase. All results were non-significant (all  $p$ 's >.05). Given the rapid changes in avoidance behaviour and behavioural ratings, this is likely due to habituation.

## 3.5. Re-extinction test

### 3.5.1. Avoidance behaviour

During *re-extinction test*, the avoidance bar would fill up once more if participants chose to press the spacebar. Comparing the first trial of *re-extinction* to the last trial of *extinction and feedback removal* revealed a significant Phase x CS interaction ( $F_{(1, 34)} = 16.14, p <.001, \eta_p^2 = .32$ ) and a significant main effect of Load ( $F_{(1, 34)} = 4.39, p <.05, \eta_p^2 = .11$ ). Follow-up pairwise comparisons showed that avoidance behavior increased significantly in the first trial of *re-extinction* for all CSs ( $p$ 's <.01). Furthermore, participants made significantly more avoidances to low effort trials.

Following this immediate increase, avoidance responses decreased throughout *extinction test* in differential rates based on CS ( $F_{(1.66, 56.46)} = 3.65, p <.05, \eta_p^2 = .10$ ), as follow-up analyses showed significantly less avoidance responses to the CS+ in the last trial of *re-extinction* for both CS+s ( $p$ 's <.05) compared to the first trial. There were no differences in avoidance responding to the CS- over time ( $p$ 's >.05). Finally, avoidance responding to the CS+ remained significantly higher than to the CS-.

### 3.5.2. Fear ratings

Fig. 2A suggests an increase in fear ratings to the CS+ and decrease to the CS-. Indeed, there was a significant interaction of Phase x CS ( $F_{(1, 34)} = 6.61, p <.05, \eta_p^2 = .16$ ) and follow-up analyses showed that participants were more fearful of the CS+ during *re-extinction* than *extinction*

and feedback removal ( $p < .01$ ) and less fearful of the CS- ( $p < .05$ ). Participants remained more fearful of the CS+ than CS- ( $F_{(1, 34)} = 27.48$ ,  $p < .001$ ,  $\eta_p^2 = .45$ ), regardless of effort ( $F_{(1, 34)} = 0.70$ ,  $p > .05$ ).

### 3.5.3. Expectancy

Participants expected shock more during CS+ trials during *re-extinction* than during *fear extinction and feedback removal*, and less during CS- trials ( $F_{(1, 34)} = 8.38$ ,  $p < .01$ ,  $\eta_p^2 = .30$ ; pairwise comparisons:  $p$ 's  $< .05$ ). Furthermore, shock expectancy remained higher for CS+ trials than CS- trials ( $F_{(1, 34)} = 24.27$ ,  $p < .001$ ,  $\eta_p^2 = .42$ ), regardless of effort ( $F_{(1, 34)} = 1.34$ ,  $p > .05$ ).

### 3.5.4. SCR

Like the preceding phase, all results were non-significant (all  $p$ 's  $> .05$ ).

## 3.6. Individual differences and pre-emptive avoidance behaviour

To investigate potential individual differences in the number of successful avoidance responses made per CS, questionnaire scores were correlated (using Spearman's rho) with the mean number of avoidance responses during *re-extinction* and behavioural ratings during *re-extinction*. To correct for the family-wise error, rate the  $\alpha$  was set to 0.017.

Table 2 shows that scores were not correlated with mean avoidance responses; however, each measure was correlated with expectancy and fear ratings. After corrections, STAI-T scores were positively correlated to shock expectancy to CS+ High and CS+ Low as well fearfulness towards the CS- High. IUS scores were positively correlated to US expectancy and fearfulness towards the CS+ High. BEAQ scores were positively correlated to fearfulness towards the CS+ High. Together, these results indicate that higher trait anxiety, intolerance of uncertainty and experiential avoidance are associated to greater levels of fear and more shock expectancy.

Finally, spacebar presses made during CS presentations before the avoidance bar appeared were recorded. To test whether this propensity to engage in pre-emptive avoidance behaviour was correlated to individual differences, mean pre-emptive avoidance behaviour per phase was correlated using Spearman's rho to questionnaire scores and an independent t-test for gender. To control the familywise error rate, the  $\alpha$  was corrected to 0.0025. There were no significant correlations (all  $p$ 's  $> .05$ ). Analyses were also conducted to compare females ( $N = 19$ ) to males ( $N = 16$ ) on pre-emptive avoidance. There were no differences in pre-emptive avoidance between these groups (all  $p$ 's  $> .05$ ).

**Table 2**

Correlations between questionnaire scores and outcomes measures during re-extinction.

|                              |      | STAI-T       | IUS          | BEAQ         |
|------------------------------|------|--------------|--------------|--------------|
|                              |      | $r$ ( $p$ )  | $r$ ( $p$ )  | $r$ ( $p$ )  |
| <b>Avoidance response</b>    |      |              |              |              |
| CS+                          | High | .00 (.99)    | .23 (.19)    | .21 (.22)    |
|                              | Low  | -.01 (.95)   | .22 (.21)    | .13 (.47)    |
| CS-                          | High | .13 (.47)    | .04 (.83)    | .04 (.82)    |
|                              | Low  | .13 (.46)    | .03 (.88)    | .06 (.72)    |
| <b>Fear ratings</b>          |      |              |              |              |
| CS+                          | High | .41 (.015)** | .44 (.009)** | .35 (.04)*   |
|                              | Low  | .42 (.012)** | .35 (.04)*   | .19 (.28)    |
| CS-                          | High | .02 (.08)    | -.01 (.95)   | .17 (.34)    |
|                              | Low  | .15 (.39)    | -.08 (.66)   | .11 (.53)    |
| <b>US Expectancy ratings</b> |      |              |              |              |
| CS+                          | High | .37 (.03)*   | .46 (.006)** | .41 (.013)** |
|                              | Low  | .28 (.10)    | .37 (.03)*   | .29 (.09)    |
| CS-                          | High | .42 (.012)** | .27 (.12)    | .39 (.02)*   |
|                              | Low  | .23 (.19)    | .08 (.66)    | .25 (.15)    |

Note. Spearman's rho coefficient is denoted by  $r$ .  $\alpha$  was corrected to 0.017. \* indicates  $p < .05$ ; \*\* indicates  $p < .017$

## 4. Discussion

The present study is, to our knowledge, the first to investigate unavoidable shock as an avoidance extinction procedure with humans. Following threat conditioning in a multi-cue procedure with several outcome measures, we observed reliable acquisition of avoidance learning across cues and effort types. Both fear ratings and expectancy remained higher for CS+ High compared to CS+ Low, while SCRs tracked these between-cue differences and gradually declined over trials. During the unavoidable shock extinction procedure, avoidance in the presence of both CS+s decreased regardless of effort, while fear, expectancy, and SCRs remained stable. During extinction with response prevention, there was an immediate decrease in avoidance in the presence of High effort cues (CS+ and CS-) only, and overall significant differences in avoidance between the CS+ High and both CS-s. Fear ratings declined regardless of effort as participants tracked the fact that although avoidance was not possible, no further shocks occurred. This trend was similarly evident for expectancy, although follow up analyses revealed that participants retained higher expectancy of shock for CS+ than CS-, suggesting some resistance to extinction, while SCR proved indistinguishable perhaps due to extinction. In the final extinction re-test phase when avoidance was possible, responding returned and increased for all cues but particularly for low effort cues. Both fear and expectancy remained high. Individual differences - higher trait anxiety, intolerance of uncertainty and experiential avoidance - were associated with greater levels of fear and threat expectancy.

Taken together, these results partially support our hypotheses. Whilst avoidance responses decreased to the high effort cues during extinction with response prevention and increased for low effort cues during the extinction re-test phase, there was no difference between avoidance to low and high effort cues during the unavoidable shock extinction phase. Similarly, our hypotheses in relation to our clinically relevant measures were also heterogenous. Although our measures of anxiety, intolerance of uncertainty and experiential avoidance were positively correlated with fear and expectancy during a standard re-extinction phase, these measures were not correlated with the number of avoidance responses participants made as we originally expected. The presents findings must therefore be considered preliminary and worthy of further replication and extension.

The onset of unavoidable shock extinction resulted in a very marginal but significant reduction in the proportion of avoidance responding to both CS+ High and CS+ Low. Marginally greater impact was observed for CS+ High, which indicates that a higher rate negative reinforcement schedule at acquisition may be resistant to change at test when the test conditions match the original learning conditions (i.e., a 100 % contingency between responding and presentation of shock). We have previously shown that higher rates of partially reinforced avoidance results in greater resistance to extinction when the aversive event is unknowingly withheld [36]. Our present findings partially extend this effect to high-rate avoidance learning and unavoidable shock extinction. Unavoidable shock extinction may have functioned as a punishment procedure, in that there was contingent delivery of the aversive event following previously effective avoidance responses and which lead to a reduction in responding [4,43-45]. Avoidance is negatively reinforced behaviour leading to an increase in responding and thus it may seem counter-intuitive to superimpose a positive punishment procedure (unavoidable shock) on an avoidance contingency. Approached in functional terms, the current procedures may be said to involve punishment if they are shown to lead to suppression of behaviour, which the onset of unavoidable shock extinction phase clearly did. Presenting response-contingent (i.e., unavoidable) shock following a period of avoidance learning does however have the potential to produce different functional effects. The absence of shock is negatively reinforcing (and likely serves to maintain ongoing avoidance; [46]), while on the other hand, the contingent presentation of shock is generally suppressive. Unlike previous research with nonhumans, the present findings

demonstrate a behaviour-reductive effect with humans akin to extinction where the functional reinforcers known to maintain the behaviour of interest (i.e., absence of shock) are clearly delineated (Azrin & Holz, 1966; Dinsmoor, 1998). The role of associative processes in the extinction of response effort avoidance such as the potential signalling properties of the avoidance bar warrant further empirical scrutiny [47,48]. Further research is warranted on the theoretical and translational implications of approaching the extinction of avoidance in this way. The present findings may have implications for treatment development in the domain of anxiety disorders. In 1970, Baum recommended combining procedures such as those of the present study with response prevention to extinguish avoidance. Our methods did not directly involve response prevention – instead, feedback following successive avoidance responses was removed and all shock withheld – but crucially this occurred *after* the unavoidable shock extinction phase. To some extent, then, while responding was not prevented it was unnecessary, and was likely suppressed by the preceding operation of punishment. Avoidance responding was at its lowest level during this phase, with responses made in the presence of CS+ High showing the greatest reduction, while responses made in the presence of CS+ Low were not as impacted. This is a notable finding and indicates that a low effort (FR5) behaviour is more resistant to extinction when combined with response prevention following a period of response-contingent punishment. The treatment implications of this finding are clear: high effort, effective avoidance is more likely to extinguish following unavoidable punishment. Given that many clinically maladaptive behaviours are likely to be low effort, effective avoidance responses then a clinician may wish to increase the response effort/rate of a problematic response prior to intervention with extinction and response prevention.

There are limitations to the current study. The sample size was relatively small which may have contributed to the heterogeneous findings observed across the different measures used during the different phases of the experiment. It is well established that small sample size can impact upon reproducibility [49]. Paradoxically, the potential translational relevance of the study is limited due to the unavoidable shock paradigm we employed. Outside of the laboratory, for instance, there are a limited range of real-world scenarios where individuals are placed in situations where they are unable to avoid aversive USs. The unavoidable shock extinction phase was relatively brief and may not have been sufficiently sensitive to detect whether extinction had occurred [4]. There are also a range of avoidance behaviours (both active and passive) that individuals are likely to make in the presence of an aversive US that were not captured in the current study.

In conclusion, we investigated unavoidable shock as an avoidance extinction procedure with humans and observed that High effort (FR-20) avoidance was more readily extinguished than Low effort (FR-5) avoidance. This procedure may hold promise for future translational investigations of avoidance learning.

### Author contributions

W.X.: Data collection, data analysis; S.D., D. Z., M.Q.: Writing – revision & editing;

### CRediT authorship contribution statement

**Simon Dymond:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Weike Xia:** Writing – original draft, Formal analysis, Data curation, Conceptualization. **Daniel V Zuj:** Writing – review & editing, Supervision, Conceptualization. **Martyn Quigley:** Writing – review & editing, Formal analysis, Conceptualization.

### Data Availability

Data will be made available on request.

### References

- [1] American Psychiatric Association, Diagnostic and statistical manual of mental disorders: DSM 5, American Psychiatric Publishing, 2013.
- [2] T.M. Ball, L.A. Gunaydin, Measuring maladaptive avoidance: from animal models to clinical anxiety, *Neuropsychopharmacology* 47 (5) (2022) 978–986, <https://doi.org/10.1038/s41386-021-01263-4>.
- [3] T. Beckers, D. Hermans, I. Lange, et al., Understanding clinical fear and anxiety through the lens of human fear conditioning, *Nat. Rev. Psychol.* 2 (2023) 233–245, <https://doi.org/10.1038/s41599-023-00156-1>.
- [4] S. Dymond, Overcoming avoidance in anxiety disorders: the contributions of Pavlovian and operant avoidance extinction methods, *Neurosci. Biobehav. Rev.* 98 (2019) 61–70, <https://doi.org/10.1016/j.neubiorev.2019.01.007>.
- [5] A.-M. Kryptos, B. Vervliet, I.M. Engelhard, The validity of human avoidance paradigms, *Behav. Res. Ther.* 111 (2018) 99–105, <https://doi.org/10.1016/j.brat.2018.10.011>.
- [6] J.E. LeDoux, J. Moscarello, R. Sears, V. Campese, The birth, death and resurrection of avoidance: a reconceptualization of a troubled paradigm, *Mol. Psychiatry* 22 (2016) 24–36, <https://doi.org/10.1038/mp.2016.166>.
- [7] D.V. Zuj, S.D. Norrholm, The clinical applications and practical relevance of human conditioning paradigms for posttraumatic stress disorder, *Prog. Neuropsychopharmacol. Biol. Psychiatry* 88 (2019) 339–351, <https://doi.org/10.1016/j.pnpbp.2018.08.014>.
- [8] A.H.K. Wong, F.M. Wirth, A. Pittig, Avoidance of learnt fear: models, potential mechanisms, and future directions, *Behav. Res. Ther.* 151 (2022) 104056, <https://doi.org/10.1016/j.brat.2022.104056>.
- [9] M.G. Wheaton, M. Gershkovich, T. Gallagher, E.B. Foa, H.B. Simpson, Behavioral avoidance predicts treatment outcome with exposure and response prevention for obsessive-compulsive disorder, *Depress Anxiety* 35 (3) (2018) 256–263, <https://doi.org/10.1002/da.22720>.
- [10] J.M. Bosch, S. Scherbaum, A. Pittig, Costly avoidance of Pavlovian fear stimuli and the temporal dynamics of its decision process, *Sci. Rep.* 12 (1) (2022) 6576, <https://doi.org/10.1038/s41598-022-09931-1>.
- [11] A.M. Hulsman, R. Kaldewaij, M.M. Hashemi, W. Zhang, S.B. Koch, B. Figner, F. Klumpp, Individual differences in costly fearful avoidance and the relation to psychophysiology, *Behav. Res. Ther.* 137 (2021) 103788, <https://doi.org/10.1016/j.brat.2020.103788>.
- [12] A. Lemmens, T. Smeets, T. Beckers, P. Dibbets, Avoiding at all costs? An exploration of avoidance costs in a novel Virtual Reality procedure, *Learn. Motiv.* 73 (2021) 101710, <https://doi.org/10.1016/j.lmot.2021.101710>.
- [13] J.A. Rattel, S.F. Miedl, J. Blechert, F.H. Wilhelm, Higher threat avoidance costs reduce avoidance behaviour which in turn promotes fear extinction in humans, *Behav. Res. Ther.* 96 (2017) 37–46, <https://doi.org/10.1016/j.brat.2016.12.010>.
- [14] B. Vervliet, E. Indekeu, Low-cost avoidance behaviors are resistant to fear extinction in humans, *Front. Behav. Neurosci.* 9 (2015) 351, <https://doi.org/10.3389/fnbeh.2015.00351>.
- [15] S. Wake, C.M. Van Reekum, H. Dodd, The effect of social anxiety on the acquisition and extinction of low-cost avoidance, *Behav. Res. Ther.* 146 (2021) 103967, <https://doi.org/10.1016/j.brat.2021.103967>.
- [16] A.H. Wong, J.C. Lee, P. Engelke, A. Pittig, Reduction of costly safety behaviors after extinction with a generalization stimulus is determined by individual differences in generalization rules, *Behav. Res. Ther.* 160 (2023) 104233, <https://doi.org/10.1016/j.brat.2022.104233>.
- [17] A.H. Wong, A. Pittig, A dimensional measure of safety behavior: a non-dichotomous assessment of costly avoidance in human fear conditioning, *Psychol. Res.* 86 (1) (2022) 312–330, <https://doi.org/10.1007/s00426-021-01490-w>.
- [18] S. Dymond, W. Xia, K. Lloyd, M.W. Schlund, D.V. Zuj, Working hard to avoid: fixed-ratio response effort and maladaptive avoidance in humans, *Q. J. Exp. Psychol.* 76 (8) (2023) 1889–1912, <https://doi.org/10.1177/17470218221127660>.
- [19] P.F. Lovibond, C.J. Mitchell, E. Minard, A. Brady, R.G. Menzies, Safety behaviours preserve threat beliefs: protection from extinction of human fear conditioning by an avoidance response, *Behav. Res. Ther.* 47 (8) (2009) 716–720, <https://doi.org/10.1037/a0029789>.
- [20] M. Baum, Extinction of avoidance responding through response prevention (flooding), *Psychol. Bull.* 74 (4) (1970) 276, <https://doi.org/10.1037/h0029789>.
- [21] L.S. Boldrin, P. Debert, S. Dymond, Avoidance extinction in equivalence classes, *Psychol. Rec.* (2024), <https://doi.org/10.1007/s40732-023-00580-y>.
- [22] S.T. Higgins, E.K. Morris, Generality of free-operant avoidance conditioning to human behavior, *Psychol. Bull.* 96 (2) (1984) 247, <https://doi.org/10.1037/0033-2909.96.2.247>.
- [23] D.G. Davenport, R.D. Olson, A reinterpretation of extinction in discriminated avoidance, *Psychon. Sci.* 13 (1968) 5–6, <https://doi.org/10.3758/BF03342380>.
- [24] D.G. Davenport, R.D. Olson, G.A. Olson, Preliminary analysis of partial reinforcement in discriminated avoidance, *Psychon. Sci.* 22 (1) (1971) 9–11, <https://doi.org/10.3758/BF03335913>.
- [25] W.A. Greene, L.T. Sutor, Stimulus control of skin resistance responses on an escape-avoidance schedule 1, *J. Exp. Anal. Behav.* 16 (2) (1971) 269–274, <https://doi.org/10.1901/jeab.1971.16.269>.
- [26] R.F. Hefferline, B. Keenan, R.A. Harford, Escape and avoidance conditioning in human subjects without their observation of the response, *Science* 130 (3385) (1959) 1338–1339, <https://doi.org/10.1126/science.130.3385.1338>.
- [27] Brunsch, M., Cooper, S.E., Abend, R., Boor, M., Chalkia, A., Ehlers, M.R., ... Lonsdorf, T.B. (2024, April 26). Associations between Anxiety-related Traits and Fear Acquisition and Extinction - an Item-based Content and Meta-analysis [Stage 1 Registered Report]. <https://doi.org/10.31234/osf.io/unx7w>.

- [28] T.B. Lonsdorf, C.J. Merz, More than just noise: inter-individual differences in fear acquisition, extinction and return of fear in humans - biological, experiential, temperamental factors, and methodological pitfalls, *Neurosci. Biobehav. Rev.* 80 (2017) 703–728, <https://doi.org/10.1016/j.neubiorev.2017.07.007>.
- [29] C.D. Spielberger, R.L. Gorsuch, R. Lushene, P.R. Vagg, G.A. Jacobs, *Manual for the State-Trait Anxiety Inventory (Form Y Self-Evaluation Questionnaire)*. Palo Alto, CA, Consulting Psychologists Press, 1983.
- [30] R.N. Carleton, M.P.J. Norton, G.J. Asmundson, Fearing the unknown: a short version of the Intolerance of Uncertainty Scale, *J. Anxiety Disord.* 21 (1) (2007) 105–117, <https://doi.org/10.1016/j.janxdis.2006.03.014>.
- [31] A. Flores, F.J. López, B. Vervliet, P.L. Cobos, Intolerance of uncertainty as a vulnerability factor for excessive and inflexible avoidance behavior, *Behav. Res. Ther.* 104 (2018) 34–43, <https://doi.org/10.1016/j.brat.2018.02.008>.
- [32] Eysenck, H.J., & Eysenck, S.B.G. (1975). *Manual of the Eysenck personality questionnaire*. Educational and Industrial Testing Service, San Diego, CA.
- [33] M.J. Lommen, I.M. Engelhard, M.A. van den Hout, Neuroticism and avoidance of ambiguous stimuli: better safe than sorry? *Personal. Individ. Differ.* 49 (8) (2010) 1001–1006, <https://doi.org/10.1016/j.paid.2010.08.012>.
- [34] W. Gámez, M. Chmielewski, R. Kotov, C. Ruggero, N. Suzuki, D. Watson, The brief experiential avoidance questionnaire: development and initial validation, *Psychol. Assess.* 26 (1) (2014) 35, <https://doi.org/10.1037/a0034473>.
- [35] T.B. Kashdan, J. Rottenberg, Psychological flexibility as a fundamental aspect of health, *Clin. Psychol. Rev.* 30 (7) (2010) 865–878, <https://doi.org/10.1016/j.cpr.2010.03.001>.
- [36] D.V. Zuj, W. Xia, K. Lloyd, B. Vervliet, S. Dymond, Negative reinforcement rate and persistence of avoidance following response-prevention extinction, *Behav. Res. Ther.* 133 (2020) 1037111, <https://doi.org/10.1016/j.brat.2020.103711>.
- [37] F. Faul, E. Erdfelder, A.G. Lang, A. Buchner, G\* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, *Behav. Res. Methods* 39 (2) (2007) 175–191, <https://doi.org/10.3758/BF03193146>.
- [38] S. Mathot, D. Schreij, J. Theeuwes, OpenSesame: an open-source, graphical experiment builder for the social sciences, *Behav. Res. Methods* 44 (2012) 314–324, <https://doi.org/10.3758/s13428-011-0168-7>.
- [39] D.T. Lykken, P.H. Venables, Direct measurement of skin conductance: a proposal for standardization, *Psychophysiology* 8 (5) (1971) 656–672, <https://doi.org/10.1111/j.1469-8986.1971.tb00501.x>.
- [40] M.E. Dawson, A.J. Rissling, A.M. Schell, R. Wilcox, Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm, *Emotion* 7 (4) (2007) 755–766, <https://doi.org/10.1037/1528-3542.7.4.755>.
- [41] M.F. Marin, F. Barbey, B.L. Rosenbaum, M.Z. Hammoud, S.P. Orr, M.R. Milad, Absence of conditioned responding in humans: a bad measure or individual differences? *Psychophysiology* 57 (2019) 1–10, <https://doi.org/10.1111/psyp.13350>.
- [42] J. Cohen. *Statistical Power Analysis for the Behavioural Sciences*, 2nd ed., Lawrence Erlbaum, 1988.
- [43] P. Jean-Richard-dit-Bressel, K. Gaetani, L. Zeng, G. Weidemann, G.P. McNally, Translational Research in Punishment Learning. *Behavioral Neuroscience*, Advance online publication, 2024, <https://doi.org/10.1037/bne0000587>.
- [44] J.A. Dinsmoor, Escape, avoidance, punishment: where do we stand? *Journal of the Experimental Analysis of Behavior* 28 (1) (1977) 83–95, <https://doi.org/10.1901/jeab.1977.28-83>.
- [45] P.N. Hineline, J. Rosales-Ruiz, Behavior in relation to aversive events: Punishment and negative reinforcement, in: G.J. Madden, W.V. Dube, T.D. Hackenberg, G. P. Hanley, K.A. Lattal (Eds.), *APA handbook of behavior analysis*, Methods and principles, 1, American Psychological Association, 2013, pp. 483–512, <https://doi.org/10.1037/13937-021>.
- [46] L. Leng, T. Beckers, B. Vervliet, What a relief! The pleasure of threat avoidance, *Emotion* 24 (3) (2024) 539–550, <https://doi.org/10.1037/emo0001312>.
- [47] P.A.F. Laing, B. Vervliet, J.E. Dunsmoor, et al., Pavlovian safety learning: an integrative theoretical review, *Psychon. Bull. Rev.* (2024), <https://doi.org/10.3758/s13423-024-02559-4>.
- [48] G.P. Urcelay, A psychological mechanism for the development of anxiety, *Behav. Neurosci.* 138 (4) (2024) 281–290, <https://doi.org/10.1037/bne0000607>.
- [49] L. Ney, P. Laing, T. Steward, D.V. Zuj, S. Dymond, B. Harrison, B. Graham, K. Felmingham, Methodological implications of sample size and extinction gradient on the robustness of fear extinction across different analytic strategies, *PLoS ONE* 17 (5) (2022) e0268814, <https://doi.org/10.1371/journal.pone.0268814>.