Original Article

QJEP

Quarterly Journal of Experimental Psychology
2023, Vol. 76(8) 1889–1912
© Experimental Psychology Society 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17470218221127660
qjep.sagepub.com
SAGE

# Working hard to avoid: Fixed-ratio response effort and maladaptive avoidance in humans

**Simon Dymond[1,2], Weike Xia[1], Keith Lloyd[3], Michael W Schlund[4] and Daniel V Zuj[1,5]**

## Abstract

Maladaptive avoidance of safe stimuli is a defining feature of anxiety and related disorders. Avoidance may involve physical effort or the completion of a fixed series of responses to prevent occurrence of, or cues associated with, the aversive event. Understanding the role of response effort in the acquisition and extinction of avoidance may facilitate the development of new clinical treatments for maladaptive avoidance. Despite this, little is known about the impact of response effort on extinction-resistant avoidance in humans. Here, we describe findings from two laboratory-based treatment studies designed to investigate the impact of high and low response effort on the extinction (Experiment 1) and return (Experiment 2) of avoidance. Response effort was operationalised as completion of fixed-ratio (FR) reinforcement schedules for both danger and safety cues in a multi-cue avoidance paradigm with behavioural, self-report, and physiology measures. Completion of the FR response requirements cancelled upcoming shock presentations following danger cues and had no impact on the consequences that followed safety cues. Both experiments found persistence of high response-effort avoidance across danger and safety cues and sustained (Experiment 1) and reinstated (Experiment 2) levels of fear and threat expectancy. Skin conductance responses evoked by all cues were similar across experiments. The present findings and paradigm have implications for translational research on maladaptive anxious coping and treatment development.

## Keywords

Avoidance; extinction; response prevention; fixed-ratio schedules; fear; anxiety

Received: 25 March 2022; revised: 30 August 2022; accepted: 2 September 2022

Avoidance of seemingly innocuous stimuli is a defining feature of anxiety-related disorders (American Psychiatric Association, 2013; Ball & Gunaydin, 2022). In maladaptive anxious coping, active avoidance of threatening stimuli is often less effortful than engaging in competing behaviours (Bennett et al., 2020; Wong & Pittig, 2021). It is relatively easy and requires minimal behavioural effort, e.g., for an individual with social anxiety to mute or turn off one's phone and hence avoid anxiety-provoking invitations to social events rather than engage with these opportunities. The range of effortful avoidance behaviour may also extend to a series of interrelated responses leading to the cancellation or postponement of an aversive event, such as the continuous responses described by ratio-based schedules of reinforcement (Catania, 1992; Ferster & Skinner, 1957; Shull & Lawrence, 1998). To illustrate, someone who is socially anxious may avoid future social invitations by first blocking or removing phone contacts, then uninstalling applications, unfollowing groups of friends on social media platforms, and adjusting notification settings, etc. These effortful sequences of avoidance behaviours require completion of a fixed series of responses to ensure the potentially catastrophic consequences of receiving a social invitation are in fact prevented. Given the contrasting

---

[1]School of Psychology, Swansea University, Swansea, UK
[2]Department of Psychology, Reykjavík University, Reykjavík, Iceland
[3]Swansea University Medical School, Swansea University, Swansea, UK
[4]Department of Psychology, Georgia State University, Atlanta, GA, USA
[5]School of Psychological Sciences, University of Tasmania, Launceston, TAS, Australia

**Corresponding author:**
Simon Dymond, School of Psychology, Swansea University, Singleton Campus, Swansea SA2 8PP, UK.
Email: s.o.dymond@swansea.ac.uk

quantitative or physical response effort involved in these instances of anxiety-related avoidance, it is possible that the effectiveness of treatment is likely to be enhanced by consideration of the prior role of effort in controlling potential threat. Here, we report the findings of two laboratory-based treatment experiments designed to investigate the impact of high and low response effort on the persistence of (extinction-resistant) avoidance in humans.

Avoidance learning is commonly studied in variants of Pavlovian threat (fear) conditioning paradigms (Ball & Gunaydin, 2022; Dymond, 2019; Krypotos et al., 2018; LeDoux et al., 2016; Zuj & Norrholm, 2019). After a neutral stimulus (i.e., conditional stimulus, CS+) becomes a reliable predictor of an aversive unconditional stimulus (US; e.g., electric shock) and another stimulus (i.e., CS−) comes to reliably predict its absence, US-avoidance is acquired by adding a discrete operant response made in the presence of the CS+ (e.g., a button press) to cancel the upcoming US. Avoidance responses made in the presence of the CS− may still occur but are deemed unnecessary since the US is withheld on all such trials regardless of responding (Krypotos et al., 2018). Response-prevention extinction procedures may be included whereby the avoidance response is no longer available or prevented from occurring and elegantly allow for the investigation of the persistence of avoidance, differential or otherwise (Dymond, 2019).

In US-avoidance studies with humans, avoidance involves performing a discrete response with minimal cost and effort such as clicking a computer mouse once or pressing a designated keyboard key which prevents occurrence of the US (Krypotos et al., 2018; Wong et al., 2022). While the translational validity of using dichotomous responses for understanding the phenomenology of anxiety-related avoidance continues to be debated, there is a growing need to incorporate continuous measures in laboratory-based treatment studies of avoidance (Wong & Pittig, 2021). Meulders et al. (2016) operationalised avoidance costs as the effort involved in moving a three degree-of-freedom, force-controlled robotic arm where greater response effort was associated with lower chances of receiving the US (van Vliet et al., 2019). Findings showed that pain-related response-effort avoidance was related to resistance to extinction, with self-reported fear and pain expectancy ratings also decreasing significantly. Flores et al. (2018) employed a discriminated free-operant paradigm in which participants could emit multiple responses on one of two marked keys in the presence of two long duration (20 s) CS+s. The aversive noise US was randomly delivered at least 8 s after CS+ onset and participants were instructed that only responses within 1 s of scheduled US delivery would prevent the upcoming US. It was found that response frequency on CS+ trials increased across blocks of avoidance learning and remained low on CS− trials. In this way, incorporating a gradient of avoidance responses permitted a more detailed examination of effects on the strength of subsequent test performance than that previously afforded by a study employing a dichotomous response. Moreover, the inclusion of multiple avoidance responses per trial approximates the costly effort involved in anxious coping whereby decisions to avoid are not only followed by the absence of threat but also of opportunities to learn new, adaptive skills (Wong & Pittig, 2020).

Several decades of basic research on response effort have shown that the physical effort or force required to perform an operant response generally produces reductive effects comparable to punishment (Ailing & Poling, 1995; Blough, 1966). That is, responding decreases as effort requirements increase and extinction proceeds more rapidly (Friman & Poling, 1995). Fixed-ratio (FR) schedules require completion of a predetermined, fixed number of responses before reinforcement is delivered and tend to produce steady rates of responding characterised by initial pauses and subsequent bouts of responding (Catania, 1992). Increased response effort is associated with shorter inter-response times (IRTs) on continuous and FR schedules of positive reinforcement (Armus, 1988), while increasing FR requirements tends to decrease the maximum force emitted on an isometric force transducer (Falligant et al., 2020; Pinkston & McBee, 2014) and influences betting size in simulated gambling (Gunnarsson et al., 2015).

The use of FR schedules to operationalise response effort therefore has many potential applied implications. Despite this, the relevance for an understanding of anxiety-related avoidance and treatment development remains underexplored. A notable exception is Augustson and Dougher (1997) who employed an FR schedule of combined CS-escape and US-avoidance following threat conditioning where 20 responses (FR-20) terminated the CS+ and prevented upcoming shock; less than 20 responses resulted in continued CS+ presentation and shock. Button-pressing avoidance learning in Augustson and Dougher's study occurred in the absence of any on-screen feedback to indicate successive responses. Yet, immediate differential acquisition was seen in the presence of the CS+ compared with the CS−, indicating discriminative FR-based avoidance. Responding also persisted across a brief period of extinction testing, where all shock deliveries were withheld, involving the directly learned (i.e., CS+) and symbolically generalised avoidance cues. Findings demonstrated the acquisition and symbolic generalisation of effortful, FR-based, avoidance in humans. However, Augustson and Dougher did not apply FR response requirements to non-avoidance safety cue trials (CS−), did not shape the avoidance response such as by using response-produced feedback to ensure FR schedule control, and their findings are limited to the small sample size employed ($N = 8$).

The objective of the present experiments was to investigate the effects of high and low response effort on the extinction and return/recovery of avoidance following response prevention. Response effort was operationalised as completion of conjunctive fixed-interval (FI) FR-20 (*High*) and FR-5 (*Low*) avoidance schedules. We adapted conjunctive FI-FR schedules where negative reinforcement (in this case, shock omission) only occurred after a fixed period had elapsed *and* a fixed number of responses have been made (Herrnstein & Morse, 1958; Katz & Barrett, 1979; Keenan & Leslie, 1984). Responding on these schedules tends to be characterised by an initial pause, followed by an abrupt increase in response rate and then a gradual lower rate that may or may not increase up to (positive) reinforcement delivery. We used a validated human threat conditioning paradigm (Xia et al., 2017; Zuj et al., 2020) and included fixed duration presentations of two danger (CS+) and two safety cues (CS−), each with high- and low-effort FR response requirements (CS+ High, CS+ Low, CS− High, and CS− Low), respectively. This permitted an assessment of the role of response effort in US-avoidance trials and during safety trials when avoidance was unnecessary and/or prevented. We measured US-avoidance as the proportion of trials in which high or low response requirements were met, threat expectancy, fear ratings, and skin conductance responses.

In Experiment 1, we expected the total proportion of avoided trials to be high (i.e., successfully prevented the US) on both *High* (FI 10 s FR-20) and *Low* (FI 10 s FR-5) response-effort CS+ trials, while avoidance was expected to be at low to zero levels during CS− trials. During *extinction*, where US deliveries were withheld, we predicted that low response-effort trials would maintain avoidance, threat expectancy, and fear to the Low-effort CS+, while there would be a reduction in all measures to the High-effort CS+. Lower avoidance was expected because removal of scheduled shock delivery following a high rate of responses effectively negates the need to expend further effort. If response effort determines at least partially the occurrence of maladaptive avoidance, then the absence of the aversive event should prompt a conservation of high-rate, effortful responding. We also predicted persistence of avoidance responding in extinction when on-screen, avoidance response-contingent feedback was removed. Phases where the presence and absence of feedback were included to enhance stimulus control over the intended discriminative, effortful avoidance. Threat expectancy and fear were predicted to reduce for the High-effort CS+ as experience tracked the changing learning and extinction contingencies. Finally, we conservatively predicted minimal differences in avoidance between High- and Low-effort CS+ trials during a return to extinction phase where the on-screen feedback was reinstated. That is, we did not expect reinstatement of the trained stimulus control over avoidance when this phase was re-presented, and we expected no differential impact across cues.

## Experiment 1

### Method

*Participants.* A total of 36 participants between 18 and 30 years of age ($M=20.2$, $SD=2.4$), 20 male and 16 female, were recruited from Swansea University. A sensitivity analysis was conducted using *G*Power 3.1* (Faul et al., 2007) with $\alpha=.05$, Power $(1-\beta)=0.8$, one group (i.e., a within-subjects design), two measurements, and a sample size of $N=36$, we should find an effect size equal to or greater than Cohen's $f=0.24$ for any effects, where such effects exist. The study was approved by the School of Psychology Research Ethics Committee at Swansea University and written consent was obtained at the outset. Participants were compensated with course credit on completion.

*Apparatus and stimuli.* On-screen stimuli consisted of four CSs, the avoidance bar, and an hourglass timer. Stimuli were presented on a 17″ computer screen with a 60 Hz refresh rate and the task was programmed in *OpenSesame* (Mathôt et al., 2012). The CSs were a grey-coloured square, diamond, circle, and triangle shapes (counterbalanced across participants). The width and height of the triangle and diamond was 240 pixels, the square was 200 pixels, and the circle had a radius of 100 pixels. CSs were presented in the middle of the screen and were allocated by type (CS+ or CS−) and by response effort (high or low), resulting in four CSs: CS+ High, CS+ Low, CS− High, and CS− Low. The avoidance bar appeared beneath the CSs and consisted of a black rectangular box with a height of 20 pixels and length of 600 pixels. 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 with a width of 60 pixels and height of 185 pixels was presented on the top-right corner of the screen and was used to indicate trial duration (Figure 1). Following a block of 12 trials (3 trials per CS), participants provided threat expectancy and fear ratings. Threat expectancy was measured via ratings made on a visual analogue scale (VAS) shown underneath each CS. Participants were instructed, "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 in a similar manner at the end of trial blocks 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."

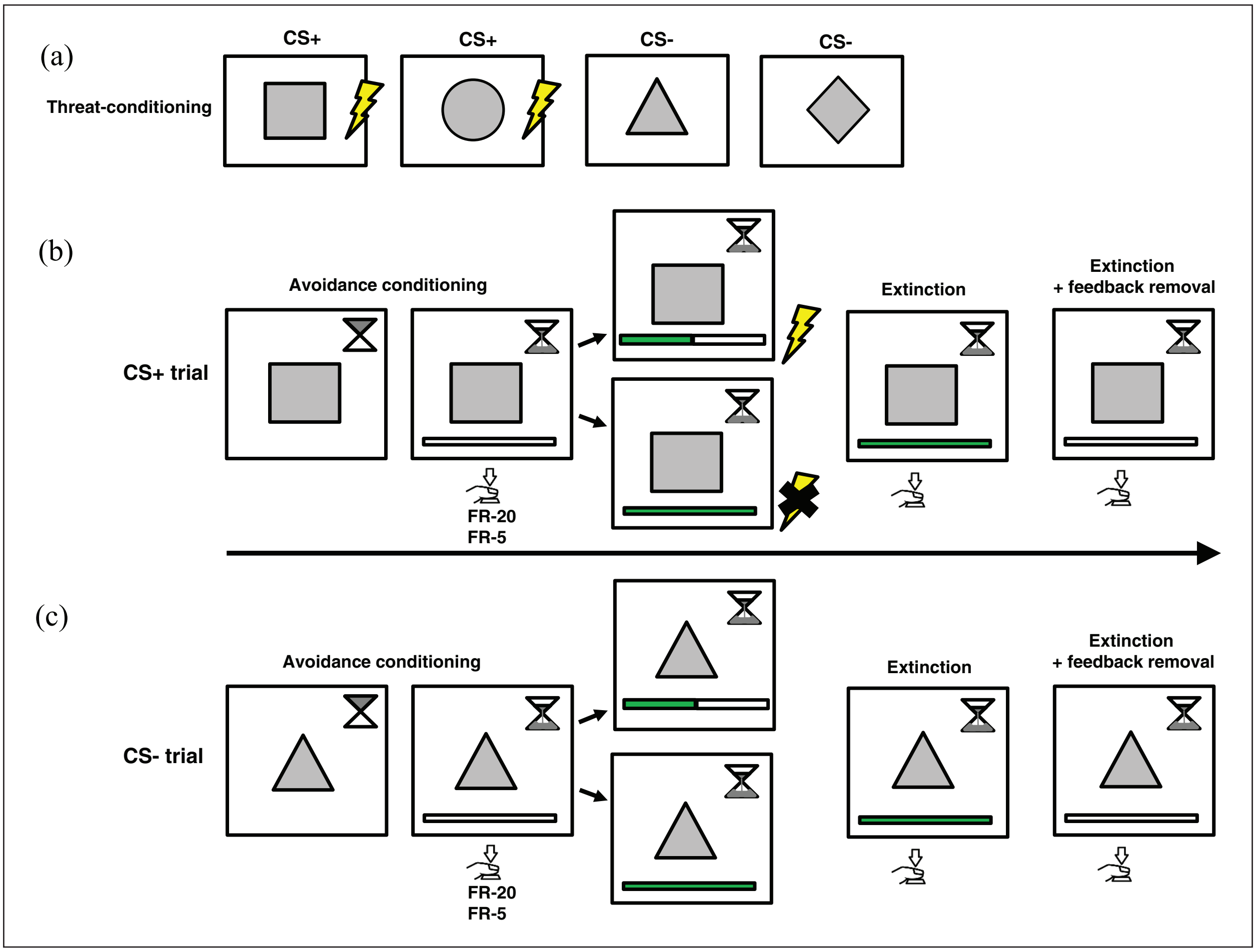

**Figure 1.** *Overview of phases and response contingencies in Experiment 1*. (a) Summarises the *threat conditioning* phase whereby two CSs were followed by shock (CS+) and two were not (CS−). (b) Illustrates the timeline of a typical CS+ trial (reading from left to right). First, a CS is presented along with the hourglass denoting the time until CS offset. Second, during *avoidance conditioning*, pressing the spacebar incrementally filled the horizontal avoidance bar according to one of the two fixed-ratio schedules. Partial responses or an incompletely filled avoidance bar was followed by shock; full responses or a filled avoidance bar cancelled shock. Third, during *extinction*, both partial and complete responses filled the avoidance bar. All shocks were withheld in this phase. Finally, in the *extinction and feedback prevention* phase, responses no longer filled up the avoidance bar and all shocks continued to be withheld. (c) Illustrates the timeline of a typical CS− trial. All features were identical to those of CS+ trials apart from the absence of shock on any trial in all phases. *Note*: FR-20 and FR-5 refer to the fixed-ratio schedules of US-avoidance conditioning, respectively.

The US was a 250 ms electric shock; intensity of the current was individually adjusted. The US was generated using a STM200 stimulator (BIOPAC Systems, Santa Barbara, CA, USA) 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 the intensity of the shock in terms of how uncomfortable they found it. When a shock 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 1,000 Hz with a notch filter of 10 Hz and recorded in micro-Siemens (μS) via the BIOPAC MP150 (BIOPAC Systems).

*Procedure.* Having signed the consent form, participants had shock electrodes applied and undertook shock calibration. Next, instructions were presented on screen regarding the shock contingencies and response requirements:

> *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. [instructions were counterbalanced across participants]*
>
> *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 SPACE BAR 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 SPACEBAR to start.*

Participants were then shown all four geometrical shapes with the hourglass in the top-right corner twice in random order for habituation purposes. Following *habituation*, participants were given four practice trials to practice filling up the bar by pressing the spacebar 5, 10, 15, or 20 times. Instructions given prior to the practice trials were as follows:

> *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 experiment consisted of five phases: *threat conditioning, avoidance conditioning, extinction, extinction* and *feedback prevention*, and re-extinction (Figure 1). Prior to the onset of threat conditioning, the following instructions were presented:

> *This is end of the practice.*
>
> *During the experiment, you will see four types of geometrical shapes. When the bar appears, you will have to fill up the bar BEFORE the hourglass runs out to successfully avoid the upcoming shock. You will also be asked to make some ratings during the experiment. Slide the mouse across the scale and press the LEFT MOUSE button to make your rating.*
>
> *The experiment takes about 45 minutes.*
>
> *Press SPACEBAR to begin the experiment.*

For 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. Simultaneous with CS onset, the hourglass appeared in the top-right corner of the screen indicating CS duration.

During the *threat conditioning* phase, each CS was presented three times for a total of 12 trials, with stimuli 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, participants were presented with four blocks of 12 trials each, with 48 trials in total. Five seconds after CS onset, the avoidance bar appeared underneath for the remainder of the trial. 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*. Following each block of 12 trials (3 trials per CS), participants made threat expectancy and fear ratings for each CS.

The *extinction* phase began without interruption and consisted of three blocks of 12 trials each (36 trials in total). Avoidance responses continued to increment the avoidance bar according to the FR schedules, and no shocks were presented regardless of responding. Following each block of trials, participants made threat expectancy and fear ratings for each CS.

*Extinction and feedback removal* also involved three blocks of 12 trials with all shocks withheld. Pressing the spacebar, however, now no longer resulted in any visual change to the avoidance bar (i.e., response-contingent feedback was prevented). Again, following each block of trials, participants made threat expectancy and fear ratings for each CS.

Finally, during *re-extinction*, which consisted of 12 trials (3 trials per CS), pressing the spacebar again filled the bar according to either FR-5 or FR-20 schedule requirements but shock was withheld on all trials. Threat expectancy and fear ratings measures were obtained once for each CS and the experiment ended. Participants then had electrodes removed, were debriefed, and compensated.

*Statistical analyses.* Threat expectancy and fear ratings were analysed using 2 (CS) × 2 (response effort) repeated-measures analysis of variance (RM-ANOVAs) in the *threat conditioning* and *re-extinction* phases, a 2 (CS) × 2 (response effort) × 4 (block) RM-ANOVA in the *avoidance conditioning* phase, and 2 (CS) × 2 (response effort) × 3 (block) RM-ANOVAs in the *extinction* and *extinction and feedback removal* phases.

Skin conductance data were processed using *AcqKnowledge* software (BIOPAC Systems) and skin conductance responses (SCRs) were calculated as the first peak (amplitude of the response) to occur within 0.5 to 5 s after CS onset. Prior to analysis, SCRs were range-corrected per participant to account for individual differences (Lykken & Venables, 1971) and square root transformed across all phases to normalise the data (Dawson et al., 2007). Response criterion was set to 0.2 µS and values less than it were scored as zero. Participants who had more than 90% zero responses in all trials were classified as physiological non-responders and excluded from analysis (Marin et al., 2019).

SCR amplitude was analysed using a 2 (CS) × 2 (response effort) × 3 (trial) RM-ANOVA for the *threat conditioning* phase, a 2 (CS) × 2 (response effort) ×12 (trial) RM-ANOVA during the *avoidance conditioning* phase, a 2 (CS) × 2 (response effort) × 9 (trial) RM-ANOVA during the *extinction* phase, and a 2 (CS) × 2 (response effort) × 9 (trial) RM-ANOVA for the *extinction and feedback removal* phase. ANOVAs were not feasible for the *re-extinction* phase due to a lack of variance.

Successful avoidance on each trial was scored as 1 if participants reached the criterion for low effort or high effort (FR-5 or FR-20, respectively), or 0 if they did not reach criterion. Proportion of avoidance and mean total of all avoidance responses (successful or otherwise) were analysed using 2 (CS) × 2 (response effort) ANOVA per phase.

To assess retention of expectancy and fear extinction, 2 (CS) × 2 (response effort) × 2 (phase) RM-ANOVAs were conducted for threat expectancy and fear ratings comparing the last block of *extinction and feedback removal* and the *re-extinction* phase. A similar ANOVA was conducted to examine changes in avoidance proportion due to feedback prevention; however, the independent variable of phase compared the proportion of avoidance for the entire phase of *extinction and feedback removal* and *re-extinction*.

Follow-up analyses were Bonferroni-corrected. Greenhouse–Geisser-corrected degrees of freedom and epsilon (ε) values are reported where the assumption of sphericity is violated. Effect sizes are reported as partial eta-squared ($\eta_p^2$) for RM-ANOVAs, and Cohen's *d* is reported for pairwise comparisons and main effects of only two levels. Cohen's *d* is interpreted according to the criteria of 0.2, 0.5, and 0.8 as small, medium, and large effects, respectively (Cohen, 1988). Alpha was $\alpha = .05$. Analyses were performed in SPSS Version 25 for Mac (IBM Corp., Armonk, NY, USA).

In addition, Bayesian RM-ANOVAs were conducted using JASP v0.13.1 (JASP Team, 2020). Here, we evaluated the weight of evidence for the alternative hypothesis against the null hypothesis ($BF_{10}$), where scores greater than 1 represent evidence for the alternative hypothesis, 1 equals no evidence for either hypothesis, and scores less than 1 represent evidence for the null hypothesis (Lee & Wagenmakers, 2013). For main effects, the $BF_{10}$ is reported, and for interactions, the inclusion Bayes factor ($BF_{incl}$) across matched models is reported.

## Results

### Threat conditioning

*Threat expectancy.* A 2 (CS) × 2 (response effort) RM-ANOVA revealed a significant main effect of CS, $F(1, 35) = 319.62$, $p < .001$, $d = 5.26$, $BF_{10} = 2.298e+54$, with significantly greater threat expectancy towards the CS+ ($M = 91.65$, 95% confidence interval [CI]=[87.41, 95.88], $SD = 12.51$) than the CS− ($M = 9.39$ [3.21, 15.56], $SD = 18.26$).

*Fear ratings.* The 2 (CS) × 2 (response effort) RM-ANOVA for fear ratings during *threat conditioning* revealed a significant main effect of CS, $F(1, 35) = 118.105$, $p < .001$, $d = 2.55$, $BF_{10} = 5.490e+29$. This main effect was superseded by a significant CS × Response Effort interaction, $F(1, 35) = 6.63$, $p = .014$, $\eta_p^2 = .159$, $BF_{incl} = 0.246$. Post hoc Bonferroni-corrected simple main effects revealed significantly greater fear ratings to the low-effort CS+ relative to the high-effort CS+ ($M_{DIFF} = 8.01$, $SEM_{DIFF} = 2.90$, $p < .001$), with no differences between the low- and high-effort CS− ($p = .300$). As the response effort manipulation does not become relevant until the *avoidance conditioning* phase, this effect is unpredictable and unlikely to represent a meaningful difference.

*SCR.* A 2 (CS) × 2 (response effort) × 3 (trial) RM-ANOVA revealed a significant main effect of CS, $F(1, 29) = 4.91$, $p = .035$, $d = 0.48$, $BF_{10} = 1.000$, with larger SCR to the CS+ ($M = 0.36$ [0.30, 0.41], $SD = 0.15$) than the CS− ($M = 0.29$ [0.24, 0.34], $SD = 0.14$). No further main effects or interactions were significant (all *p*s > .077).

### Avoidance conditioning

*Threat expectancy.* During the *avoidance conditioning* phase, a 2 (CS) × 2 (response effort) × 4 (block) RM-ANOVA revealed a significant main effect of CS, $F(1, 35) = 39.82$, $p < .001$, $d = 1.43$, $BF_{10} = 2.216e+62$, with greater threat expectancy to the CS+ ($M = 51.43$ [39.27, 63.60], $SD = 35.96$) than the CS− ($M = 10.78$ [4.70, 16.86], $SD = 17.98$). There was also a significant main effect of trial block, $F(2.05, 71.78) = 5.08$, $p = .008$, $\eta_p^2 = .127$, $\varepsilon = .684$, $BF_{10} = 0.024$, with decreasing threat expectancy ratings throughout the phase. Furthermore, there was a significant CS × Response Effort interaction, $F(1, 35) = 5.29$, $p = .028$, $\eta_p^2 = .131$, $BF_{10} = 0.249$.

*Fear ratings.* All significant main effects and first-order interactions were superseded by a significant CS × Response Effort × Block interaction, $F(2.28, 79.95) = 5.74$, $p = .003$, $\eta_p^2 = .141$, $\varepsilon = .761$, $\text{BF}_{\text{incl}} = 0.104$. Bonferroni-corrected simple interaction effects revealed that there were no significant differences in fear ratings to the low- and high-effort CS− throughout the *avoidance conditioning* phase (all $p\text{s} > .098$). For the CS+, while there was no significant difference in fear ratings between the high- and low-effort stimuli during the first block of trials ($p = .736$), the high-effort CS+ was rated significantly more fearful than the low-effort CS+ for the remaining trial blocks with this difference increasing from Block 2 ($M_{DIFF} = 11.31$, $SEM_{DIFF} = 4.11$, $p = .009$) to Block 3 ($M_{DIFF} = 13.16$, $SEM_{DIFF} = 3.55$, $p = .001$), with the greatest difference in ratings following trial Block 4 ($M_{DIFF} = 14.69$, $SEM_{DIFF} = 3.24$, $p < .001$).

*SCR.* The 2 (CS) × 2 (response effort) × 12 (trial) RM-ANOVA also showed significantly larger SCRs to the CS+ ($M = 0.27$ [0.22, 0.32], $SD = 0.13$) than the CS− ($M = 0.23$ [0.18, 0.27], $SD = 0.12$), $F(1, 29) = 6.40$, $p = .017$, $d = 0.32$, $\text{BF}_{10} = 0.332$. Furthermore, there was a significant trial main effect, $F(7.60, 220.53) = 3.09$, $p = .003$, $\eta_p^2 = .096$, $\varepsilon = .691$, $\text{BF}_{10} = 0.030$, with a decreasing, non-differential, pattern of SCRs throughout the *avoidance conditioning* phase. There were no other significant main effects or interactions (all $p\text{s} > .129$).

*Proportion of avoidance.* All participants met the learning criteria during avoidance conditioning. A 2 (CS) × 2 (response effort) RM-ANOVA revealed a significant main effect of CS, $F(1, 35) = 73.03$, $p < .001$, $d = 1.98$, $\text{BF}_{10} = 2.316\text{e}{+}25$, with greater proportion of avoidance to the CS+ ($M = 97.45$ [95.76, 99.15], $SD = 5.01$) than the CS− ($M = 34.95$ [19.91, 50.00], $SD = 44.47$). There was no significant main effect of response effort, $F(1, 35) = 3.33$, $p = .076$, $d = 0.06$, $\text{BF}_{10} = 0.177$, or CS × Response Effort interaction, $F(1, 35) < .001$, $p = 1.00$, $\eta_p^2 = .00$, $\text{BF}_{\text{incl}} = 0.238$.

Analysis of mean number of avoidance responses revealed a significant Cue × Effort interaction, $F(1, 11) = 971.405$, $p < .001$, $\eta_p^2 = .117$, $\text{BF}_{10} = 5.059\text{e}{-}29$, with follow-up tests highlighting significantly greater responding on CS+ High-effort trials relative to all other trial types ($p < .001$).

### Extinction

*Threat expectancy.* The 2 (CS) × 2 (response effort) × 3 (block) RM-ANOVA revealed a significant main effect of CS, $F(1, 35) = 19.67$, $p < .001$, $d = 1.53$, $\text{BF}_{10} = 2.264\text{e}{+}20$, with significantly greater threat expectancy to the CS+ ($M = 31.25$ [20.16, 42.34], $SD = 19.08$) compared with the CS− ($M = 6.56$ [2.29, 10.83], $SD = 12.624$). There was also a significant main effect of block, $F(1.42, 49.84) = 6.46$, $p = .007$, $\eta_p^2 = .156$, $\varepsilon = .712$, $\text{BF}_{10} = 0.943$, with a reduction in threat expectancy throughout the phase (see Figure 2). No further main effects or interactions were significant (all $p\text{s} > .084$).

*Fear ratings.* A 2 (CS) × 2 (response effort) × 3 (block) RM-ANOVA revealed significant main effects of CS, $F(1, 35) = 23.14$, $p < .001$, $d = 1.01$, $\text{BF}_{10} = 1.545\text{e}{+}24$; response effort, $F(1, 35) = 11.92$, $p = .001$, $d = 0.47$, $\text{BF}_{10} = 268.974$; and block, $F(1.43, 50.08) = 9.86$, $p = .001$, $\eta_p^2 = .220$, $\varepsilon = .715$, $\text{BF}_{10} = 1.486$. These main effects were superseded by significant first-order CS × Response Effort, $F(1, 35) = 6.11$, $p = .018$, $\eta_p^2 = .149$, $\text{BF}_{\text{incl}} = 0.596$, and CS × Block interactions, $F(1.52, 53.11) = 8.01$, $p = .002$, $\eta_p^2 = .186$, $\varepsilon = .759$, $\text{BF}_{\text{incl}} = 0.153$. Follow-up Bonferroni-corrected tests of simple main effects for the CS × Response Effort interaction revealed that the high-effort CS+ was rated significantly more fearful than the low-effort CS+ ($M_{DIFF} = 11.61$, $SEM_{DIFF} = 3.00$, $p < .001$). Similarly, the high-effort CS− was also rated more fearfully than the low-effort CS−, although avoidance was not necessary ($M_{DIFF} = 5.61$, $SEM_{DIFF} = 2.52$, $p = .033$). Follow-up Bonferroni-corrected tests of simple main effects for the CS × Block interaction revealed that, for the CS+, fear ratings reduced significantly from trial Block 1 to Block 2 ($M_{DIFF} = 5.38$, $SEM_{DIFF} = 1.78$, $p = .014$), and from Block 2 to Block 3 ($M_{DIFF} = 5.87$, $SEM_{DIFF} = 2.14$, $p = .029$). For the CS−, the reductions in fear ratings were not significant from Block 1 to Block 2 ($M_{DIFF} = 2.90$, $SEM_{DIFF} = 1.30$, $p = .097$) or from Block 2 to Block 3 ($M_{DIFF} = 1.38$, $SEM_{DIFF} = 1.12$, $p = .676$). No further interactions were significant (all $p\text{s} > .133$).

*SCR.* Similar to the *avoidance conditioning* phase, the 2 (CS) × 2 (response effort) × 9 (trial) RM-ANOVA revealed a significant main effect of CS, $F(1, 29) = 14.54$, $p = .001$, $d = 0.57$, $\text{BF}_{10} = 6.479\text{e}{-}6$, with larger SCRs to the CS+ ($M = 0.13$ [0.07, 0.18], $SD = 0.15$) than the CS− ($M = 0.06$ [0.03, 0.10], $SD = 0.09$). There was also a significant main effect of trial, $F(3.86, 111.99) = 7.23$, $p < .001$, $\eta_p^2 = .199$, $\varepsilon = .483$, $\text{BF}_{10} = 2.040\text{e}{-}8$, with a non-differential reduction in SCRs throughout the *extinction* phase (see Figure 3). No further main effects or interactions were significant (all $p\text{s} > .131$).

*Proportion of avoidance.* During the *extinction* phase, there was a significant main effect of CS, $F(1, 35) = 50.16$, $p < .001$, $d = 1.49$, $\text{BF}_{10} = 1.381\text{e}{+}19$, with greater avoidance to the CS+ ($M = 87.35$ [77.83, 96.86], $SD = 28.13$) than the CS− ($M = 31.02$ [15.65, 46.39], $SD = 45.43$). There was no significant main effect of response effort, $F(1, 35) = 2.41$, $p = .130$, $d = 0.05$, $\text{BF}_{10} = 0.180$, or CS × Response Effort interaction, $F(1, 35) = 0.09$, $p = .768$, $\eta_p^2 = .003$, $\text{BF}_{\text{incl}} = 0.229$.

Analysis of mean number of avoidance responses (Figure 4) revealed a significant Cue × Effort interaction, $F(1, 8) = 701.217$, $p < .001$, $\eta_p^2 = .121$, $\text{BF}_{10} = 5.214\text{e}{-}19$, with follow-up tests highlighting significantly greater responding

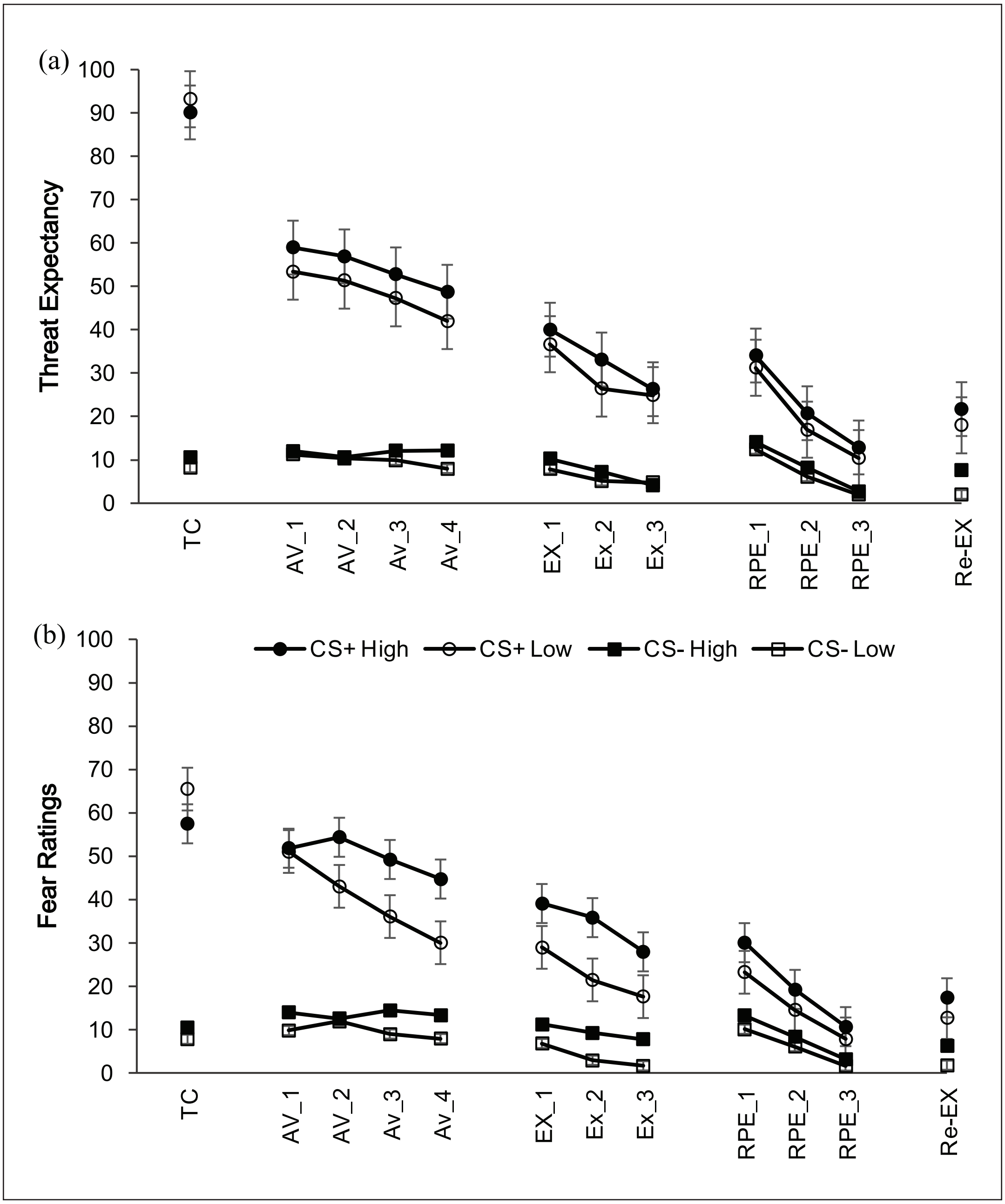

**Figure 2.** Threat expectancy (a) and fear (b) ratings in Experiment 1. TC refers to threat conditioning, AV refers to avoidance conditioning, EX refers to extinction, EX + FR refers to the extinction and feedback removal phase, and Re-EX refers to the re-extinction test phase. Numbers represent blocks per phase. Error bars represent SEM.

on CS+ High-effort trials relative to all other trial types ($p < .001$).

### *Extinction and feedback removal*

*Threat expectancy.* While participants did not receive feedback for making avoidance responses during the *extinction and feedback removal* phase, a 2 (CS) × 2 (response effort) × 3 (block) RM-ANOVA revealed significant main effects of CS, $F(1, 35) = 10.15$, $p = .003$, $d = 0.62$, $BF_{10} = 2.887e+8$, and block, $F(1.66, 58.03) = 14.56$, $p < .001$, $\eta^2_p = .294$, $\varepsilon = .829$, $BF_{10} = 1.382e+7$. These main effects were superseded by a significant CS × Block interaction, $F(1.50, 40.17) = 7.26$, $p = .008$, $\eta^2_p = .172$, $\varepsilon = .574$, $BF_{incl} = 0.594$. Post hoc Bonferroni-corrected tests of simple main effects revealed, for the CS+, a significant decrease in threat expectancy ratings from trial Block 1 to Block 2 ($M_{DIFF} = 13.79$, $SEM_{DIFF} = 3.61$, $p = .002$), and from Block 2 to Block 3 ($M_{DIFF} = 7.23$, $SEM_{DIFF} = 2.80$, $p = .043$). For the CS−, the reduction in threat expectancy was not significant from trial Block 1 to Block 2 ($M_{DIFF} = 6.11$, $SEM_{DIFF} = 2.47$, $p = .055$), or from Block 2 to Block 3 ($M_{DIFF} = 4.80$, $SEM_{DIFF} = 2.53$, $p = .198$). These results

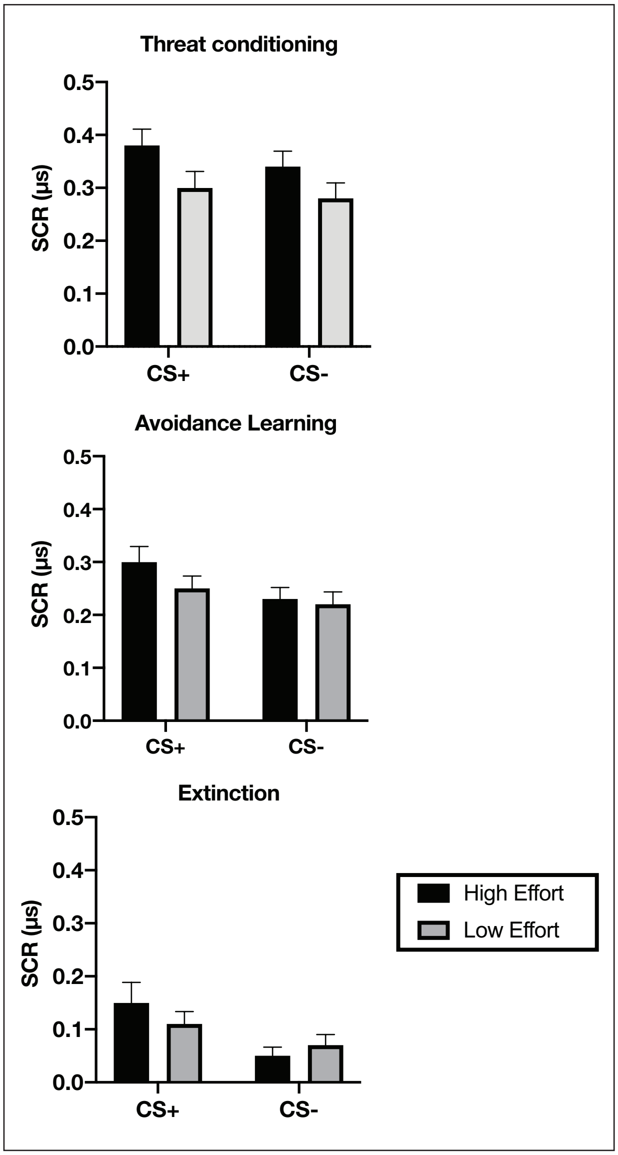

**Figure 3.** Skin conductance responses in Experiment 1. Error bars represent SEM.

show that during the *extinction and feedback removal* phase, there was a steeper reduction in threat expectancy for the CS+ than the CS− (see Figure 2). There were no further significant main effects or interactions during the *extinction and feedback removal* phase (all $ps > .147$).

*Fear ratings.* Similar to the *extinction* phase, a 2 (CS) × 2 (response effort) × 3 (block) RM-ANOVA revealed significant main effects of CS, $F(1, 35) = 9.34$, $p = .004$, $d = 0.53$, $\mathrm{BF}_{10} = 1.395\mathrm{e}{+}7$, response effort, $F(1, 35) = 11.86$, $p = .002$, $d = 0.21$, $\mathrm{BF}_{10} = 0.802$, and block, $F(1.54, 53.77) = 11.35$, $p < .001$, $\eta_p^2 = .245$, $\varepsilon = .768$, $\mathrm{BF}_{10} = 8.114\mathrm{e}{+}6$, and these main effects were superseded by First-Order CS × Response Effort, $F(1, 35) = 5.38$, $p = .026$, $\eta_p^2 = .133$, $\mathrm{BF}_{\mathrm{incl}} = 0.217$, and CS × Block interactions, $F(1.17, 40.18) = 5.07$, $p = .025$, $\eta_p^2 = .126$, $\varepsilon = .583$, $\mathrm{BF}_{\mathrm{incl}} = 0.364$. Follow-up tests of simple main effects for the CS × Response Effort interaction revealed that the fear ratings were significantly higher for the high-effort CS+ than the low-effort CS+ ($M_{DIFF} = 4.81$, $SEM_{DIFF} = 1.33$, $p = .001$), with a similar pattern for the high- and low-effort CS− ($M_{DIFF} = 2.30$, $SEM_{DIFF} = 0.98$, $p = .024$). Follow-up tests for the CS × Block interaction found that, like the *extinction* phase, fear ratings to the CS+ decreased significantly from Block 1 to Block 2 ($M_{DIFF} = 9.74$, $SEM_{DIFF} = 2.96$, $p = .007$) and from Block 2 to Block 3 ($M_{DIFF} = 7.62$, $SEM_{DIFF} = 2.81$, $p = .031$). The reduction in fear ratings to the CS− was not significant from Block 1 to Block 2 ($M_{DIFF} = 4.48$, $SEM_{DIFF} = 2.11$, $p = .121$) and from Block 2 to Block 3 ($M_{DIFF} = 4.80$, $SEM_{DIFF} = 2.64$, $p = .231$).

*SCR.* There were no significant main effects or interactions during the *extinction and feedback removal* phase (all $ps > .069$).

*Proportion of avoidance.* The 2 (CS) × 2 (response effort) RM-ANOVA revealed significantly greater avoidance to the CS+ ($M = 26.08$ [14.84, 37.32], $SD = 33.22$) than the CS− ($M = 13.58$ [4.10, 23.06], $SD = 28.03$), $F(1, 35) = 8.71$, $p = .006$, $d = 0.41$, $\mathrm{BF}_{10} = 220.596$. Furthermore, there was also a significant main effect of response effort, $F(1, 35) = 5.35$, $p = .027$, $d = 0.20$, $\mathrm{BF}_{10} = 0.703$, with greater avoidance to the low-effort stimuli ($M = 22.69$ [12.97, 32.40], $SD = 28.71$) than the high-effort stimuli ($M = 16.98$ [7.10, 26.85], $SD = 29.18$). There was no significant CS × Response Effort interaction, $F(1, 35) = 0.50$, $p = .483$, $\eta_p^2 = .014$, $\mathrm{BF}_{\mathrm{incl}} = 0.231$.

Analysis of mean number of avoidance responses revealed a significant Cue × Effort interaction, $F(1, 8) = 18.501$, $p = .003$, $\eta_p^2 = .032$, $\mathrm{BF}_{10} = 0.051$, with follow-up tests highlighting significantly greater responding on CS+ High-effort trials relative to all other trial types ($p < .001$ to $p = .002$).

### *Re-extinction*

*Threat expectancy.* Threat expectancy ratings remained significantly higher to the CS+ ($M = 19.84$ [9.78, 29.90], $SD = 29.74$) than the CS− ($M = 6.25$ [1.84, 10.66], $SD = 13.03$) during the *re-extinction* phase, $F(1, 35) = 8.74$, $p = .006$, $d = 0.59$, $\mathrm{BF}_{10} = 138.719$. There was no significant main effect of response effort, $F(1, 35) = 0.09$, $p = .770$, $d = 0.05$, $\mathrm{BF}_{10} = 0.179$ or CS × Response Effort interaction, $F(1, 35) = 1.84$, $p = .183$, $\eta_p^2 = .050$, $\mathrm{BF}_{\mathrm{incl}} = 0.316$.

A 2 (CS) × 2 (response effort) × 2 (phase) RM-ANOVA comparing the final block of the *extinction and feedback removal* phase with the first block of the *re-extinction* phase revealed significantly greater threat expectancy ratings to the CS+ ($M = 15.74$ [3.38, 24.10], $SD = 24.71$) than the CS− ($M = 4.30$ [1.35, 7.24], $SD = 8.71$), $F(1, 35) = 8.57$, $p = .006$, $d = 0.62$, $\mathrm{BF}_{10} = 45{,}325.419$. Furthermore, there was a significant main effect of phase, $F(1, 35) = 6.26$, $p = .017$, $d = 0.62$, $\mathrm{BF}_{10} = 3.916$, showing a significant increase in threat expectancy in the *re-extinction* phase ($M = 13.04$ [6.83, 19.26], $SD = 3.06$) relative to the final block of the *extinction and feedback removal* phase ($M$=6.99 [2.45, 11.53], $SD$=13.42). No further main effects or interactions were significant, with no differences due to high or low response effort (all $ps > .142$).

*Fear ratings.* During the *re-extinction* phase, fear ratings remained significantly higher to the CS+ ($M$=15.04 [6.48, 23.60], $SD$=25.31) compared to the CS− ($M$=4.00 [0.79, 7.20], $SD$=9.46), $F(1, 35)=7.89$, $p=.008$, $d=0.37$, $BF_{10}=268.885$. It is worth noting that, despite the significant difference between fear ratings to the CS+ and CS−, fear ratings here are quite low, relatively speaking. There was no significant main effect of response effort, $F(1, 35)=3.46$, $p=.071$, $\eta_p^2=.090$, $BF_{10}=0.552$, or CS × Response Effort interaction, $F(1, 35)=0.02$, $p=.894$, $\eta_p^2=.001$, $BF_{incl}=738{,}374.575$.

The 2 (CS) × 2 (response effort) × 2 (phase) RM-ANOVA comparing the final block of the *extinction and feedback removal* phase with the *re-extinction* phase revealed a significant main effect of CS, $F(1, 35)=7.68$, $p=.009$, $d=0.58$, $BF_{10}=31{,}093.278$, with greater fear reported to the CS+ ($M$=12.15 [5.22, 12.08], $SD$=20.48) than the CS− ($M$=3.18 [0.63, 5.74], $SD$=7.55). There was also a significant main effect of response effort, $F(1, 35)=4.87$, $p=.034$, $d=0.26$, $BF_{10}=0.639$, with greater fear reported to the high response-effort stimuli ($M$=9.34 [4.50, 14.18], $SD$=14.30) than the low-effort stimuli ($M$=6.00 [2.22, 9.78], $SD$=11.17). No further main effects or interactions were significant (all $ps > .081$).

*SCR.* There were no significant main effects or interactions during *re-extinction* (all $ps > .05$).

*Proportion of avoidance.* During the *re-extinction* phase, the 2 (CS) × 2 (response effort) RM-ANOVA revealed a significant main effect of CS, $F(1, 35)=11.29$, $p=.002$, $d=0.55$, $BF_{10}=25{,}358.563$, with significantly greater avoidance to the CS+ ($M$=51.39 [35.88, 66.90], $SD$=45.84) than the CS− ($M$=27.32 [13.30, 41.33], $SD$=41.44). There was no significant main effect of response effort, $F(1, 35)=0.49$, $p=.487$, $d=0.05$, $BF_{10}=0.189$, and no significant CS × Response Effort interaction, $F(1, 35)=0.25$, $p=.624$, $\eta_p^2=.007$, $BF_{incl}=0.274$.

A 2 (CS) × 2 (response effort) × 2 (phase) RM-ANOVA was conducted to compare the change in proportion of avoidance from the *extinction and feedback removal* phase to the *re-extinction* phase. There was a significant main effect of phase, $F(1, 35)=17.42$, $p<.001$, $d=0.58$, $BF_{10}=962{,}772.064$, with an increase in avoidance from the *extinction and feedback removal* ($M$=19.83 [10.36, 29.30], $SD$=27.98) to the *re-extinction* phase ($M$=39.35 [26.48, 52.22], $SD$=38.04). As expected, there was significantly greater avoidance to the CS+ ($M$=38.74 [26.80, 50.67], $SD$=35.27) than the CS− ($M$=20.45 [9.42, 31.47], $SD$=32.59), $F(1, 35)=12.81$, $p=.001$, $d=0.54$, $BF_{10}=122{,}314.300$. Similarly, there was significantly greater avoidance to the low-effort stimuli ($M$=31.48 [20.89, 42.07], $SD$=31.30) than the high-effort stimuli ($M$=27.70 [17.46, 37.94], $SD$=30.27), $F(1, 35)=4.29$, $p=.046$, $d=0.12$, $BF_{10}=0.224$, although Bayesian analysis indicated no evidence for the alternative hypothesis over the null. No further interactions were significant (all $Fs < 3.92$, all $ps > .056$).

Analysis of mean number of avoidance responses revealed a significant Cue × Effort interaction, $F(1, 2)=44.212$, $p=.022$, $\eta_p^2=.020$, $BF_{10}=0.362$, with follow-up tests highlighting significantly greater responding on CS+ High-effort trials relative to all other trial types ($p$=.019 to $p$=.008) with comparable and non-significant avoidance only on CS+ Low and CS− Low trials ($p$=.094).

## Discussion

In Experiment 1, we conducted a within-subjects differential conditioned avoidance task, whereby the CS+ and CS− each had *Low* (FR-5) and *High* (FR-20) response effort requirements. We hypothesised that low response effort would result in ongoing avoidance, threat expectancy, and fear ratings to the *Low*-effort CS+, while there would be a reduction in avoidance, threat expectancy, and fear to the *High*-effort CS+ throughout the *extinction* phase. This hypothesis was not supported, with a stimulus non-specific reduction in threat expectancy throughout *extinction*. Alternatively, fear ratings showed that participants reported greater fear of the High-effort CS+ and the High-effort CS− compared with their low-effort counterparts. This effect did not result in a greater or lesser proportion of avoidance during the *extinction* phase. These results from Experiment 1 bear similarity with previous demonstrations of avoidance and protection from extinction (Lovibond et al., 2009; Rattel et al., 2017; Volders et al., 2012).

Second, we also predicted that, during the *extinction and feedback removal* phase, there would be ongoing avoidance, threat expectancy, and fear ratings to the low-effort CS+ with a reduction in responding to the *High*-effort CS+. This hypothesis was partially supported. While participants avoided the *Low*-effort CS+ more frequently than the *High*-effort CS+ during the *extinction and feedback removal* phase, participants continued to report greater fear to the *High*-effort CS+ compared with the *Low*-effort CS+. These findings may be related, at least in part, to participants perceiving avoidance as ineffective given the absence of on-screen visual changes in the avoidance bar. This possibility, combined with the fact that this phase was not, strictly speaking, an extinction with response-prevention procedure (ERP) because participants could still make the avoidance response, may explain the abrupt decrease in avoidance. In future work, actively preventing avoidance (rather than signalling that

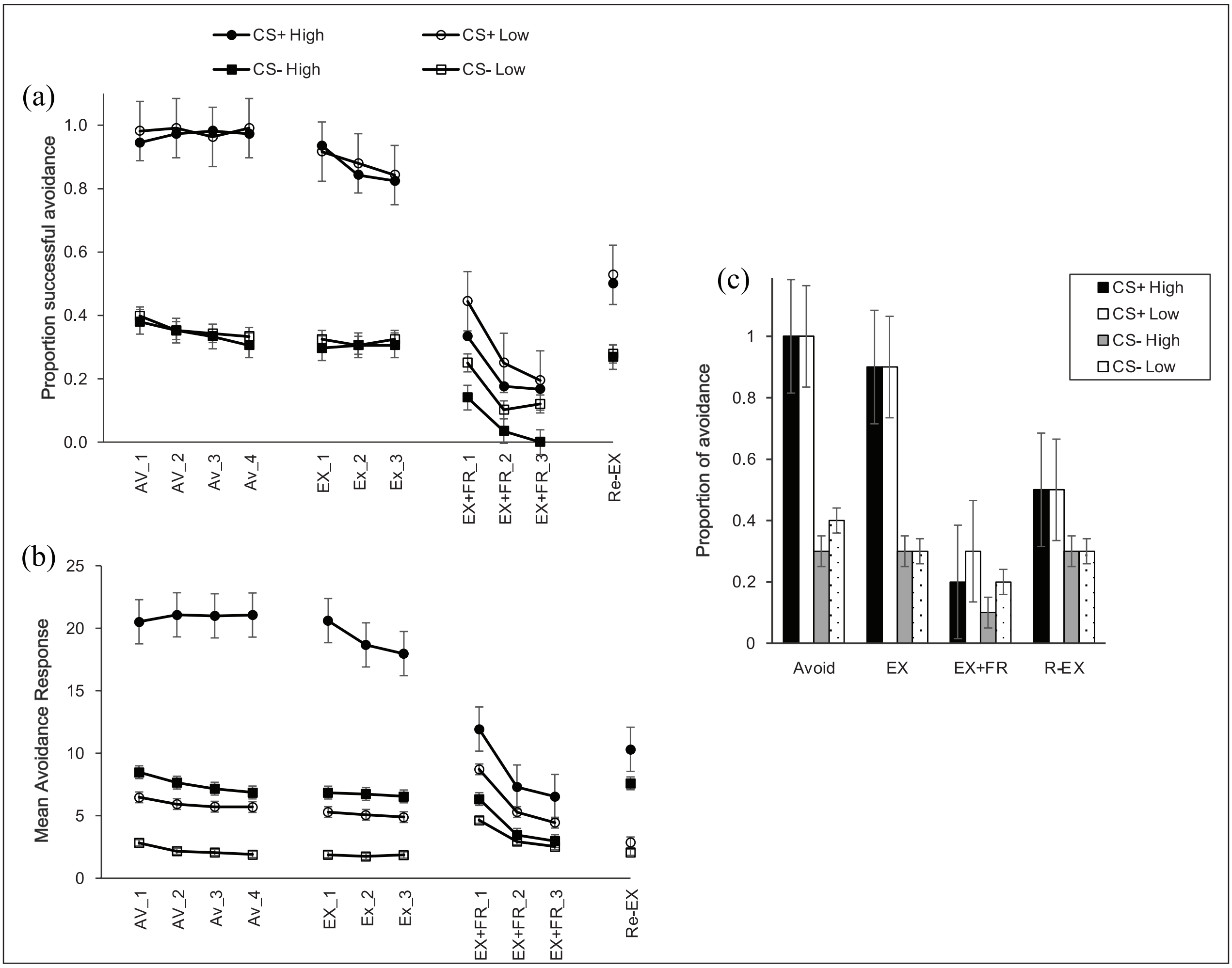

**Figure 4.** *Proportion of avoidance responses and the mean number of responses per CS across all phases in Experiment 1.* (a) Successful avoidance was achieved by pressing the spacebar either 5 times (Low) or 20 times (High) in the presence of the respective CS+. (b) Total mean number of avoidance responses per phase. (c) Proportion of avoidance per phase. Error bars represent SEM.

avoidance was no longer available; Dymond, 2019; Vervliet & Indekeu, 2015) would constitute a stronger test of competing accounts.

Third, our hypothesis was supported that there were no differences between the *High*- or *Low*-effort CS+ during the *re-extinction* phase, as indicated by Bayesian analysis. We did, however, find that rates of avoidance increased from the *extinction and feedback removal* phase to the *re-extinction* phase. It could be that the removal of on-screen visual feedback increased the perceived threat value of stimuli or functioned as a safety signal (Fernando et al., 2014; Weisman & Litner, 1972) or functioned in compound with a form of ERP to generate the observed effects. While frequentist analysis indicated that this effect was most prominent for low-effort stimuli, Bayesian statistics did not support the evidence for the alternative hypothesis over the null. Overall, the return to a *re-extinction* phase may have been comparable to removing response prevention which is known to promote recovery or renewal of fear and low-cost avoidance (Vervliet & Indekeu, 2015).

In summary, the key finding from Experiment 1 was that removal, or prevention, of visual feedback about the potential success of avoidance at cancelling shock resulted in return of CS+ avoidance during the *re-extinction* phase. Furthermore, there was an increase in threat expectancy from the *extinction and feedback removal* phase to the *re-extinction* phase, but this effect did not differ based on CS+/− or the response effort required for avoidance.

In Experiment 2, the *extinction and feedback removal* phase was extended, and, as there were no differences in proportion of avoidance between the *avoidance* and *extinction* phases in Experiment 1, the *extinction* phase was removed, and participants moved directly to *extinction and feedback removal* (Figure 5). To test the persistence of extinction, avoidance was observed during return of fear conditions. Return of fear is a common phenomenon observed following anxiety treatment and diminishing return of fear is a priority in improving therapeutic efficacy (Vervliet et al., 2013). We used reinstatement as the laboratory model of return of fear, which involves unsignalled

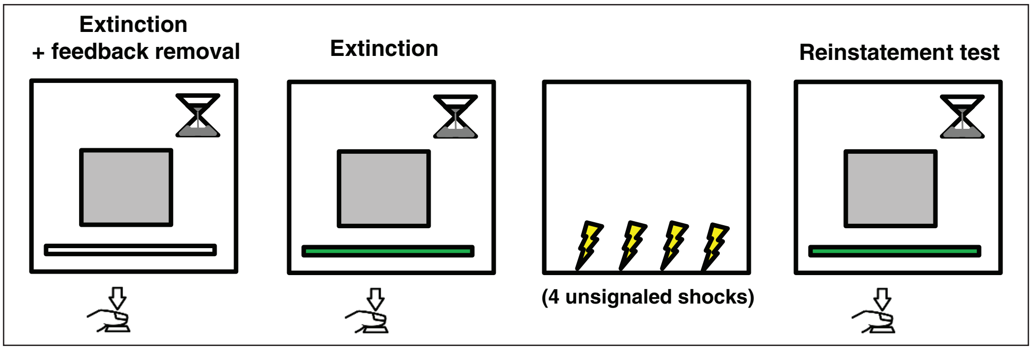

**Figure 5.** *Partial overview of test phases in Experiment 2*. Following *threat conditioning* and *avoidance conditioning* as in Experiment 1, during the *fear extinction and feedback removal* phase, responses made in the presence of both CS+ and CS− no longer filled up the avoidance bar and all shocks were withheld. During *extinction*, both partial and complete responses filled the avoidance bar and shocks continued to be withheld. Then, in the presence of a blank screen, participants received four unsignalled shocks before being re-tested for reinstatement of avoidance under conditions identical to *extinction*.

presentations of the US following fear extinction (Bouton, 2002; Haaker et al., 2014; Zuj et al., 2018). Reinstatement is analogous to situations in which a person is exposed to the feared outcome without prior CS signalling, such as experiencing a panic attack in the absence of expected situational triggers (Zuj & Norrholm, 2019). Little research has been conducted on the reinstatement of avoidance (Cameron et al., 2015; Urcelay et al., 2019). Krypotos and Engelhard (2018) found that their reinstatement test showed an increase in anxiety ratings, but no marked difference in avoidance responding, compared with the preceding extinction with response-prevention phase. The authors hypothesised that avoidance responding remained low as participants were not explicitly instructed on the effectiveness of the avoidance response during the test phase.

In Experiment 2, therefore, reinstatement stimuli were presented to investigate whether fear ratings would similarly increase following reinstatement without discernible difference in avoidance responding and whether response effort would influence avoidance responding following reinstatement. The primary hypothesis for Experiment 2 was that following an extended *extinction and feedback removal* phase, avoidance, threat expectancy, and fear ratings of the CS+ High and CS+ Low would both decrease. The secondary hypothesis was that following reinstatement, return of elevated threat expectancy, and fear ratings would be observed but with no effect on proportion of avoidance responding.

## Experiment 2

### Method

*Participants.* A total of 41 participants were recruited from Swansea University, aged from 18 to 32 years ($M$=20.2 years, $SD$=2.8 years), with 26 females and 15 males. Three participants failed to meet the learning criteria in *avoidance conditioning* and one participant was excluded due to missing data, resulting in a total of 37 participants' data eligible for analysis. The study was approved by the School of Psychology Research Ethics Committee, Swansea University and informed consent was provided at the outset. Participants were compensated with course credit on completion.

*Apparatus.* Apparatus used, measures obtained, and stimuli presented were identical to those described in Experiment 1.

*Procedure.* As before, participants completed the consent form upon arrival and then underwent shock calibration and the attachment of skin conductance electrodes. Participants were given identical instructions, habituation, and training trials as Experiment 1. Experiment 2 consisted of five phases: *threat conditioning*, *avoidance conditioning*, *extinction and feedback removal*, *extinction*, and *re-extinction*. Notable changes to Experiment 2 from Experiment 1 include swapping the order of phases with the *extinction and feedback removal phase* coming before the *extinction* phase, and the *re-extinction* phase coming after unsignalled reinstatement stimuli.

Reinstatement began immediately following the *extinction* phase and involved a blank screen for 2 min with four unsignalled deliveries of the US in the absence of a preceding CS. The timing of shocks was pseudo-randomised with an interval ranging between 20 and 40 s between shocks. The *re-extinction* phase began immediately following reinstatement, whereby participants were presented with a total of 12 trials (3 trials per CS). Like the *avoidance conditioning* phase, the avoidance bar would appear 5 s after CS onset and would "fill up" incrementally following each spacebar press dependent on response effort condition. Regardless of the number of avoidance responses made, CSs were never followed by the US throughout the *re-extinction* phase.

*Statistical analyses.* SCR amplitude data were range-corrected for each participant and square root transformed prior to analyses. Due to missing data, 15 participants were excluded from analyses of SCR, and a further 5 participants were identified as "non-responders," which resulted in a final sample for SCR analyses of $n = 17$.

Analysis strategies were the same as Experiment 1, with the addition of a pre- and post-reinstatement analysis. Reinstatement effects were analysed using a 2 (CS) × 2 (response effort) × 2 (time) RM-ANOVA, where time used the average of the final trial block of the *extinction* phase and the first block of the *re-extinction* phase (Haaker et al., 2014; Kindt et al., 2009; Soeter & Kindt, 2010; Zuj et al., 2018). Greenhouse–Geisser-corrected degrees of freedom and epsilon values are reported for all RM-ANOVAs. Effect sizes are reported as partial eta-squared $(\eta_p^2)$ and Cohen's *d*. As with Experiment 1, the Bayes factor is reported for all analyses, assessing the weight of the evidence for the alternative hypothesis against the null hypothesis ($BF_{10}$). For main effects, the $BF_{10}$ is reported, and for interactions, the inclusion Bayes factor ($BF_{incl}$) across matched models is reported.

## Results

### Threat conditioning

*Threat expectancy.* A 2 (CS) × 2 (response effort) RM-ANOVA revealed a significant main effect of CS, $F(1, 36) = 214.23$, $p < .001$, $d = 4.56$, $BF_{10} = 5.339e+43$, with higher threat expectancy ratings to the CS+ ($M = 87.24$ [81.48, 93.00], $SD = 17.28$) than the CS− ($M = 11.36$ [6.03, 16.69], $SD = 15.97$). No further main effects or interactions were significant (*p*s > .399).

*Fear ratings.* The 2 (CS) × 2 (response effort) RM-ANOVA revealed a significant main effect of CS, $F(1, 36) = 95.32$, $p < .001$, $d = 2.27$, $BF_{10} = 1.000$, and a significant CS × Response Effort interaction, $F(1, 36) = 4.72$, $p = .037$, $\eta_p^2 = .116$, $BF_{incl} = 1.729$, showing significantly lower fear ratings for the high-effort CS+ than the low-effort CS+ ($p = .017$), and no significant difference between the low- and high-effort CS− ($p = .239$). As the low- and high-effort contingencies are not encountered until the *avoidance conditioning* phase, we expect this difference is due to random variation. The response effort main effect was not significant ($p = .323$).

*SCR.* During *threat conditioning*, there was a significant main effect of CS, $F(1, 13) = 15.28$, $p = .002$, $d = 1.13$, $BF_{10} = 1.000$, which was superseded by a significant CS × Trial interaction, $F(1.62, 21.05) = 5.60$, $p = .015$, $\eta_p^2 = .301$, $\varepsilon = .810$, $BF_{incl} = 5.847$. Follow-up test of simple main effects showed that there was no significant difference between the CS+ and CS− at Trial 1 ($p = .702$), but there was at Trial 3 ($p < .001$) suggesting differential conditioned responding in SCR amplitude data (Figure 7). No other main effects or interactions were significant (all *p*s > .100).

### Avoidance conditioning

*Threat expectancy.* Here, threat expectancy was significantly higher for the CS+ ($M = 49.69$ [39.73, 59.65], $SD = 29.87$) than the CS− ($M = 6.71$ [2.82, 10.60], $SD = 11.66$), $F(1, 36) = 83.59$, $p < .001$, $d = 1.90$, $BF_{10} = 1.242e+93$. Furthermore, there was a significant main effect of response effort, $F(1, 36) = 8.96$, $p = .005$, $d = 0.23$, $BF_{10} = 0.355$, with higher threat expectancy for high-effort stimuli ($M = 30.27$ [24.01, 36.53], $SD = 18.78$) than low-effort stimuli ($M = 26.13$ [20.33, 31.93], $SD = 17.38$). Finally, there was a significant main effect of trial block, $F(2.20, 79.04) = 7.72$, $p = .001$, $\eta_p^2 = .177$, $\varepsilon = .732$, $BF_{10} = 0.054$, with threat expectancy decreasing throughout the *avoidance conditioning* phase (see Figure 6). No further main effects or interactions were significant (all *p*s > .079).

*Fear ratings.* The 2 (CS) × 2 (response effort) × 4 (trial block) RM-ANOVA revealed significant main effects of CS, $F(1, 35) = 85.31$, $p < .001$, $d = 1.96$, $BF_{10} = 0.825$, and block, $F(1.58, 55.27) = 5.28$, $p = .013$, $\eta_p^2 = .131$, $\varepsilon = .602$, $BF_{10} = 1.870e{-}89$. These main effects were superseded by a significant CS × Block interaction, $F(1.81, 63.22) = 3.48$, $p = .041$, $\eta_p^2 = .090$, $\varepsilon = .602$, $BF_{incl} = 0.152$, showing a significant reduction in fear ratings to the CS+ from Block 1 to Block 4 ($p = .030$) with no significant change for fear ratings to the CS− ($p = 1.000$). There was also a significant main effect of response effort, $F(1, 35) = 9.90$, $p = .003$, $d = 0.23$, $BF_{10} = 1.012e{-}88$, with higher fear ratings reported to the high-effort stimuli ($M = 31.47$ [25.71, 37.22], $SD = 17.01$) than the low-effort stimuli ($M = 27.57$ [21.92, 33.23], $SD = 16.70$). No further main effects or interactions were significant (*p*s > .094).

*Proportion of avoidance.* Three participants failed to meet the learning criteria and were excluded from further analysis. As participants learned the avoidance response, there was a significant main effect of CS, $F(1, 36) = 141.29$, $p < .001$, $d = 2.77$, $BF_{10} = 1.000$, with greater proportion of avoidance to the CS+ ($M = 96.73$ [94.17, 99.30], $SD = 7.68$) than the CS− ($M = 23.20$ [10.93, 35.47], $SD = 36.81$). Furthermore, there was a significant main effect of response effort, $F(1, 36) = 6.93$, $p = .012$, $d = 0.20$, $BF_{10} = 9.493e{-}39$, with significantly greater proportion of avoidance to the low-effort stimuli ($M = 61.94$ [55.49, 68.38], $SD = 19.32$) than the high-effort stimuli ($M = 58.00$ [51.55, 64.44], $SD = 19.33$). The CS × Response Effort interaction was not significant, $F(1, 36) = 2.40$, $p = .130$, $\eta_p^2 = .062$, $BF_{incl} = 0.260$.

Analysis of mean number of avoidance responses (binned per three trials) revealed a significant Cue × Effort interaction, $F(1, 3) = 159.06$, $p = .001$, $\eta_p^2 = .981$, $BF_{10} = 1.000$, with follow-up tests highlighting significantly greater responding on CS+ High-effort trials ($p < .001$).

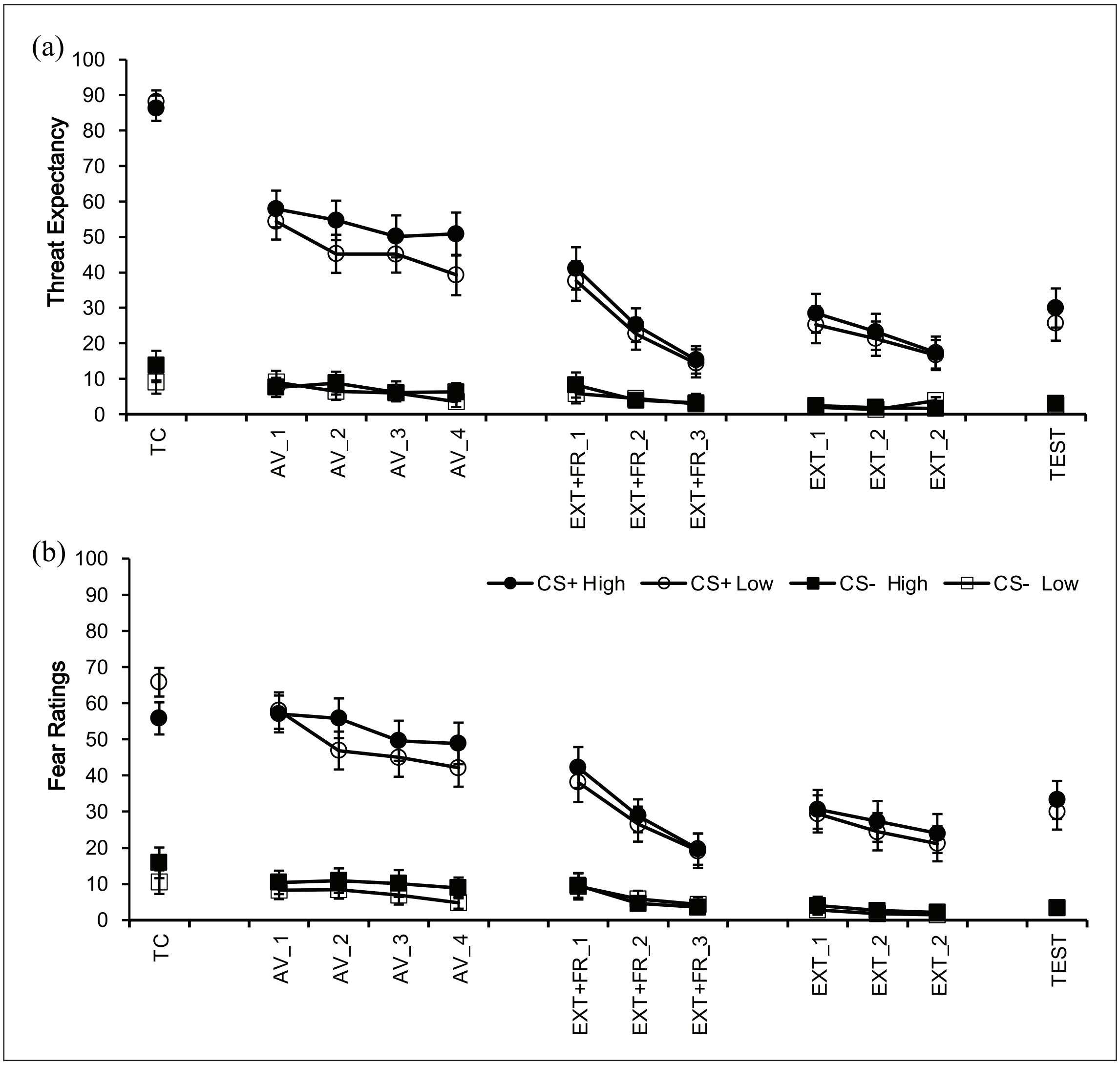

**Figure 6.** Threat expectancy (a) and fear (b) ratings in Experiment 2. TC refers to threat conditioning, AV refers to the avoidance conditioning, EXT + FR refers to the fear extinction and feedback removal, EXT refers to fear extinction, and TEST refers to the reinstatement test phase. Numbers represent blocks per phase. Error bars represent SEM.

*SCR.* The 2 (CS) × 2 (response effort) × 12 (trial) RM-ANOVA showed a significant main effect of trial, $F(5.50, 71.47)=2.41$, $p=.039$, $\eta_p^2=.157$, $\varepsilon=.500$, $BF_{10}=0.018$, with a general decrease in SCR amplitude across the *avoidance conditioning* phase. No further main effects or interactions were significant (all $ps>.110$).

### *Extinction and feedback removal*

*Threat expectancy.* During the *extinction and feedback removal* phase, the 2 (CS) × 2 (response effort) × 3 (trial block) RM-ANOVA found a significant main effect of CS, $F(1, 36)=32.49$, $p<.001$, $d=1.07$, $BF_{10}=4.816e+25$, with higher threat expectancy for the CS+ ($M=26.02$ [17.48, 34.56], $SD=25.60$) than the CS− ($M=4.77$ [0.89, 8.64], $SD=11.62$). Furthermore, there was a significant main effect of trial block, $F(1.53, 54.97)=18.39$, $p<.001$, $\eta_p^2=.338$, $\varepsilon=.763$, $BF_{10}=339{,}227.328$, with decreasing threat expectancy over the phase (Figure 6). There was also a significant CS × Block interaction, $F(1.62, 58.31)=14.11$, $p<.001$, $\eta_p^2=.282$, $\varepsilon=.810$, $BF_{incl}=0.057$. Tests of simple main effects revealed there were significant reductions in threat expectancy to the CS+ from Block 1 to Block 2 ($p=.001$) and from Block 2 to Block 3 ($p=.018$), with no significant changes in responding to the CS− ($ps>.434$). No further main effects or interactions were significant (all $ps>.089$).

*Fear ratings.* During the *extinction and feedback removal* phase, there was a significant main effect of CS, $F(1, 35)=32.68$, $p<.001$, $d=1.09$, $BF_{10}=2.560e{-}9$, and a significant CS × Response Effort interaction, $F(1, 35)=6.01$, $p=.019$, $\eta_p^2=.146$, $BF_{incl}=0.214$. Tests of simple main effects showed the high-effort CS+ was rated more fearfully than the low-effort CS+ ($p=.041$) with no significant difference between the high- and low-effort CS− ($p=.511$). ANOVA also revealed a significant main

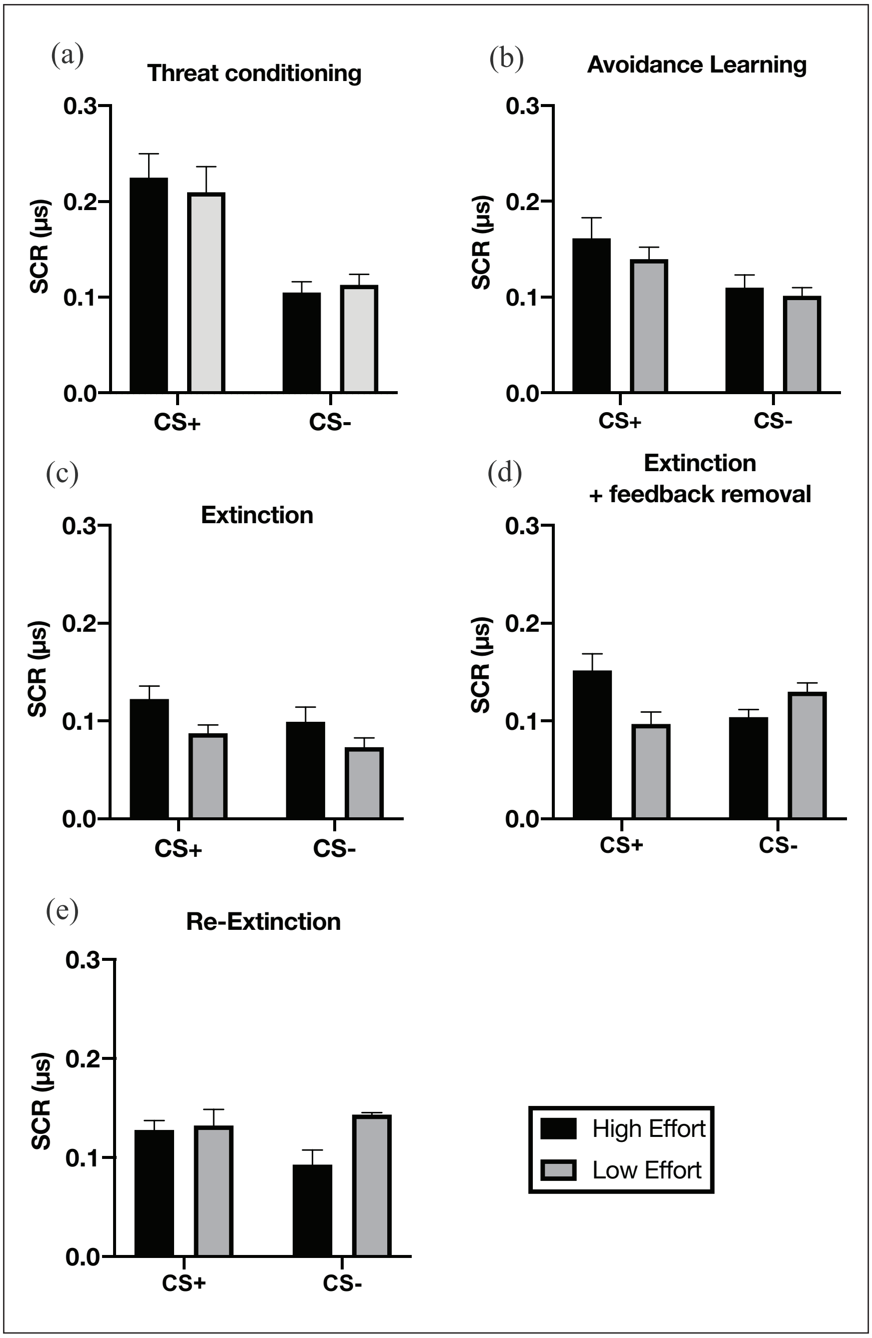

**Figure 7.** Skin conductance responses in Experiment 2. Error bars represent SEM.

effect of trial block, $F(1.44, 50.54)=19.35$, $p<.001$, $\eta_p^2=.356$, $\varepsilon=.722$, $BF_{10}=2.812e{-}35$, showing a reduction in fear ratings throughout the *extinction and feedback removal* phase irrespective of CS and response effort. No further main effects or interactions were significant ($ps>.304$).

*Proportion of avoidance.* During the *extinction and feedback removal* phase, there were significant main effects of CS, $F(1, 36)=23.35$, $p<.001$, $d=1.00$, $BF_{10}=0.011$, and response effort, $F(1, 36)=18.26$, $p<.001$, $d=0.64$, $BF_{10}=6.442e{-}8$; however, interpretation of these main effects was superseded by a significant CS × Response Effort interaction, $F(1, 36)=4.40$, $p=.043$, $\eta_p^2=.109$, $BF_{incl}=0.397$. Tests of simple main effects show that avoidance proportions were significantly higher for the low-effort CS+ than the high-effort CS+ ($M_{DIFF}=17.72$, $SEM=4.42$, $p<.001$), and this pattern was the same for the low- and high-effort CS− ($M_{DIFF}=9.61$, $SEM=2.89$, $p=.002$).

Analysis of mean number of avoidance responses (binned per three trials) revealed a main effect of CS, $F(1, 2)=3.000$, $p=.225$, $d=-0.091$, $BF_{10}=0.570$.

*SCR.* There were no significant main effects or interactions during the *extinction and feedback removal* phase ($p\text{s} > .088$).

### *Extinction*

*Threat expectancy.* Similar to the *extinction and feedback removal* phase, the *extinction* phase showed significant main effects of CS, $F(1, 36)=23.23$, $p<.001$, $d=0.99$, $\text{BF}_{10}=2.465\text{e}+30$, and trial block, $F(1.51, 54.28)=6.34$, $p=.007$, $\eta_p^2=.150$, $\varepsilon=.754$. These main effects were superseded by a significant CS × Block interaction, $F(1.59, 57.30)=9.82$, $p=.001$, $\eta_p^2=.214$, $\varepsilon=.796$, $\text{BF}_{\text{incl}}=0.057$. Tests of simple main effects showed that for the CS+ there was no significant change in threat expectancy from Block 1 to Block 2 ($p=.192$), but there was a significant reduction from Block 2 to Block 3 ($p=.007$). There were no significant changes in responding to the CS− ($p\text{s} > .689$). No further main effects or interactions were significant ($p=.161$).

*Fear ratings.* The 2 (CS) × 2 (response effort) × 3 (trial block) RM-ANOVA found a significant main effect of CS, $F(1, 35)=27.09$, $p<.001$, $d=1.09$, $\text{BF}_{10}=1.000$, with significantly higher fear ratings to the CS+ ($M=26.15$ [16.12, 36.19], $SD=29.66$) than the CS− ($M=2.48$ [−0.15, 5.12], $SD=7.78$). Furthermore, there were significantly higher fear ratings to the high-effort stimuli ($M=15.12$ [9.19, 21.05], $SD=17.53$) than the low-effort stimuli ($M=13.52$ [7.96, 19.08], $SD=16.43$), $F(1, 35)=5.14$, $p=.030$, $d=0.09$, $\text{BF}_{10}=5.766\text{e}-38$. Finally, there was also a significant main effect of trial block, $F(1.23, 43.05)=5.40$, $p=.019$, $\eta_p^2=.134$, $\varepsilon=.615$, $\text{BF}_{\text{incl}}=5.378\text{e}-38$, showing a general reduction in fear ratings throughout the *extinction* phase irrespective of CS type and response effort (see Figure 6).

*Proportion of avoidance.* During the *extinction* phase, the 2 (CS) × 2 (response effort) ANOVA revealed a significant main effect of CS, $F(1, 36)=52.90$, $p<.001$, $d=1.48$, $\text{BF}_{10}=1.000$, with significantly higher proportion of avoidance persisting for the CS+ ($M=65.92$ [51.01, 80.83], $SD=44.71$) compared with the CS− ($M=11.11$ [2.00, 20.23], $SD=27.34$). Furthermore, there was significantly higher proportion of avoidance for the low-effort stimuli ($M=40.09$ [30.18, 50.00], $SD=29.72$) than the high-effort stimuli ($M=36.94$ [27.30, 46.57], $SD=28.91$), $F(1, 36)=8.40$, $p=.006$, $d=0.11$, $\text{BF}_{10}=4.929\text{e}-21$. The CS × Response Effort interaction was not significant, $F(1, 36)=0.15$, $p=.702$, $\eta_p^2=.004$, $\text{BF}_{\text{incl}}=0.247$.

Analysis of mean number of avoidance responses (binned per three trials) revealed a significant Cue × Effort interaction, $F(1, 2)=38.368$, $p=.025$, $\eta_p^2=.174$, $\text{BF}_{10}=0.245$, with post hoc tests showing that responding only differed on CS+ High and CS− High ($p=.007$) trials (Figure 8).

*SCR.* There were no significant main effects or interactions during *extinction* ($p\text{s} > .148$).

### *Re-extinction*

*Threat expectancy.* The 2 (CS) × 2 (response effort) × 2 (time) RM-ANOVA found significant main effects of CS, $F(1, 36)=21.89$, $p<.001$, $d=0.98$, $\text{BF}_{10}=0.029$, and time, $F(1, 36)=9.00$, $p=.005$, $d=0.33$, $\text{BF}_{10}=3.328\text{e}-19$. These main effects were superseded by a significant CS × Time interaction, $F(1, 36)=12.18$, $p=.001$, $\eta_p^2=.253$, $\text{BF}_{\text{incl}}=$. Bonferroni-corrected test of simple main effects showed a significant increase in threat expectancy ratings from *extinction* to *re-extinction* for the CS+ ($p=.001$) but not the CS− ($p=.978$). In addition, there was a significant Response Effort × Time interaction, $F(1, 36)=4.32$, $p=.045$, $\eta_p^2=.107$, $\text{BF}_{\text{incl}}=6.503$, showing a significant threat expectancy increase following reinstatement for the high-effort ($p=.001$) but not the low-effort stimuli ($p=.069$). The CS × Response Effort interaction approached statistical significance, $F(1, 36)=4.02$, $p=.053$, $\eta_p^2=.100$, $\text{BF}_{10}=0.242$. No further main effects or interactions were significant ($p\text{s} > .334$).

*Fear ratings.* The 2 (CS) × 2 (response effort) × 2 (time) RM-ANOVA revealed significant main effects of CS, $F(1, 35)=30.33$, $p<.001$, $d=1.17$, $\text{BF}_{10}=0.228$, and time, $F(1, 35)=9.95$, $p=.003$, $d=0.32$, $\text{BF}_{10}=4.826\text{e}-26$. These main effects were superseded by a significant CS × Time interaction, $F(1, 35)=8.15$, $p=.007$, $\eta_p^2=.189$, $\text{BF}_{\text{incl}}=0.905$. Tests of simple main effects showed a significant increase in fear ratings following reinstatement for the CS+ ($p=.003$) but not the CS− ($p=.133$). There was also a significant main effect of response effort, $F(1, 35)=4.38$, $p=.044$, $d=0.10$, $\text{BF}_{10}=6.281\text{e}-27$, showing higher fear ratings for the high-effort stimuli ($M=15.68$ [9.96, 21.41], $SD=16.93$) compared with the low-effort stimuli ($M=14.04$ [8.72, 19.36], $SD=15.72$). No further interactions were significant ($p\text{s} > .126$).

*Proportion of avoidance.* The 2 (CS) × 2 (response effort) × 2 (time) RM-ANOVA revealed a significant main effect of response effort, $F(1, 36)=8.18$, $p=.007$, $d=0.08$, $\text{BF}_{10}=5.947\text{e}-44$, showing higher avoidance proportion for the low-effort ($M=39.19$ [29.35, 49.03], $SD=29.51$) than the high-effort stimuli ($M=36.71$ [26.94, 46.48], $SD=29.29$). While there was a significant main effect of CS, $F(1, 36)=49.68$, $p<.001$, $d=1.41$, $\text{BF}_{10}=1.000$, there was a significant CS × Time interaction, $F(1, 36)=4.23$, $p=.047$, $\eta_p^2=.105$, $\text{BF}_{\text{incl}}=0.239$. Bonferroni-corrected tests of simple main effects showed that there were significant differences between the CS+ and CS− both pre- and post-reinstatement ($p\text{s} < .001$); however, there was no significant increase in proportion of avoidance post-reinstatement for the CS+ ($p=.163$) or the CS− ($p=.422$). No further main effects or interactions were significant (all $p\text{s} > .214$).

One-way ANOVA of mean number of avoidance responses (trial-by-trial) revealed a significant main effect of cue, $F(3, 8)=6.499$, $p=.015$, $\eta_p^2=.709$, $\text{BF}_{10}=1.00$, with responding on CS+ High trials significantly greater than on all other trial types ($p\text{s} < 0.05$).

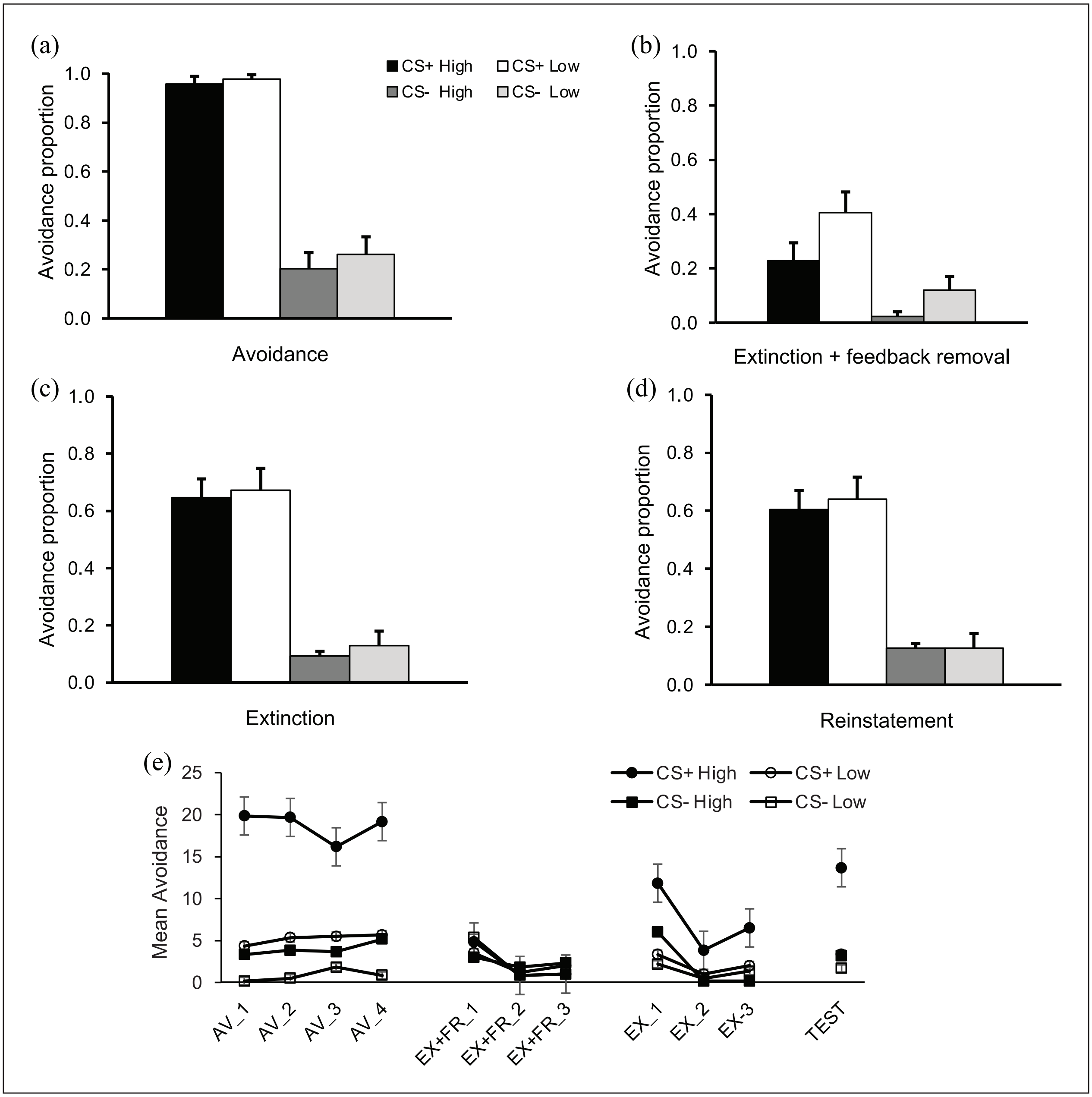

**Figure 8.** Proportion of avoidance responses and the mean number of responses per CS across all phases in Experiment 2. Panels a to d represent the following phases: avoidance, extinction and feedback removal, extinction, and reinstatement test (re-extinction), respectively. (e) Shows the total mean number of avoidance responses per phase. Successful avoidance was achieved by pressing the spacebar either 5 times (Low) or 20 times (High). Error bars represent SEM.

*SCR.* There were no significant main effects or interactions in assessing the effect of reinstatement during *re-extinction* (*p*s > .267).

## Discussion

In Experiment 2, we conducted a further within-subjects investigation with both the CS+ and CS− having low (FR-5) and high (FR-20) response effort requirements. It was hypothesised that following *extinction and feedback removal*, avoidance, threat expectancy, and fear of the *High* and *Low* CS+s would decrease. We again observed near-perfect rates of successful avoidance during *avoidance conditioning* for both CS+ High and Low trials, while maintaining low rates of avoidance in the presence of the two CS−s. Threat expectancy was higher for CS+ *High* than CS+ *Low*. Similarly, avoidance, expectancy and fear measures all decreased during *extinction and feedback removal*—again, like in Experiment 1, this may have been influenced by the perception that US-avoidance was ineffective given the absence of on-screen visual changes in the avoidance bar. Interestingly, fear and threat expectancy remained elevated for the CS+ *High* than all other CSs. During the *extinction test*, avoidance responses during low-effort trials, regardless of CS (i.e., both CS+ *Low* and CS− *Low*), remained significantly higher than high-effort trials, whereas avoidance responses made in the presence of CS+ trials tended to be significantly higher than those for CS− trials. Threat expectancy declined on CS+ trials, but both expectancy and fear remained higher on CS+ than CS− trials. During the *reinstatement test*, as expected, there were no differences in the proportion of avoidance between phases, while both fear and threat expectancy increased for the CS+s relative to the CS−.

Our secondary hypothesis was also supported. That is, we observed a return of elevated threat expectancy and fear with no effect on proportion of avoidance. Responding on low-effort trials remained higher than high-effort trials during reinstatement testing, indicating differential persistence and recovery of avoidance. The undifferentiated SCR data seen from *avoidance conditioning* perhaps reflected the complex experimental design involving multiple phases of acquisition, extinction, recovery, and re-extinction (Lonsdorf et al., 2017; Ney et al., 2020; Sjouwerman & Lonsdorf, 2020).

## General discussion

This study reports the findings of two laboratory-based analogue studies of the maintenance, reduction, and renewal of high and low response effort avoidance on FR-20 and FR-5 schedules of shock cancellation. Experiment 1 found non-differential persistence of avoidance and increased fear and threat expectancy during *re-extinction*, whereas Experiment 2 found an absence of discriminated avoidance responding and increased fear and threat expectancy during *reinstatement* testing. We will discuss each of these main findings in turn.

### *Response effort and persistent avoidance*

In Experiment 1, we expected the response effort manipulation would have greater impact on extinction of the High-effort relative to the Low-effort CS+ across all measures. In fact, the manipulation worked similarly across cues. Threat expectancy declined similarly for both cues, while the High-effort CSs each retained greater fear than the Low-effort CSs, and avoidance responding remained unchanged. The observed decline in threat expectancy is common in avoidance learning research (San Martín et al., 2020; Vervliet & Indekeu, 2015; Vervliet et al., 2017). However, the comparable levels of US-avoidance evoked by the CS+s more likely explain the non-differential decline in threat expectancy; ratings thus reflected the perceived likelihood of shock following High or Low CS+s. Despite threat expectancy being assessed at the end of trial blocks and not trial-by-trial (Boddez et al., 2013), participants still reliably learned of the modulatory effects of avoidance on US cancellation. It would be helpful in future research on avoidance response effort to employ trial-by-trial expectancy measures capable of fully differentiating the modulatory (presence/absence) effects of avoidance from other task-specific features such as our feedback removal phase.

Avoidance was comparable across the response effort manipulation in all phases except during the *extinction and feedback removal* phase, where the proportion of avoidance responding was lower for the High-effort CS+ than the Low-effort CS+ (Figure 4). Renewed levels of avoidance were seen during the brief, final *re-extinction* phase with non-discriminated responding across both CS+ types. The temporary and unpredicted change in the proportion of avoidance during *extinction and feedback removal* likely reflected the impact of the original schedule of reinforcement histories for high-effort avoidance. That is, in the absence of the on-screen feedback, and with all US deliveries withheld, high-effort FR-20 avoidance responding was briefly more responsive to extinction learning and decreased across trials. The number of CS+ high-effort responses was lower on these trials than during the previous *extinction* phase, which may indicate an immediate molar effect of the extinction and feedback removal contingencies on the early bouts of FR responding (i.e., within the first 10 or so trials) rather than molecular contingencies governing the FR schedule requirement (Baum, 2002; Vaughan & Miller, 1984). Cleary, this enhanced extinction procedure worked to reduce responding across extended time periods (i.e., experimental phases) and response sequences in an integrated, molar-like manner in much the same way as an ERP (Rattel et al., 2017). Of course, the effect was transient and likely partially controlled by the context provided by the removal and subsequent reinstatement of the on-screen response-produced feedback (Hineline, 1981). To that extent, the on-screen feedback may have exerted both a molecular and discriminative-like (i.e., response-prevention) function, promoting poor schedule control/sensitivity in its presence and greater sensitivity to extinction in its absence.

Our findings did reveal higher, sustained fear ratings for CSs requiring high compared with low response effort. This provides supportive evidence for a reinforcement schedule effect on the maintenance of self-reported fear and likely mirrors the range of avoidant behaviours performed by individuals with anxiety in coping with threat. The completion of the FR-20 response requirement for both CS+ and CS− (although response rates were predictably low and approaching zero for CS−) is analogous to the maladaptive behavioural routines of anxious coping whereby the feared event is prevented or avoided while high levels of fear remain intact. For instance, an individual with social anxiety disorder may fear social interactions and situations in the workplace with potential for scrutiny or negative judgement by others and hence engage in excessive avoidance to reduce threat (Penninx et al., 2021). The present findings indicate that performing multiple, high-rate avoidance responses like disabling phone notifications, deleting email invitations to social events with colleagues, avoiding shared social spaces (e.g., kitchens), working at unsocial hours or with one's office door closed may only serve to maintain the perceived fearfulness of social situations and interactions. Fear is likely to be sustained at greater levels following high rates of excessive avoidance like this compared with low rates of discrete avoidance responses (e.g., working from home), leading to obvious treatment implications (Dymond, 2019).

Elevated fear and threat expectancy for high-effort over low-effort cues has implications for understanding the complex, bidirectional relationship between fear and avoidance (Pittig et al., 2020). Maintained fear of the high-effort CS+ may have been the result of extended exposure to the cues during completion of the FR response requirements, with a resulting weaker association between the low-effort CS+ and fear. It is possible that the impact of cue exposure during the high and low FR schedules may interact with other mechanisms such as the reliability or effectiveness of avoidance at cancelling shock (Leng & Vervliet, 2022; Xia et al., 2017; Zuj et al., 2020). Here, low- or high-effort avoidance CS+s were always followed by shock cancellation when the response requirements were met, but the impact of partial avoidance schedules on fear levels remains to be determined.

The present FR response-effort paradigm differs considerably from existing studies of low-cost avoidance. These low-cost avoidance tasks usually require minimal cost or response effort (e.g., button-pressing once, on an FR1) and have been criticised for failing to address the complexity of real-world, maladaptive avoidance repertoires (Ball & Gunaydin, 2022; Krypotos et al., 2018). Costly avoidance paradigms, on the contrary, contrast the availability of avoidance behaviour with loss of opportunity to obtain rewards. Recent studies have contrasted shock probabilities with reward (usually monetary) magnitude (Wong & Pittig, 2020, 2021) and found enhanced rates of avoidance as reward magnitude increased in people high in trait anxiety (Pittig & Scherbaum, 2020). Such paradigms are useful analogues of clinically significant maladaptive avoidance, yet they have tended to adopt dichotomous measures (i.e., avoidance is either performed or it is not) with low response effort (e.g., FR1 response requirement) on all trials. The study of costly avoidance may benefit from incorporation of different FR schedules of shock cancellation and/or loss of monetary rewards. For instance, it is possible to conceptualise of a study where every response completing an FR schedule requirement simultaneously reduced an accumulating reward amount (i.e., response-cost punishment; Fontes & Shahan, 2020; Gandarela et al., 2020; Jean-Richard-Dit-Bressel et al., 2018). Indeed, research on costly avoidance generally superimposes a punishment contingency on the availability of avoidance responding. Early operant research findings on response suppression effects of punishment and its impact on avoidance may be salutary in that regard. Using an FR1 punishment schedule on one key and avoidance responses on a second key that could delay or prevent responses on the other key from being punished, a seminal study by Azrin et al. (1965) found that increased punishment intensity increased avoidance in pigeons. Notably, avoidance was maintained even though it reduced available reinforcement rates (i.e., it incurred greater costs) and preventing avoidance led to greater resistance to punishment. Although the relation between negative reinforcement and punishment is complex, further studies are needed of high-cost/response effort avoidance responding in humans informed by findings from the non-human operant literature.

## *Response effort and reinstatement of avoidance*

In Experiment 2, we found no differences during *reinstatement testing* for avoidance of the low or high response effort cues, whereas fear and threat expectancy increased for both CS+s relative to CS−. These findings add to the growing literature on the reinstatement of avoidance and extend them to a novel paradigm using FR response-effort avoidance extinction (Cameron et al., 2015; Dirikx et al., 2004; Norrholm et al., 2006; Urcelay et al., 2019; Zuj et al., 2018).

The return of elevated threat expectancy and fear with no effect on avoidance that we observed was likely to have arisen, at least in part, due to the sequence of extinction procedures employed. That is, the *extinction and feedback removal phase* occurred before *extinction* and *re-extinction*, and it is possible that the combined absence of on-screen FR response-contingent feedback and US deliveries may have exerted a facilitative effect on subsequent reinstatement. There has been surprisingly little empirical work conducted on reinstatement in humans and many of the boundary conditions underlying the return and reinstatement of fear and avoidance remain to be delineated (Sjouwerman & Lonsdorf, 2020). Our findings suggest that the *extinction and feedback removal phase* was effective at extinguishing the previously acquired avoidance response effort, which was further undermined and extinguished during *extinction* prior to the crucial reinstatement testing. In this way, a potential boundary condition for the reinstatement of avoidance may be the confirmed extinction of responding by faded stimulus control (i.e., removal of on-screen feedback) and withholding of US deliveries. While the number of unsignalled reinstatement USs used are unlikely to have made a difference (Haaker et al., 2014; Sjouwerman & Lonsdorf, 2020), further empirical research is needed on both sequence and format of the procedural parameters necessary for the reinstatement of avoidance.

### Maladaptive and extinction-resistant response-prevention avoidance

Ball and Gunaydin (2022) outlined a neurobehavioural model of maladaptive avoidance based on three central features: heightened threat appraisal, habitual avoidance, and trait avoidance tendency. Maladaptive avoidance refers to avoidance of safe stimuli, usually with costs incurred by the individual for so doing. How might we interpret the present findings in terms of maladaptive avoidance? According to Ball and Gunaydin, heightened threat appraisal involves a tendency to overestimate threat from stimuli mistakenly perceived as fearful and potentially dangerous and likely promotes maladaptive avoidance through high levels of fear. This feature is consistent with classic, two-factor theories of avoidance positing that ongoing avoidance of safe stimuli is motivated and maintained by fear reduction (Dymond & Roche, 2009), but fails to accommodate both extensions of such traditional accounts (e.g., Lovibond, 2006) and instances where fear and avoidance are decoupled (Pittig et al., 2020; Wong et al., 2022). Habitual avoidance refers to repeated instances of avoidance behaviour seemingly divorced from its outcomes or consequences and unrelated to perceived levels of fear. Accounts of habitual avoidance have emphasised the non-goal-directed basis of these behaviours that rapidly become a default means of coping with uncertainty (Gillan & Robbins, 2014). Finally, trait avoidance tendency refers to individual differences resulting in an increased likelihood of maladaptive avoidance. Again, here the trait avoidance tendency overrides level of fear and promotes a “better safe than sorry” form of coping across dissimilar situations (Lommen et al., 2010).

Let us address the implications and contributions of the present experiments on extinction-resistant avoidance with response prevention considering these features. In Experiment 1, we found extinction-resistant high- and low-effort avoidance with sustained levels of fear and threat expectancy. This indicates a non-differential impact on maladaptive avoidance of heightened threat appraisal. The observation of consistently elevated fear levels for CSs requiring high compared with low response effort does however indicate that the interaction between avoidance response effort and maladaptive avoidance was maintained, at least in part, by heightened threat appraisal. High-effort avoidance may have protected participants’ responding from extinction or fostered emergence of an early form of habitual avoidance.

In Experiment 2, we found some evidence of the reinstatement of avoidance responding but did observe increased levels of fear and threat expectancy. These findings highlight, for the first time, reinstatement in an extinction-resistant avoidance with response-prevention paradigm. However, the levels of avoidance we found suggest a differential impact of reinstatement on the different outcome measures employed. Again, it is difficult to interpret the findings of Experiment 2 in terms of habitual avoidance given that avoidance did not differ across danger and safety cues during the brief *re-extinction* test phase. Similarly, the role of trait avoidance measures in generating the observed reinstatement effects remains unclear (Haaker et al., 2014; Sjouwerman & Lonsdorf, 2020; Sjouwerman et al., 2020).

The present paradigm has potential for future research on maladaptive avoidance and testing the assumptions of neurobehavioural models (Ball & Gunaydin, 2022). For instance, the relative independence of the three central features of heightened threat appraisal, habitual avoidance, and trait avoidance tendency may be investigated in an extinction-resistant avoidance with response-prevention paradigm where avoidance response effort is varied at the outset. Investigating maladaptive avoidance in this way also allows for the inclusion of avoidance costs in a parametric analysis of negative reinforcement and punishment schedules (Dymond, 2019; Jean-Richard-Dit-Bressel et al., 2018). By extending or repeating test phases and overtraining avoidance, it is possible to facilitate analyses of the potential decoupling of both performance measures and control by goal- or non-goal-based habitual mechanisms (Pool et al., 2022). The present paradigm also has potential for use with functional neuroimaging methods and in so doing revealing more of the neural structures underlying maladaptive avoidance (Chase et al., 2020; Schlund et al., 2016).

### US-avoidance and the validity of human avoidance paradigms

The present paradigm was developed to study *US-avoidance* where completion of the FR response requirement in the presence of CS+ prevents the scheduled occurrence of the shock US (Krypotos et al., 2018; Wong et al., 2022). In this way, US-avoidance is distinct from *CS-avoidance* which involves responding that prevents the occurrence of both the CS+ itself and the US. Thus, it is possible to distinguish between US- and CS-avoidance by whether it results in simultaneous CS termination; in the former, US-avoidance does not impact on CS presentation, whereas in the latter, CS-avoidance directly influences the extent of exposure to the CS. Our approach here, of incorporating FR-based response effort within a US-avoidance paradigm, is however open to alternative interpretation. For instance, avoidance conditioning trials began with presentation of the CS for a brief period before presentation of the “avoidance bar”; successful avoidance involved completing the ratio-based response requirement which produced on-screen, incremental changes in the avoidance bar. Filling up the bar prevented occurrence of the US and simultaneous CS+ offset. If the response requirement was not met, and the avoidance bar was incomplete, the US was presented at CS+ offset. In this way, both the simultaneous removal of CS+ and US cancellation on FR-avoidance trials we employed may better resemble a form of CS-escape procedure.

Avoidance in our paradigm is therefore likely to have been maintained by the combined negative reinforcement processes of US omission and reduction in the fear-eliciting properties of the CS (Dymond & Roche, 2009; Wong et al., 2022). Identifying the relative contributions of these potentially separable behavioural mechanisms warrants further investigation in laboratory-based treatment studies, as does the general interplay between escape and avoidance in human learning (Haskell et al., 2019). The inclusion of costs or similar FR response effort manipulations in studies of avoidance with humans increases face validity and may also benefit assessments of diagnostic validity aimed at identifying clinical markers of vulnerable individuals (Dymond, 2019; Krypotos et al., 2018; Wong et al., 2022). Clinically anxious participants may show excessive acquisition of avoidance with the present procedures and have greater difficulty subsequently overcoming extinction-resistant high-rate responding. This possibility warrants empirical examination, as are the development of novel treatments aimed at disrupting or extinguishing high-rate maladaptive avoidance. Indeed, the treatment of low-rate and high-rate avoidance may require different approaches that target the combined impact of the original acquisition schedule and the client's previous, unsuccessful treatment attempts.

### *Limitations*

This study has several limitations with implications for further research. First, we failed to find significant evidence for SCR measures which may have been due to the lengthy experimental session and multiple cue presentation format. Further research on the role of response effort in avoidance extinction should incorporate alternative physiology measures such as fear-potentiated startle, respiration rate, and heart rate variability, perhaps in shorter sessions (or multi-day studies). Second, the experiments employed a predetermined learning criterion of a minimum 80% response rate (i.e., avoidance responding on 10 out of 12 trials) for at least one of the CS+s during *avoidance conditioning*. All participants met criterion in Experiment 1, while three failed to do so in Experiment 2. We chose to implement a relatively conservative criterion to ensure minimal contact with the FR-avoidance schedules at cancelling upcoming shock in the presence of either CS+. Of course, the fact that the criterion applied *to one but not both* CS+s may have resulted in partial schedule control and warrants further investigation, such as in a future study with criteria applied to all cues discriminative for avoidance and non-avoidance responding, respectively (e.g., Dymond et al., 2007). Third, the translational validity of the response effort requirements may be lacking in that the FR schedules were applied to responding in the presence of both danger and safety cues. In the real world, active avoidance responding combined with passive non-avoidance contingencies may be more likely (Cornwell et al., 2013). For instance, the socially anxious individual may take steps to avoid fearful social interactions or people yet withhold responding or do nothing in the presence of unfearful people or situations. In this way, there is greater response effort involved in actively avoiding feared situations than in approaching safety, which can have important treatment implications. Further research should develop targeted clinical treatments for extensions to vulnerable populations. Fourth, the crucial extinction test phases consisted of relatively few trials which may have been inadequate to detect stable patterns of extinguished responding. Finally, this study did not administer trait measures as potential predictors of behavioural performance.

## Conclusion

The present findings illustrate, for the first time, the impact of high and low FR response effort on the extinction and return of avoidance. Results demonstrate the effectiveness of an empirical paradigm capable of further adaptations to better understand the role of reinforcement schedules in the persistence of maladaptive avoidance.

### 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) received no financial support for the research, authorship, and/or publication of this article.

### ORCID iD

Simon Dymond https://orcid.org/0000-0003-1319-4492

### References

Ailing, K., & Poling, A. (1995). The effects of differing response-force requirements on fixed-ratio responding of rats. *Journal of the Experimental Analysis of Behavior*, *63*, 331–346. https://doi.org/10.1901/jeab.1995.63-331

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM 5*. American Psychiatric Publishing.

Armus, H. L. (1988). Effect of response effort requirement on relative frequency of short interresponse times: CRF and FR-5 reinforcement schedules. *Bulletin of the Psychonomic Society*, *26*, 139–140. https://doi.org/10.3758/BF03334886

Augustson, E. M., & Dougher, M. J. (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *Journal of Behavior Therapy and Experimental Psychiatry*, *3*, 181–191. https://doi.org/10.1016/S0005-7916(97)00008-6

Azrin, N. H., Hake, D. F., Holz, W. C., & Hutchinson, R. R. (1965). Motivational aspects of escape from punishment. *Journal of the Experimental Analysis of Behavior*, *8*, 31–44. https://doi.org/10.1901/jeab.1965.8-31

Ball, T. M., & Gunaydin, L. A. (2022). Measuring maladaptive avoidance: From animal models to clinical anxiety. *Neuropsychopharmacology*, *47*, 978–986. https://doi.org/10.1038/s41386-021-01263-4

Baum, W. M. (2002). From molecular to molar: A paradigm shift in behavior analysis. *Journal of the Experimental Analysis of Behavior*, *78*, 95–116. https://doi.org/10.1901/jeab.2002.78-95

Bennett, M., Roche, B., Baeyens, F., Vervliet, B., Dymond, S., Whelan, R., & Hermans, D. (2020). Transitions from avoidance: Reinforcing competing behaviours reduces generalized avoidance in new contexts. *Quarterly Journal of Experimental Psychology*, *73*, 2119–2131. https://doi.org/10.1177/1747021820943148

Blough, D. S. (1966). The study of animal sensory processes by operant methods. In W K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 345–379). Appleton-Century-Crofts.

Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: Validity of US expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, *44*, 201–206. https://doi.org/10.1016/j.jbtep.2012.08.003

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*, 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9

Cameron, G., Schlund, M. W., & Dymond, S. (2015). Generalization of socially transmitted and instructed avoidance. *Frontiers in Behavioral Neuroscience, 9*, Article 159. https://doi.org/10.3389/fnbeh.2015.00159

Catania, A. C. (1992). *Learning* (3rd ed.). Prentice-Hall.

Chase, H. W., Graur, S., Versace, A., Greenberg, T., Bonar, L., Hudak, R., Quirk, G. J., Greenberg, B. D., Rasmussen, S. A., Haber, S. N., & Phillips, M. L. (2020). Neural mechanisms of persistent avoidance in OCD: A novel avoidance devaluation study. *Neuroimage Clinical, 28*, Article 102404. https://doi.org/10.1016/j.nicl.2020.102404

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Lawrence Erlbaum.

Cornwell, B. R., Overstreet, C., Krimsky, M., & Crillon, C. (2013). Passive avoidance is linked to impaired fear extinction in humans. *Learning & Memory*, *20*, 164–169. https://doi.org/10.1101/lm.028902.112

Dawson, M. E., Rissling, A. J., Schell, A. M., & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion*, *7*(4), 755–766. https://doi.org/10.1037/1528-3542.7.4.755

Dirikx, T., Hermans, D., Vansteenwegen, D., Baeyens, F., & Eelen, P. (2004). Reinstatement of extinguished conditioned responses and negative stimulus valence as a pathway to return of fear in humans. *Learning & Memory*, *11*, 549–554. https://doi.org/10.1101/lm.78004

Dymond, S. (2019). Overcoming avoidance in anxiety disorders: The contributions of Pavlovian and operant avoidance extinction methods. *Neuroscience and Biobehavioral Reviews*, *98*, 61–70. https://doi.org/10.1016/j.neubiorev.2019.01.007

Dymond, S., & Roche, B. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Anayst*, *32*, 7–28. https://doi.org/10.1007/BF03392173

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior*, *88*, 249–262.

Falligant, J. M., Rapp, J. T., Brogam, K. M., & Pinkston, J. W. (2020). Extinction and response force in conjugate schedules. *The Psychological Record*, *70*, 109–121. https://doi.org/10.1007/s40732-020-00374-6

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146

Fernando, A. B. P., Urcelay, G. B., Mar, A. C., Dickinson, A., & Robbins, T. W. (2014). Safety signals as instrumental reinforcers during free-operant avoidance. *Learning & Memory*, *21*, 488–497. https://doi.org/10.1101/lm.034603.114

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Croft.

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 & Therapy*, *104*, 34–43. https://doi.org/10.1016/j.brat.2018.02.008

Fontes, R. M., & Shahan, T. A. (2020). Punishment and its putative fallout: A reappraisal. *Journal of the Experimental Analysis of Behavior*, *115*, 185–203. https://doi.org/10.1002/jeab.653

Friman, P. C., & Poling, A. (1995). Making life easier with effort: Basic findings and applied research on response effort. *Journal of Applied Behavior Analysis*, *28*, 583–590. https://doi.org/10.1901/jaba.tive reinforcement rate and persiste1995.28-583

Gandarela, L., Boldrin, L. S., & Debert, P. (2020). Transfer of the avoidance function in equivalence classes using loss of points as the aversive stimulus. *Psychological Record*, *70*, 471–479. https://doi.org/10.1007/s40732-019-00365-2

Gillan, C. M., & Robbins, T. W. (2014). Goal-directed learning and obsessive–compulsive disorder. *Philosophical Transactions of the Royal Society: B, 369*, Article 20130475. https://doi.org/10.1098/rstb.2013.0475

Gunnarsson, K. F., Rowsey, K. E., & Dixon, M. R. (2015). Increasing response effort impacts wager sizes of slot machine gamblers. *Analysis of Gambling Behavior*, *9*, 19–25.

Haaker, J., Golkar, A., Hermans, D., & Lonsdorf, T. B. (2014). A review on human reinstatement studies: An overview and methodological challenges. *Learning & Memory*, *21*, 424–440. https://doi.org/10.1101/lm.036053.114

Haskell, A. M., Britton, P. C., & Servatius, R. J. (2019). Toward an assessment of escape/avoidance coping in depression. *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2019.112363

Herrnstein, R. J., & Morse, W. H. (1958). A conjunctive schedule of reinforcement. *Journal of the Experimental Analysis of Behavior*, *1*, 15–24. https://doi.org/10.1901/jeab.1958.1-15

Hineline, P. N. (1981). The several roles of stimuli in negative reinforcement. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behaviour: Predictability, correlation, and contiguity* (*Vol. 2*., pp. 203–246). John Wiley & Sons.

JASP Team. (2020). *JASP (Version 0.13.1)* [Computer software].

Jean-Richard-Dit-Bressel, P., Killcross, S., & McNally, G. P. (2018). Behavioral and neurobiological mechanisms of punishment: Implications for psychiatric disorders. *Neuropsychopharmacology*, *43*, 1639–1650. https://doi.org/10.1038/s41386-018-0047-3

Katz, J. L., & Barrett, J. E. (1979). Conjunctive schedules of reinforcement IV: Effects on the pattern of responding of changes in requirement at reinforcement. *Animal Learning & Behavior*, *7*, 483–488. https://doi.org/10.3758/BF03209706

Keenan, M., & Leslie, J. (1984). Separating response dependence and response-reinforcer contiguity within a recycling conjunctive schedule. *Journal of the Experimental Analysis of Behavior*, *41*, 203–210. https://doi.org/10.1901/jeab.1984.41-203

Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: Erasing human fear responses and preventing the return of fear. *Nature Neuroscience*, *12*, 256–258. https://doi.org/10.1038/nn.2271

Krypotos, A.-M., & Engelhard, I. M. (2018). Testing a novelty-based extinction procedure for the reduction of conditioned avoidance. *Journal of Behavior Therapy & Experimental Psychiatry*, *60*, 22–28. https://doi.org/10.1016/j.jbtep.2018.02.006

Krypotos, A.-M., Vervliet, B., & Engelhard, I. M. (2018). The validity of human avoidance paradigms. *Behaviour Research & Therapy*, *111*, 99–105. https://doi.org/10.1016/j.brat.2018.10.011

LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2016). The birth, death and resurrection of avoidance: A reconceptualization of a troubled paradigm. *Molecular Psychiatry*, *22*, 24–36. https://doi.org/10.1038/mp.2016.166

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Leng, L., & Vervliet, B. (2022). More engagement in inefficient avoidance through partial reinforcement. *Journal of Behavior Therapy and Experimental Psychiatry, 76*, Article 101751. https://doi.org/10.1016/j.jbtep.2022.101751

Lommen, M. J. J., Engelhard, I. M., & van den Hout, M. A. (2010). Neuroticism and avoidance of ambiguous stimuli: Better safe than sorry? *Personality and Individual Differences*, *49*, 1001–1006. https://doi.org/10.1016/j.paid.2010.08.012

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., & Merz, C. J. (2017). Don't fear "fear conditioning": Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Lovibond, P. F. (2006). Fear and avoidance: An integrated expectancy model. In M. Craske, D. Hermans & D. Vansteenwegen (Eds.), *Fear and learning* (pp. 117–132). American Psychological Association.

Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A., & Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research & Therapy*, *47*, 716–720. https://doi.org/10.1016/j.brat.2009.04.013

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

Marin, M. F., Barbey, F., Rosenbaum, B. L., Hammoud, M. Z., Orr, S. P., & Milad, M. R. (2019). Absence of conditioned responding in humans: A bad measure or individual differences? *Psychophysiology*, *57*, 1–10. https://doi.org/10.1111/psyp.13350

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324.

Meulders, A., Franssen, M., Fonteyne, R., & Vlaeyen, J. W. (2016). Acquisition and extinction of operant pain-related avoidance behavior using a 3 degrees-of-freedom robotic arm. *Pain*, *157*, 1094–1104. https://doi.org/10.1097/j.pain.0000000000000483

Ney, L. J., Laing, P. A. F., Steward, T., Zuj, D. V., Dymond, S., & Felmingham, K. L. (2020). Inconsistent analytic strategies reduce robustness in fear extinction via skin conductance response. *Psychophysiology, 57*, Article e13650. https://doi.org/10.1111/psyp.13650

Norrholm, S. D., Jovanovic, T., Vervliet, B., Myers, K. M., Davis, M., Rothbaum, B. O., & Duncan, E. J. (2006). Conditioned fear extinction and reinstatement in a human fear-potentiated startle paradigm. *Learning & Memory*, *13*, 681–685. https://doi.org/10.1101/lm.393906

Penninx, B. W. J. H., Pine, D. S., Holmes, E. A., & Reif, A. (2021). Anxiety disorders. *The Lancet*, *397*, 914–927. https://doi.org/10.1016/S0140-6736(21)00359-7

Pinkston, J. W., & McBee, L. N. (2014). Force dynamics in fixed-ratio schedules. *Behavioural Processes*, *103*, 112–116. https://doi.org/10.1016/j.beproc.2013.11.00

Pittig, A., & Scherbaum, S. (2020). Costly avoidance in anxious individuals: Elevated threat avoidance in anxious individuals under high, but not low competing rewards. *Journal of Behavior Therapy and Experimental Psychiatry*, *66*, Article 101524.

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

Pool, E. R., Gera, R., Fransen, A., Perez, O. D., Cremer, A., Aleksic, M., Tanwisuth, S., Quail, S., Ceceli, A. O., Manfredi, D. A., Nave, G., Tricomi, E., Balleine, B., Schonberg, T., Schwabe, L., & O'Doherty, J. P. (2022). Determining the effects of training duration on the behavioral expression of habitual control in humans: A multilaboratory investigation. *Learning & Memory*, *29*, 16–28. https://doi.org/10.1101/lm.053413.12

Rattel, J. A., Miedl, S. F., Blechert, J., & Wilhelm, F. H. (2017). Higher threat avoidance costs reduce avoidance behaviour which in turn promotes fear extinction in humans. *Behaviour Research and Therapy*, *96*, 37–46. https://doi.org/10.1016/j.brat.2016.12.010

San Martín, C., Jacobs, B., & Vervliet, B. (2020). Further characterization of relief dynamics in the conditioning and generalization of avoidance: Effects of distress tolerance and intolerance of uncertainty. *Behaviour Research and Therapy, 124*, Article 103526. https://doi.org/10.1016/j.brat.2019.103526

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*, 94–105. https://doi.org/10.1016/j.neuroimage.2016.04.070

Shull, R. L., & Lawrence, P. S. (1998). Reinforcement. In K. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 95–129). Springer.

Sjouwerman, R., & Lonsdorf, T. B. (2020). Experimental boundary conditions of reinstatement-induced return of fear in humans: Is reinstatement in humans what we think it is? *Psychophysiology, 57*, e13549. https://doi.org/10.1111/psyp.13549

Sjouwerman, R., Scharfenort, R., & Lonsdorf, T. B. (2020). Individual differences in fear acquisition: Multivariate analyses of different emotional negativity scales, physiological responding, subjective measures, and neural activation. *Scientific Reports, 10*, Article 15283. https://doi.org/10.1038/s41598-020-72007-5

Soeter, M., & Kindt, M. (2010). Dissociating response systems: Erasing fear from memory. *Neurobiology of Learning and Memory*, *94*, 30–41. https://doi.org/10.1016/j.nlm.2010.03.004

Urcelay, G. P., Symmons, K., & Prével, A. (2019, August 28). *Renewal of instrumental avoidance in humans*. https://doi.org/10.31234/osf.io/2nxkh

van Vliet, C. M., Meulders, A., Vancleef, L. M. G., Meyers, E., & Vlaeyen, J. Y. S. (2019). Changes in pain-related fear and pain when avoidance behaviour is no longer effective. *Journal of Pain, 21*, 494–505. doi: https://doi.org/10.1016/j.jpain.2019.09.002

Vaughan, W., & Miller, H. L. (1984). Optimization versus response strength accounts of behavior. *Journal of the Experimental Analysis of Behavior*, *42*, 337–348. https://doi.org/10.1901/jeab.1984.42-337

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology*, *9*, 215–248. https://doi.org/10.1146/annurev-clinpsy-050212-185542

Vervliet, B., & Indekeu, E. (2015). Low-cost avoidance behaviors are resistant to fear extinction in humans. *Frontiers in Behavioral Neuroscience, 9*, Article 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*, 66–78. https://doi.org/10.1016/j.brat.2017.04.011

Volders, S., Meulders, A., De Peuter, S., Vervliet, B., & Vlaeyen, J. W. (2012). Safety behavior can hamper the extinction of fear of movement-related pain: An experimental investigation in healthy participants. *Behaviour Research and Therapy*, *50*(11), 735–746. https://doi.org/10.1016/j.brat.2012.06.004

Weisman, R. G., & Litner, J. S. (1972). *The role of Pavlovian events in avoidance training: Inhibition and learning*. Academic Press.

Wong, A. H. K., & Pittig, A. (2021). A dimensional measure of safety behavior: A nondichotomous assessment of costly avoidance in human fear conditioning. *Psychological Research*, *86*, 312–330. https://doi.org/10.1007/s00426-021-01490-w

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*, Article 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*, 79–89. https://doi.org/10.1016/j.brat.2017.04.002

Zuj, D. V., & Norrholm, S. D. (2019). The clinical applications and practical relevance of human conditioning paradigms for posttraumatic stress disorder. *Progress in Neuropsychopharmacology & Biological Psychiatry*, *88*, 339–351. https://doi.org/10.1016/j.pnpbp.2018.08.0

Zuj, D. V., Palmer, M. A., Malhi, G. S., Bryant, R. A., & Felmingham, K. L. (2018). Greater sleep disturbance and longer sleep onset latency facilitate SCR-specific fear reinstatement in PTSD. *Behaviour Research and Therapy*, *110*, 1–10. https://doi.org/10.1016/j.brat.2018.08.005

Zuj, D. V., Xia, W., Lloyd, K., Vervliet, B., & Dymond, S. (2020). Negative reinforcement rate and persistence of avoidance following response-prevention extinction. *Behaviour Research and Therapy*, *133*, Article 1037111. https://doi.org/10.1016/j.brat.2020.103711