University of Reading

Intolerance of uncertainty, and not social anxiety, is associated with compromised extinction of social threat

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Wake, S., Morriss, J., Johnstone, T., Van Reekum, C. M. ORCID: https://orcid.org/0000-0002-1516-1101 and Dodd, H. ORCID: https://orcid.org/0000-0003-1446-5338 (2021) Intolerance of uncertainty, and not social anxiety, is associated with compromised extinction of social threat. Behaviour Research and Therapy, 139. 103818. ISSN 0005-7967 doi: 10.1016/j.brat.2021.103818 Available at https://centaur.reading.ac.uk/96483/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1016/j.brat.2021.103818

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

REVISED MANUSCRIPT: Intolerance of uncertainty, and not social anxiety, is associated with compromised extinction of social threat

Shannon Wake* [1], Jayne Morriss [1], Tom Johnstone [1,2], Carien M. van Reekum [1] & Helen Dodd* [1]

[1] School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK

[2] Swinburne University of Technology, John Street, Hawthorn, Victoria 3122, Australia

* Correspondence:
Shannon Wake
School of Psychology and Clinical Language Sciences
University of Reading
Earley Gate, Whiteknights Campus
RG6 6AH Reading
United Kingdom
s.j.wake@pgr.reading.ac.uk

Helen Dodd
School of Psychology and Clinical Language Sciences
University of Reading
Earley Gate, Whiteknights Campus
RG6 6AH Reading
United Kingdom
h.f.dodd@reading.ac.uk

Abstract

Extinction-resistant threat is regarded as a central hallmark of pathological anxiety. However, it remains relatively under-studied in social anxiety. Here we sought to determine whether self-reported trait social anxiety is associated with compromised threat extinction learning and retention. We tested this hypothesis within two separate, socially relevant conditioning studies. In the first experiment, a Selective Extinction Through Cognitive Evaluation (SECE) paradigm was used, which included a cognitive component during the extinction phase, while experiment 2 used a traditional threat extinction paradigm. Skin conductance responses and subjective ratings of anxiety (experiment 1 and 2) and expectancy (experiment 2) were collected across both experiments. The findings of both studies demonstrated no effect of social anxiety on extinction learning or retention. Instead, results from experiment 1 indicated that individual differences in Intolerance of Uncertainty (IU) were associated with the ability to use contextual cues to decrease a conditioned response during SECE. However, during extinction retention, high IU predicted greater generalisation across context cues. Findings of experiment 2 revealed that higher IU was associated with impaired extinction learning and retention. The results from both studies suggest that compromised threat extinction is likely to be a characteristic of high levels of IU and not social anxiety.

Keywords: Threat Acquisition, Extinction Learning, Extinction Retention, Social Anxiety, Intolerance of Uncertainty, Skin Conductance

1. General Introduction

Social anxiety disorder (SAD) is characterised by the persistent and intense fear of social or performance situations in which the individual is exposed to potential scrutiny and negative evaluation from others (American Psychiatric Association, [APA], 2013). Fear in social anxiety is often expressed as severe distress in social contexts and the avoidance of situations in which there is a perceived social threat (APA, 2013). Previous models have noted that social anxiety exists along a severity continuum, with many individuals experiencing severe symptoms without meeting the threshold for clinical diagnosis of SAD (Bögels et al., 2010; Spence & Rapee, 2016; Stein et al., 2000). The UK's National Institute for Health and Care Excellence (NICE) guidelines indicate that graduated exposure can be used in the treatment of SAD. Although exposure is an effective and evidence-based treatment for SAD (Jorstad-Stein & Heimberg, 2009; Ponniah & Hollon, 2008), relapse after exposure therapy is common (Craske et al., 2014; Graham & Milad, 2011; Hofmann & Smits, 2008), and can affect as many as 33-50% of successfully treated individuals (Craske & Rachman, 1987; Smith et al., 2009).

Principles of threat extinction serve as a model for exposure therapy (Dunsmoor et al., 2015; Foa et al., 1989; Milad & Quirk, 2012). Classical conditioning occurs when a previously neutral stimulus (conditioned stimulus, CS+) is associated with an aversive stimulus (unconditioned stimulus, US) so that the CS+ acquires the ability to elicit a defensive response when presented alone (conditioned response, CR) (Pavlov, 1927). The reduction of the conditioned response is typically achieved through threat extinction. During threat extinction, the CS+ is repeatedly presented in the absence of the US, leading to a decline in the conditioned response, as the CS+ loses its predictive value concerning the US. Threat extinction

does not erase the learned threat association; instead, it involves new safety learning which inhibits the expression of the original threat association (Bouton, 1993; Milad & Quirk, 2012). Exposure-based therapies use the threat extinction principle to oppose flawed associations between intrinsically safe situations (CS) and imagined dangerous outcomes (US), by repeatedly exposing the patient to the objects or situations that generate fear (Vervliet et al., 2013).

Across threat conditioning and extinction research, anxiety is generally defined broadly, and studies have included dimensional assessments of individual differences in trait anxiety as well as comparisons across clinical and non-clinical samples. This broad definition of anxiety is consistent with The National Institute of Mental Health Research Domain Criteria (RDoc; National Institute of Mental Health, 2011) and theory regarding the relationship between trait anxiety and anxiety disorders (e.g., Raymond et al., 2017). The threat extinction literature has shown that for individuals with elevated levels of anxiety, affective responses during extinction remain elevated and sustained to cues that no longer signal threat, suggesting impaired threat extinction (See Lonsdorf & Merz, 2017, for a recent review). As relapse after exposure therapy is common in SAD, further research examining the circumstances under which extinction learning and retention is compromised in socially anxious individuals is of value to inform clinical avenues aimed at improving the efficacy of exposure-based treatments for social anxiety.

Previous work has reported that social anxiety is associated with compromised extinction of the conditioned response, however, there are discrepancies in findings across the literature. A study that measured US expectancy ratings found that patients with SAD did not show evidence of successful extinction of the CS+ and US contingency (Rabinak et al., 2016). Two studies have further

demonstrated that elevated levels of social anxiety are associated with delayed extinction of the conditioned response, indicated by a maintained differential skin conductance response throughout the extinction phase, in both patients with SAD (Hermann et al., 2002), and participants with elevated self-reported levels of rejection sensitivity (Olsson et al., 2013). In contrast, several studies have failed to find a relationship between elevated levels of social anxiety and differences in threat extinction processes across either psychophysiological or subjective measures, in both clinical and analogue samples (Lissek et al., 2008; Reichenberger et al., 2017; Tinoco-Gonzalez et al., 2015).

One explanation for this discrepancy across findings in the literature may be that the aforementioned studies did not investigate whether the reported effects were related to social anxiety specifically and not the result of more general or transdiagnostic processes that underpin anxiety more broadly (i.e., trait anxiety and intolerance of uncertainty). This is an important limitation because it is well-established that social anxiety is related to transdiagnostic traits such as trait anxiety and Intolerance of Uncertainty (IU) (Boelen & Reijntjes, 2009; Carleton, Collimore & Asmundson, 2007; McEvoy et al., 2019; Mogg & Bradley, 2002; Whiting et al., 2014), and these same trait variables are associated with compromised threat extinction learning and retention.

Intolerance of uncertainty (IU) has been defined as "an individual's dispositional incapacity to endure the aversive response triggered by the perceived absence of salient, key, or sufficient information, and is sustained by the associated perception of uncertainty." (Carleton, 2016a, p.31). IU is recognised as an important component of anxiety and stress disorders (Carleton, 2016a, 2016b; Dugas et al., 2004; Grupe & Nitschke, 2013). A series of recent experiments have revealed that

higher IU is associated with reduced safety-learning indicated by greater skin conductance responding and pupil dilation to cues that no longer signal threat throughout extinction (Morriss et al., 2015, 2016; Morriss & van Reekum, 2019). Previous research has also shown that individuals with higher IU demonstrate impaired extinction retention of learned threat (Dunsmoor et al., 2015; Lucas et al., 2018). Similarly, high levels of self-reported trait anxiety have also been associated with compromised threat extinction learning and retention (Barrett & Armony, 2009; Gazendam et al., 2013; Haaker et al., 2015). Hence, the inconsistencies in findings across the social anxiety and threat extinction literature may be the consequence of variance in dispositional traits, such as IU and trait anxiety, that have not been accounted for in previous work across the field. For example, levels of IU or trait anxiety may have been high across participants recruited in studies reporting an association between social anxiety and compromised threat extinction, but low in studies that do not find the same effect. As a result, there is a need to disentangle the transdiagnostic and disorder specific vulnerabilities associated with social anxiety when examining its effect on threat extinction processes (Shihata et al., 2016). Such an approach would provide the field with a starting point for understanding the contribution of transdiagnostic and specific components of social anxiety on extinction learning and retention, which will have relevance for targets of exposure-based treatments for social anxiety.

A further limitation of the existing literature on extinction learning and retention in social anxiety is that, despite relapse after exposure therapy being common for individuals with social anxiety (Craske & Mystkowski, 2006), research examining extinction retention in social anxiety is extremely sparse. It could be argued that social anxiety is unique in terms of the amount of exposure to the feared stimulus the

individual experiences in their daily life. It is almost impossible to avoid social interactions entirely, but despite exposure, social anxiety is maintained. To our knowledge, only two studies have investigated the relationship between social anxiety and extinction retention within discriminative learning paradigms and both studies have recruited relatively small sample sizes (Pejic et al., 2013; Rabinak et al., 2017). Therefore, further examination of the relationship between social anxiety and extinction retention processes is warranted to inform new approaches to exposure-based therapies that promote the retention of the extinction memory.

Therefore, the current research aimed to address a number of limitations in the existing literature on social anxiety and extinction learning and retention, across two studies. The primary research questions were: 1) Are individual differences in social anxiety related to compromised threat extinction and retention?; 2) Are previous findings of a relationship between transdiagnostic measures of anxiety (i.e., trait anxiety and IU) and extinction learning and retention replicated?; 3) Are any effects of social anxiety on extinction learning and retention robust after controlling for transdiagnostic features of anxiety (i.e., trait anxiety and IU)? To examine these questions, we carried out two separate socially relevant conditioning experiments designed to target extinction learning and extinction retention processes. The first experiment employed a cognitive component during extinction, while the second experiment comprised a standard extinction paradigm.

2. **Introduction Experiment 1**

Cognitive mechanisms play a vital role in the development and maintenance of social anxiety (Clark & Wells, 1995; Heimberg et al., 2014; Hirsch & Clark, 2004; Hofmann, 2007). The cognitive model assumes that socially anxious individuals allocate attentional resources towards detecting social threat in the environment and

there is increasing evidence that such attentional biases towards threat play a causal role in the maintenance of social anxiety (Heimberg et al., 2014; Hofmann, 2007).

Across the conditioning literature, previous studies have found that highly socially anxious individuals demonstrate difficulties differentiating between threat and safety cues (Ahrens et al., 2014; Hermann et al., 2002; Sachs et al., 2003). These findings suggest that compromised discrimination learning may be a characteristic of elevated social anxiety and imply an increased tendency in socially anxious individuals to generalise conditioned fear across safe cues (Ahrens et al., 2016).

In experiment 1, we used a Selective Extinction through Cognitive Evaluation (SECE) task (development of task outlined in Macdonald et al., 2020), that adds additional cognitive load into an extinction study. Within this task, contextual information was presented with the CS+ and CS- during the extinction phase. Participants were required to allocate attention towards and cognitively evaluate these contexts on a trial-by-trial basis to determine the threat level of the CS+. Therefore, as in a real-life situation, participants were required to assess all the information presented to predict whether the US is likely to occur. If socially anxious individuals exhibit attentional biases towards threat during extinction, we might predict that such individuals will have difficulty using contextual "safety" information to determine the threat level of a stimulus previously associated with an aversive event.

Therefore, in this study, we hypothesised that high social anxiety, relative to low social anxiety, would be associated with larger conditioned responding to the CS+ associated with a safe context during the late part of SECE learning (day 1) and the early part of the SECE retention (day 2), as demonstrated by greater skin conductance responsivity. We also hypothesised that ratings of anxiety would be

higher toward the CS+ relative to the CS- throughout the experiment for high socially anxious participants compared to low socially anxious participants. Following previous research, we hypothesised that trait anxiety and IU would have these same effects on extinction learning and retention. Further, we aimed to explore whether any effects of social anxiety would be consistent over and above IU and trait anxiety (Dunsmoor et al., 2015; Lucas et al., 2018; Morriss et al., 2015, 2016; Morriss & van Reekum, 2019).

3. Method Experiment 1

3.1 Participants

The sample consisted of 83 female students (age *M* = 19.61, *SD* = 1.39; Ethnicity: 66 White, 9 Asian/Pacific Islander, 4 Black, 4 Multi-ethnic). Participants were recruited if they were female, aged 18 to 35, and not currently receiving treatment for a psychiatric disorder, including medication. Females were recruited due to the consistently higher prevalence of social anxiety in females compared to males (Remes et al., 2016). Females also demonstrate higher levels of social anxiety when using a dimensional approach (Sosic et al., 2008). Further, female faces and voices were used as conditioned and unconditioned stimuli, and it was thought that a female voice administering critical statements would have a different threat value to male participants compared to female participants. All participants had normal or corrected to normal vision. Participants provided written informed consent and received £15 or course credit for their participation.

There were two participants that did not return to the laboratory for the SECE retention phase. Further, six participants were missing ratings data due to incorrect button presses that were not recorded. Therefore, 81 participants were included in SCR analyses, and 75 participants were included in the analyses of anxiety ratings.

The power analyses suggested a sample size of 67 participants for this experiment based on an effect size of d = 0.89. This effect size was taken from the original SECE experiment that reported a main effect of context (CS+ dangerous vs CS- dangerous) for SCR magnitude during SECE learning (Macdonald et al., 2020). The following parameters were used: f = 0.45 (converted from d = 0.89), α error probability = 0.05, Power (1-β error probability) = 0.95, number of groups = 4 (CS: CS+, CS-; Context: Dangerous, Safe), numerator df = 1, number of covariates = 3 (SPIN, IU, Trait Anxiety). We oversampled by 14 participants for SCR analyses and 8 participants for analyses of ratings data[1].

The procedure was approved by the University of Reading Research Ethics Committee.

3.2 Procedure

Upon arrival on day 1, participants were informed about the experimental procedures and were provided with a verbal description of the SECE task (Macdonald et al., 2020). They were then asked to complete a consent form and a series of questionnaires (see 'Questionnaires' below). Participants were seated in a testing booth, and physiological sensors were placed on the participant's index, middle and ring finger. The stimulator electrode was placed on the little finger of the left hand. The shock level for each participant was set following procedures outlined in Delgado et al (2008) (See 'Shock Build-Up Procedure' in Supplementary Materials). Participants were asked whether they had any questions and were instructed to respond to the ratings using the number keys on the keyboard with their right hand. They were asked to stay as still as possible. At this point, the SECE task

[1] We oversampled in experiment 1 as an initial power analysis, conducted before data collection, used an estimated effect size due to the lack of published research using the SECE task. The power analysis has since been updated to use the most appropriate effect size from the original SECE experiment (Macdonald et al., 2020).

(see 'SECE Task' below for details) was presented on a computer screen while skin conductance response (SCR), pulse, and behavioural ratings were recorded.

On the second day (24 hours later), participants were provided with similar instructions of the SECE task (see ‘SECE Instructions’ below). The same physiological setup was used as above. Each session took approximately 45 minutes in total.

3.3 SECE Instructions

The instructions provided for the SECE task were adapted from those administered in Macdonald et al's (2020) experiment. Participants were told that they would see two female identities during the first phase and that one identity would be associated with the risk of electric shock and negative statement. Participants were informed that, after a break, the identities would again be presented, but an image of a 'place' would also be briefly displayed. Most of the places would belong to two distinct categories: houses or skyscrapers. Participants were instructed that one of these categories was "safe", meaning that regardless of the identity presented they would not receive the shock or statement. They were asked to determine which place category was safe during this phase and to keep the contingencies in mind throughout. Participants were also asked to focus on whether or not they were likely to receive the shock and statement during each trial. Finally, participants were asked to press the space bar on the rare occasion that they saw an image that did not belong to one of the two place categories (i.e., distractor trials).

Participants were provided with the same instructions on day 2. They were informed that they would not see the identities alone, but always alongside the 'place' images. Participants were not instructed that the contingencies between the identity and context images would be the same as on day 1.

3.4 SECE Task

The SECE task was designed using Matlab 2016b software (Mathworks, Ltd) and was presented using Psychtoolbox-3 (Kleiner et al., 2007). Visual stimuli were presented using a screen resolution of 800 X 600 with a 60 Hertz refresh rate. Participants sat approximately 60 cm from the computer screen. Visual stimuli included two photographs of neutral expressions of two white female identities taken from the Chicago Face Database (Ma et al., 2015). Actors were chosen from a set of 37 white female faces based on normative data collected from over 90 individuals (96 raters for identity 1 and 91 raters for identity 2). The two identities were chosen based on having comparable subjective ratings of age and expressions of happiness, anger, and disgust presented in the neutral expression rated on a 7-point Likert scale. One was brunette, the other blonde. Aversive Statements were delivered throughout headphones by a female voice and were taken from previous studies carried out by Ahrens et al (2014), Lissek et al (2008) and Wiggert et al (2017). The volume of the statements was standardised across participants by using fixed volume settings on the presentation computer.

The paradigm comprised 3 phases, threat acquisition, Selective Extinction through Cognitive Evaluation (SECE) learning, and SECE retention. Acquisition and SECE learning took place on day 1, and SECE retention took place 24 hours later on day 2. During acquisition, one of the female identities (blonde or brunette) was paired with the electric shock and negative statement 50% of the time (CS+). In contrast, the other identity (brunette or blonde) was always presented alone (CS-). The 50% pairing rate was designed to maximise the unpredictability of the CS+/US contingency (Grady et al., 2016; Jenkins & Stanley, 1950; Leonard, 1975). Conditioning contingencies were counterbalanced across subjects. Following

acquisition, the SECE learning phase took place during which participants were presented with both identities (CS+ and CS-). However, one of two 'context' stimuli (houses or skyscrapers) was also presented for a period of the trial. Context stimuli consisted of 12 different images of houses and 12 different images of skyscrapers that were randomly presented alongside the CS. One context category represented a dangerous trial, i.e., when the dangerous context was paired with the CS+, participants received the US in 50% of trials. The other context category represented a safe trial, i.e., when the safe context was paired with the CS+, there was not a risk of receiving the US. Dangerous and safe context categories were counterbalanced across participants. The SECE procedure was repeated when participants returned on day 2.

The acquisition phase comprised 40 trials (10 CS+ reinforced, 10 CS+ non-reinforced, 20 CS-). The SECE learning and SECE retention phases both consisted of 84 trials;16 CS+ dangerous reinforced, 16 CS+ dangerous non-reinforced, 16 CS+ safe, 16 CS- dangerous, 16 CS- safe and 4 distractor trials (2 CS+ and 2 CS-). We included one trial of each condition as orientation trials (CS+/CS-; dangerous/safe), none of which were reinforced, at the start of each SECE phase. The fifth trial of SECE was always a reinforced CS+ dangerous trial to inform participants of the CS/context contingencies.

Early SECE learning and SECE retention was defined as the first seven trials (after orientation) of each trial type, and late SECE learning and SECE retention was defined as the last eight trials of each trial type. Two distractor trials (i.e., 1 CS+ and 1 CS-) were presented during early SECE learning and SECE retention, and two distractor trials were presented during late SECE learning and SECE retention. Distractor trials were included to ensure that participants were

paying attention to the CS with context contingencies throughout the task. These trials were not included in analyses.

Experimental trials were pseudo-randomised. Two separate pseudo-randomised trial sequences were generated for acquisition, SECE learning and SECE retention before data collection. The task was programmed so that a maximum of 3 trials of the same trial type would be presented in a row and sequences were counterbalanced across participants.

During acquisition, the CS was presented on the screen for 4000 ms. During reinforced trials, the statement was presented 3000 ms after CS onset and co-terminated with the trial. The electric shock lasted for 200 ms and also co-terminated with the trial. Trials during the SECE learning and SECE retention phases had a duration of 5500 ms. The identity was presented alone for 1000 ms; the context was presented on the screen below the CS for a duration of 500 ms before disappearing. The identity then remained on screen alone for a further 4000 ms. During reinforced CS+ dangerous trials, the negative statement was presented 4500 ms into the trial, and the onset of the statement was followed by the electric shock that lasted for 200 ms and co-terminated with the trial. A jittered ITI, ranging between 6000ms and 8000ms, consisted of a black screen and followed each stimulus presentation throughout the task.

During acquisition, blocks were made up of 20 trials, and during SECE learning and SECE retention blocks were made up of 42 trials. Participants were given a break between blocks in each phase. At certain time points throughout the task (Day 1: before acquisition, after acquisition and after SECE learning; Day 2: before SECE retention and after SECE retention) participants were asked to rate

how anxious each identity made them feel on a scale ranging from 1 ("not at all") to 9 ("extremely").

3.5 Questionnaires

To assess social anxiety, we administered the Social Phobia Inventory (SPIN) (Connor et al., 2000), which consists of 17 items that are rated on a 5-point Likert scale[2]. We also administered the Intolerance of Uncertainty Scale (Freeston et al., 1994), which includes 27 items that are rated on a 5-point Likert scale, and the State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1983), which is made up of 20 State items and 20 Trait items rated on a 4-point Likert scale. Although collected, state anxiety scores were not included in the analysis as not deemed relevant to the research questions. Scores to all questionnaires were mean-centred before being entered into the analyses. Cronbach's alphas for all scales were >.9. Means and distributions were as follows, SPIN (*M* = 20.61; *SD* = 11.50; Range = 0 - 58), IU (*M* = 61.55; *SD* = 18.68; Range = 29 - 106) and trait anxiety (*M* = 41.12; *SD* = 11.10; Range = 22 - 68) (See Supplementary Materials for the distribution of questionnaire scores).

3.6 Skin Conductance Acquisition and Scoring

Physiological recordings were obtained using AD Instruments (AD Instruments Ltd, Chalgrove, Oxfordshire) hardware and software. Electrodermal activity was measured with dry MLT118F stainless steel bipolar finger electrodes that were attached to the distal phalanges of the index and middle fingers of the non-dominant hand. A low constant-voltage AC excitation of 22 mVrms at 75 Hz was

[2] The SPIN was initially designed to establish threshold scores that might best differentiate between individuals with and without pathological social anxiety, such that the scale could act as a useful screening instrument within the assessment of SAD (Connor et al., 2000). Further, the SPIN has been used to examine individual differences in social anxiety on extinction learning in previous work (Pejic et al., 2013; Reichenberger et al., 2017; Shiban et al., 2015).

passed through the electrodes, which were connected to a PowerLab 8/35, and converted to DC before being digitised and stored. The electrodermal signal was converted from volts to microSiemens using AD Instruments software (AD Instruments Ltd, Chalgrove, Oxfordshire).

We used a similar scoring procedure to previous studies (Morriss, 2019; Morriss et al., 2019). Skin conductance responses were marked using AD Instruments software (AD Instruments Ltd, Chalgrove, Oxfordshire) and extracted using Matlab R2017a software (The MathWorks, Inc., Natick, Massachusetts, United States). Skin conductance responses (SCR) were scored when there was an increase in skin conductance level exceeding 0.03 microsiemens (Dawson et al., 2000). The amplitude of each response was scored as the difference between the onset and the maximum deflection before the signal flattened out or decreased. SCR onsets and respective peaks were counted if the SCR onset was within 0.5-3.5 seconds (CS response) following CS onset during acquisition and 0.5-5 seconds (CS response) following CS onset during SECE learning and SECE retention. Trials with no discernible SCRs were scored as zero. SCR magnitudes were square-root transformed to reduce skew and z-scored (across trials for each participant) to control for interindividual differences in skin conductance responsiveness (Ben-Shakhar, 1985). CS+ non-reinforced and CS- trials were included in the analysis, but CS+ reinforced trials were discarded to avoid confounds from the sound and electric shock. Orientation trials were also discarded from analyses as participants were not yet aware of the Stimulus/Context contingencies. SCR magnitudes were calculated from remaining trials by averaging SCR-transformed values for each condition. Non-responders were defined as those who responded to 10% or less of the CS+ unpaired and CS- trials (Morriss et al., 2018; Xia et al., 2017). Eight non-responders

were identified in this experiment. As excluding the non-responders did not alter the pattern or significance of SCR findings, for completeness non-responders were included in the analysis of the SCR data. There were five participants that did not learn the association between the neutral cue and the US (see 'Learning Assessment Criteria' in Supplementary Materials). Again, as removing these participants did not change the results reported, for completeness, these participants were included in the analyses.

3.7 SCR and Ratings Analysis

The analyses were conducted using the mixed procedure in SPSS 25.0 (SPSS, Inc; Chicago, Illinois). We conducted separate MLMs for SCR magnitude and subjective ratings during threat acquisition, SECE learning and SECE retention. For SCR magnitude and ratings during the acquisition phase, we entered Stimulus (CS+, CS-) at level 1 and individual subjects at level 2. For SCR magnitude and ratings during SECE learning and SECE retention, we entered Stimulus (CS+, CS-), Context (Dangerous, Safe), and Time (Early, Late) at level 1 and individual subjects at level 2. SPIN, IU and STAI-T scores were included as individual difference predictor variables in MLMs examining SCR magnitude and ratings during SECE learning and SECE retention. Separate MLMs were carried out to investigate the effect of each individual difference variable on dependent variables (i.e., separate models for each predictor; SPIN, IU and STAI-T score)[3].

Fixed effects included Stimulus, Context and Time. A diagonal covariance matrix for level 1 was used in all models. A random intercept for each participant was included as random effects, where a variance components covariance structure was

[3] As there was not an effect of social anxiety on SCR magnitude or self-report measures of anxiety or US expectancy in either experiment 1 or experiment 2, we did not run additional MLM analyses to examine whether effects of social anxiety were robust when controlling for IU and STAI-T scores.

used. We used a maximum likelihood estimator for the MLMs and corrected post-hoc tests for multiple comparisons using the Benjamini-Hochberg False Discovery Rate procedure (Benjamini & Hochberg, 1995). Level 1 variables were categorical and therefore contrast coded. Further, separate models were carried out for the three-way interactions between Stimulus, Context and Time that examined the effects of the task (i.e., section 4.2) and the four-way interactions that include individual differences variables (i.e., sections 4.3 and 4.4).

Where an interaction with an individual difference variable (SPIN, IU, STAI-T) was observed, follow-up pairwise comparisons were performed on the estimated marginal means of the relevant conditions at specific values of + or – 1 SD of the mean individual difference score. These values are estimated from the multilevel model of the complete sample, not unlike performing a simple slopes analysis in a multiple regression analysis. Similar analyses have been published elsewhere (Morriss et al., 2016; Morriss et al., 2020).

4. Results Experiment 1

4.1 Self-reported reactions to Unconditioned Stimuli

Participants reported that the electric shock (*M* = 5.6, *SD* = 1.88) and negative statement (*M* = 4.78, *SD* = 2.42) made them feel anxious (where 1 = “not at all” and 9 = “extremely”), after SECE learning on day 1. Participants rated the electric shock as significantly more anxiety-provoking compared to the negative vocal statements, $t(80) = 3.56$, $p = .001$. Individual differences in trait social anxiety were not significantly correlated with ratings of anxiety elicited by the socially relevant negative statements, $r(81) = -0.02$, $p = .89$.

4.2 Effects of the task

4.2.1 Acquisition

SCR magnitude was significantly greater towards the CS+ compared to the CS- during the acquisition phase [Stimulus, $F(1, 132.95) = 8.47$, $p = .004$, see Table 1].

There was not a significant difference in anxiety ratings towards the CS+ and CS- before acquisition, however, after the acquisition phase, anxiety ratings were significantly higher towards the CS+ compared to the CS- [Stimulus, $F(1, 163.77) = 116.77$, $p < .001$; Time, $F(1, 163.77) = 65.78$, $p < .001$; Stimulus x Time, $F(1, 163.77) = 160.78$, $p < .001$, see Table 2].

These findings indicate that conditioning was effective during the acquisition phase.

4.2.2 SECE Learning

There was a significant stimulus x context x time interaction for SCR during SECE learning [Stimulus x Context x Time, $F(1, 503.72) = 4.53$, $p = .03$, see Table 1]. SCR was significantly greater towards the CS+ dangerous compared to the CS- dangerous during early SECE learning, $p = .03$. However, SCR significantly reduced towards the CS+ dangerous between early SECE learning and late SECE learning, $p = .01$. Further, SCR declined towards the CS- safe between early SECE learning and late SECE learning, $p = .02$.

During SECE learning, anxiety ratings were significantly higher towards the CS+ compared to the CS-, $p < .001$ [Stimulus, $F(1, 161.85) = 384.45$, $p < .001$]. There was not a main effect of time or a significant stimulus x time interaction for anxiety ratings towards the CS+ and CS- across the SECE learning phase [Time,

F(1, 162.01) = 0.06, *p* = .8; Stimulus x Time, *F*(1, 161.85) = 1.15, *p* = .29, see table 2].

4.2.3 SECE Retention

There was not a significant stimulus x context x time interaction for SCR during SECE retention [Stimulus x Context x Time, *F*(1, 613.2) = 0.01, *p* = .91, see Table 1]. There was a main effect of stimulus, *p* = .04 [Stimulus, *F*(1, 613.2) = 4.46, *p* = .04], indicating significant differential responding between CS+ (dangerous and safe, early and late) and CS- (dangerous and safe, early and late) trials throughout SECE retention.

During SECE retention, participants demonstrated significantly higher anxiety ratings towards the CS+ compared to the CS-, *p* < .001 [Stimulus, *F*(1, 168.52) = 292.58, *p* < .001]. Further, ratings of anxiety towards the CS+ increased marginally over time, *p* = .056, while there was a decline at trend for ratings of anxiety towards the CS- over time, *p* = .088 [Time, *F*(1, 168.52) = 1.98, *p* = .16; Stimulus x Time, *F*(1, 168.52) = 5.32, *p* = .02, see Table 2].

4.3 Aim 1: The effect of social anxiety on SECE learning and SECE retention

4.3.1 SCR Magnitude

Contrary to hypotheses, individual differences in social anxiety were not related to SCR during SECE learning, [Stimulus x Context x Time x SPIN, *F*(1, 639.28) = 0.74, *p* = .39]. There were no other significant interactions between social anxiety and stimulus, context or time in this analysis, max *F* = 0.74.

Further, there was not a significant relationship between individual differences in social anxiety and SCR during SECE retention, [Stimulus x Context x Time x

SPIN, $F(1, 639.83) = 0.35$, $p = .56$]. There were no other significant interactions between social anxiety, stimulus, context or time in this analysis, max $F = 1.42$.

4.3.2 Self-reported anxiety elicited by Conditioned Stimuli

There was not a significant relationship between individual differences in social anxiety and anxiety ratings towards the CS+ and CS- across the SECE learning phase, [Stimulus x Time x SPIN, $F(1, 161.82) = 0.01$, $p = .92$]. There were no other significant interactions between social anxiety, stimulus or time in this analysis, max $F = 0.68$.

Again, during SECE retention, there was not a significant relationship between individual differences in social anxiety and anxiety ratings towards the CS+ and CS-, [Stimulus x Time x SPIN, $F(1, 160.23) = 0.001$, $p = .97$]. There were no other significant interactions between social anxiety, stimulus and time in this analysis, max $F = 0.01$.

4.4 Aim 2: Replication of previous literature – The effect of IU and trait anxiety on SECE learning and SECE retention

4.4.1 IU and SCR magnitude

Individual differences in IU were marginally related to SCR magnitude during the SECE learning phase, [Stimulus x Context x Time x IU, $F(1, 639.34) = 3.43$, $p = .06$, see Figure 2]. There were not any significant or marginal interactions between stimulus, context and time at lower IU, $p > .11$. However, higher IU scores were associated with significantly increased SCR towards the CS+ dangerous compared to the CS- dangerous during early SECE learning, $p = .02$. Further, higher IU was marginally associated with increased SCR towards the CS+ dangerous compared to

the CS+ safe during early SECE learning, *p* = .08. SCR magnitude towards the CS+ dangerous decreased at trend between early and late SECE at higher IU, *p* .07.

Individual differences in IU were also related to SCR magnitude during SECE retention, [Stimulus x IU, *F*(1, 609.27) = 5.42, *p* = .02; Stimulus x Context, *F*(1, 609.27) = 4.90, *p* = .03, see Figure 2]. Lower IU scores were associated with increased SCR magnitude towards CS+ dangerous compared to CS+ safe trials throughout the SECE retention phase, *p* = .004. Further, there was no discernible difference in SCR between CS- dangerous and CS- safe trials throughout SECE retention at low IU, *p* = .45. However, higher IU scores were not affected by context during SECE retention, *p* = .57, as higher IU scores were associated with significantly increased SCR magnitude during CS+ (dangerous and safe) compared to CS- (dangerous and safe) trials throughout the extinction retention phase, *p* = .002.

4.4.2 IU and self-reported anxiety elicited by Conditioned Stimuli

Individual differences in IU were related to ratings of anxiety elicited by the CS+ and CS- during SECE learning, [Stimulus x IU, *F*(1, 162.42) = 11.38, *p* = .001, see Figure 3]. Parameter estimates indicated that higher IU compared to lower IU was associated with significantly higher ratings of anxiety towards the CS+ both before*, p* = .001, and after, *p* = .004, the SECE learning phase. This effect was present for anxiety ratings toward the CS- pre-SECE learning, *p* =.01, but not post-SECE learning, p = .23).

There was also a significant relationship between individual differences in IU and ratings of anxiety towards the CS during SECE retention [Stimulus x IU, *F*(1, 165.2) = 30.63, *p* < .001, see Figure 3]. Parameter estimates demonstrated that

higher IU compared to lower IU predicted higher ratings of anxiety towards the CS+ both before, p = .003, and after, p = .002, the SECE retention phase. This effect was not present for anxiety ratings toward the CS- (pre-SECE retention, p =.38, post-SECE retention, p = .5).

4.4.3 Trait Anxiety and SCR

There was not a significant relationship between individual differences in trait anxiety and SCR during SECE learning, [Stimulus x Context x Time x STAI-T, F(1, 639.28) = 0.48, p = .49]. There were no other significant interactions between social anxiety and stimulus, context and time in this analysis, max F = 1.41.

Further, there was not a significant relationship between individual differences in trait anxiety and SCR during SECE retention, [Stimulus x Context x Time x STAI-T, F(1, 611.72) = 0.81, p = .37]. There were no other significant interactions between trait anxiety, stimulus, context and time in this analysis, max F = 1.46.

3.4.4 Trait Anxiety and self-reported anxiety elicited by Conditioned Stimuli

There was not a significant relationship between trait anxiety scores and anxiety ratings towards the CS+ and CS- across the SECE learning phase, [Stimulus x Time x STAI-T, F(1, 162.51) = 0.03, p =.86]. There were no other significant interactions between trait anxiety, stimulus and time in this analysis, max F = 0.27.

Again, during SECE retention, there was not a significant relationship between individual differences in trait anxiety and anxiety ratings towards the CS+ and CS-, [Stimulus x Time x STAI-T, F(1, 204.28) = 0.43, p =.52]. There were no other significant interactions between trait anxiety, stimulus and time in this analysis, max F = 0.43.

4.5 Aim 3: The effect of social anxiety on SECE learning and SECE retention when controlling for IU and trait anxiety

As we did not find any significant effects of social anxiety on SCR magnitude or self-reported reactions to the CS during extinction learning or retention when carrying out analyses to examine Aim 1, we did not conduct follow up analyses to investigate whether effects of social anxiety were robust when controlling for IU and trait anxiety scores within a single model.

5. Experiment 1 Conclusion

Contrary to hypotheses, the results from experiment 1 suggest that social anxiety does not play a role in the maintenance of learned threat during SECE learning or SECE retention. Further, we did not replicate findings from the previous literature (Barrett & Armony, 2009; Gazendam et al., 2013; Haaker et al., 2015) that demonstrated that trait anxiety is associated with compromised extinction processes. Instead, SCR magnitude results and findings across anxiety ratings indicate that IU is associated with extinction processes during SECE learning and retention. During the early part of SECE learning, higher IU was marginally associated with the adaptive use of contextual cures to update a conditioned response to a learnt threat cue, as indexed by larger differential SCR magnitude responding during CS+ dangerous trials compared to CS+ safe trials, suggesting that participants with higher IU can successfully decrease their emotional arousal to the CS+ when they identify, based on the context presented, that there is no risk of receiving the US. In contrast, during SECE retention, higher IU was associated with increased SCR magnitude to CS+ compared to CS- trials, with no effect of context. Such a result suggests that higher IU predicted the generalisation of threat cues across contexts in the presence

of the CS+ during SECE retention. These findings support previous research that suggests compromised extinction processes are related to broader and transdiagnostic features of anxiety, such as IU (Dunsmoor et al., 2015; Lucas et al., 2018; Morriss et al., 2015, 2016; Morriss & van Reekum, 2019), and build upon this work through the use of a socially relevant paradigm that includes a partial instruction and a cognitive component.

6. Experiment 2 Introduction

The results of study 1 suggest that individual differences in IU and not social anxiety are associated with the maintenance of learned threat in a socially relevant extinction task that incorporates a cognitive component. However, it is possible that features of the SECE task masked an effect of social anxiety on extinction learning and retention in experiment 1. For instance, there were a high number of trials during the SECE learning and retention phases that could have led to fatigue and habituation of the US. Further, partial instruction regarding the contingencies between CS and contexts might have played a role in the absence of social anxiety related findings.

Previous work has demonstrated that when instructed to refrain from using safety behaviours, socially anxious individuals were less negative in judgements of their performance and rated the likelihood of a negative outcome as less than those who were not instructed (Taylor & Alden, 2010). A further study demonstrated that when instructed to direct their attention to components of the external environment, socially anxious individuals reported lower anxiety and less negative beliefs (Wells & Papageorgiou, 1998). It may be that individuals with elevated levels of social anxiety can use instructions to reduce anxiety during extinction learning and retention, resulting in successful extinction of the conditioned response in experiment 1. This

may not be the case at high levels of IU, where the partial instruction might have induced uncertainty during the SECE retention phase. Before SECE retention in experiment 1, participants were not explicitly instructed that the CS with context contingencies would be the same as during day 1, perhaps resulting in the perception of missing information in individuals with high IU. This perceived uncertainty may have resulted in generalisation across safe and dangerous contexts for individuals with higher IU during extinction retention. Therefore, during experiment 2, we removed possibly confounding task features and employed a standard threat extinction paradigm with socially relevant stimuli.

Experiment 2 aimed to further examine whether individual differences in social anxiety are related to impaired extinction learning and retention as well as to follow up on the effects of IU that were found during experiment 1. We hypothesised that high socially anxious participants, compared to low socially anxious participants, would demonstrate sustained skin conductance responding to the CS+ relative to the CS- during late extinction learning (day 1) and early extinction retention (day 2). We further expected that high social anxiety, relative to low social anxiety, would be associated with higher ratings of anxiety as well as higher US expectancy ratings towards the CS+ compared to the CS- throughout the experiment. Given that the findings of experiment 1 suggest that IU plays a role in the maintenance of conditioned threat within a socially relevant paradigm, and in line with prior work on IU and trait anxiety, we examined whether the hypothesised effects outlined above could also be true of IU and trait anxiety. Hence, in this simplified social extinction study, we expected that any effect of social anxiety would not be specific to social anxiety, but that IU would also be associated with deficits in extinction learning and retention.

7. **Method Experiment 2**

The method was identical to experiment 1, except as follows.

7.1 Participants

The sample consisted of 92 female students (*M* age = 19.66, SD = 1.38; Ethnicity: 67 White, 14 Asian/Pacific Islander, 7 Black, 3 Multi-ethnic, 1 Middle Eastern/Arab) recruited from the University of Reading. Participants were not eligible to participate if they had taken part in Study 1. Participants provided written informed consent and received course credit for their participation.

There were six participants that did not return to the laboratory for the extinction retention phase, and there were software errors for two participants. Therefore, 84 participants were included in the analysis of SCR and ratings data.

The power analysis is based on an average effect size (η2p = 0.16) taken from the Stimulus x Time x IU interaction for SCR magnitude from five published experiments using highly similar experimental designs as the current experiment (4/5 experiments reported significant effects of IU) (Morriss et al., 2015; 2016; Morriss & van Reekum, 2019). The following parameters were used: *f* = 0.43 (converted from η2p = 0.16), α error probability = 0.05, Power (1-β error probability) = 0.95, number of groups = 2 (CS: CS+, CS), numerator df = 1, number of covariates = 3 (SPIN, IU, Trait Anxiety). The suggested sample size was 73 participants. Based on the updated power analysis, we oversampled by 10 participants.[4]

The procedure was approved by the University of Reading Research Ethics Committee.

7.2 Procedure

[4] We have oversampled in experiment 2 as the initial power analysis, conducted prior to data collection, included an effect size of *f* = 0.22, reported in Morriss et al (2016). We have updated the power analysis to include a more reliable effect size averaged across five similar experiments.

Experiment 2 used a standard extinction paradigm. The same stimuli, experimental parameters and physiological setup were used as in Experiment 1. However, context images were not presented in the extinction learning and extinction retention phase, and participants were not given any verbal instructions about what to expect during the task.

7.3 Conditioning Paradigm

The paradigm was comprised of three phases: acquisition, extinction learning, and extinction retention. Following acquisition, extinction learning took place during which both identities (CS+ and CS-) were presented in the absence of the shock and negative statement (US). During the extinction retention phase on day 2, participants were again presented with both identities in the absence of the US.

The acquisition phase consisted of 24 trials (6 CS+ paired, 6 CS+ unpaired, and 12 CS-). The extinction learning phase was comprised of 32 trials (16 CS+ unpaired and 16 CS-), where early extinction learning was defined as the first 8 CS+/ CS- trials and late extinction learning was defined as the last 8 CS+/CS- trials. The extinction retention phase included of 32 trials (16 CS+ unpaired and 16 CS-), again where early extinction retention was defined as the first 8 CS+/CS- trials and late extinction retention was defined as the last 8 CS+/CS- trials.

The identities were presented on the screen for a total of 4000 ms during all phases of the task. During reinforced trials, the negative statement was presented after 3000 ms, and the electric shock lasted for 200 ms. Both US co-terminated with the trial. A jittered ITI of between 6000 ms and 8000 ms occurred between trials and consisted of a black screen.

Blocks of trials in acquisition were made up of 12 trials and in extinction learning and extinction retention blocks included 16 trials. Participants were asked to

rate how anxious each identity made them feel before acquisition, after acquisition, after extinction learning on day 1, and before extinction retention and after extinction retention on day 2. The scale ranged from 1 ("not at all") to 9 ("extremely"). Participants were also asked to rate whether they expected to receive negative stimuli in the presence of each identity at the above time points throughout the experiment. The scale also ranged from 1 ("do not expect") to 9 ("do expect").

7.4 Questionnaires

Means and distributions for the anxiety measures were as follows, SPIN (*M* = 22.85; *SD* = 11.75; Range = 2 - 49), IU (*M* = 64.68; *SD* = 17.91; Range = 35 - 113) and trait anxiety (*M*= 44.63; *SD* = 10.87; Range = 22 - 73). Cronbach's alphas for all scales were >.9.

7.5 Physiological Acquisition and Scoring

SCR onsets and respective peaks were included if the SCR onset was within 0.5-3.5 seconds (CS response) following CS onset (Morriss et al., 2018) in all three phases. As excluding non-responders (N=14) did not alter the pattern or significance of SCR findings, for completeness non-responders were included in the analysis of the SCR data. No participants were identified as non-learners in this experiment (see ‘Learning Assessment Criteria’ in Supplementary Materials).

7.6 SCR and Ratings Analysis

The procedure of data analysis was the same in experiment 2 as in experiment 1, apart from the following: context was not included as a variable in the model at level 1 and MLMs were also conducted for expectancy ratings during acquisition, extinction learning and extinction retention. Further, we did not expect specificity for SPIN but also expected to observe significant interactions with IU.

8. Results Experiment 2

8.1 Self-reported reactions to Unconditioned Stimuli

Participants rated the electric shock (*M* = 5.61, *SD* = 1.67) and negative statements (*M* = 4.46, *SD* = 2.15) as making them feel anxious (where 1 = “not at all”, 9 = “extremely”) after extinction learning on day 1. The electric shock was rated as more anxiety-provoking compared to the negative vocal statements, $t(82) = 5.48$, $p < .001$, Cohen’s $d = 0.59$. Individual differences in social anxiety were not significantly correlated with ratings of anxiety elicited by the socially relevant negative statements, $r(83) = 0.08$, $p = .46$. There was missing stimulus rating data from one participant.

8.2 Effects of the task

8.2.1 Acquisition

SCR was significantly greater towards the CS+ compared to the CS- during the acquisition phase [$F(1, 84) = 10.13$, $p = .002$, see Table 3].

There was no significant difference in anxiety ratings between the CS+ and CS- before acquisition, however, after acquisition anxiety ratings were significantly higher towards the CS+ versus the CS- [Stimulus, $F(1, 191.72) = 165.44$, $p < .001$, Time, $F(1, 191.72) = 132.03$, $p < .001$, Stimulus x Time, $F(1, 191.72) = 161.06$, $p < .001$, see Table 3].

Further, US expectancy ratings were significantly greater towards the CS+ compared to the CS- after acquisition [$F(1, 167.39) = 590.08$, $p < .001$, see Table 3].

These findings indicate that conditioning was effective during the acquisition phase.

8.2.2 Extinction Learning

During extinction learning, SCR was significantly higher to the CS+ compared to the CS- [Stimulus, $F(1, 247.91) = 7.54$, $p = .006$]. There was no effect of time and no stimulus x time interaction for SCR during extinction learning [Time, $F(1, 247.91) = 0.37$, $p = .55$; Stimulus x Time, $F(1, 247.91) = 0.16$, $p = .67$, see Table 3].

Participants rated the CS+ as more anxiety provoking compared to the CS-, $p < .001$. Anxiety ratings towards the CS+ significantly decreased across the extinction phase, $p < .001$ [Stimulus, $F(1, 185.16) = 357.01$, $p < .001$; Time, $F(1, 185.16) = 41.28$, $p < .001$; Stimulus x Time, $F(1, 185.16) = 14.17$, $p < .001$, see Table 3].

During the extinction learning phase, participants demonstrated significantly higher US expectancy ratings towards the CS+ compared to the CS-, $p < .001$. Further, expectancy ratings towards the CS+ significantly reduced over time, $p < .001$ [Stimulus, $F(1, 219.42) = 505.09$, $p < .001$; Time, $F(1, 219.42) = 85.09$, $p < .001$; Stimulus x Time, $F(1, 219.42) = 89.83$, $p < .001$, see Table 3].

8.2.3 Extinction Retention

SCR magnitude during extinction retention was significantly greater for the CS+ compared to the CS-, $p < .001$ [Stimulus, $F(1, 325.97) = 20.52$, $p < .001$, see table 3]. Further, SCR towards the CS+ and CS- dropped across the extinction retention phase [Time, $F(1, 325.97) = 5.25$, $p = .02$]. The interaction between stimulus x time was not significant [Stimulus x Time, $F(1, 325.97) = 0.51$, $p = .48$, see Table 3].

During extinction retention, anxiety ratings were significantly higher towards the CS+ compared to the CS-, $p < .001$ [Stimulus, $F(1, 204.15) = 132.52$, $p < .001$, see Table 3]. There was no main effect of time and the stimulus x time interaction

was not significant for anxiety ratings towards the CS+ and CS- [Time, $F(1, 204.15) = 2.39$, $p = .12$; Stimulus x Time, $F(1, 204.15) = 2.66$, $p = .1$, see Table 3].

During the extinction retention phase, participants demonstrated significantly higher US expectancy ratings towards the CS+ compared to the CS-, $p < .001$. Further, expectancy ratings towards the CS+ significantly reduced over time, $p < .001$ [Stimulus, $F(1, 179.85) = 307.61$, $p < .001$; Time, $F(1, 179.85) = 77.22$, $p < .001$; Stimulus x Time, $F(1, 179.85) = 77.22$, $p < .001$, see Table 3].

8.3 Aim 1: The effect of social anxiety on extinction learning and extinction retention

8.3.1 SCR Magnitude

During extinction learning, individual differences in social anxiety were not related to SCR magnitude towards the CS+ and CS-, [Stimulus x Time x SPIN, $F(1, 313.75) = 0.14$, $p = .71$]. There were no other significant interactions between social anxiety, stimulus and time in this analysis, max $F = 0.14$.

Further, there was not a significant relationship between individual differences in social anxiety and SCR towards the CS+ and CS- during extinction retention, [Stimulus x Time x SPIN, $F(1, 326.04) = 0.40$, $p = .53$]. There were no other significant interactions between social anxiety, stimulus and time in this analysis, max $F = 1.24$.

8.3.2 Self-reported reactions toward Conditioned Stimuli

8.3.2.1 Anxiety ratings

There was not a significant relationship between individual difference in social anxiety and anxiety ratings towards the CS+ and CS- across the extinction learning

phase, [Stimulus x Time x SPIN, $F(1, 184.91) = 0.29$, $p = .59$]. There were no other significant interactions between social anxiety, stimulus and time in this analysis, max $F = 0.29$.

Again, during extinction retention, there was not a significant relationship between individual differences in social anxiety and anxiety ratings towards the CS+ and CS-, [Stimulus x Time x SPIN, $F(1, 204.17) = 0.43$, $p = .51$]. There were no other significant interactions between social anxiety, stimulus and time in this analysis, max $F = 0.63$.

8.3.2.2 US Expectancy ratings

There was not a significant relationship between individual difference in social anxiety and US expectancy ratings towards the CS+ and CS- across the extinction learning phase, [Stimulus x Time x SPIN, $F(1, 219.76) = 2.31$, $p = .13$]. There were no other significant interactions between social anxiety, stimulus and time in this analysis, max $F = 2.31$.

Further, there was an interaction at trend between individual differences in social anxiety and US expectancy ratings towards the CS+ and CS- across the extinction retention phase, [Stimulus x Time x SPIN, $F(1, 180.84) = 3.03$, $p = .08$]. However, parameter estimates did not demonstrate that there was a significant difference in US expectancy ratings towards the CS+ and CS- at high social anxiety compared to low social anxiety before extinction retention or after the extinction retention phase, $p > .11$.

8.4 Aim 2: Replication of previous literature – The effect of IU and trait anxiety on extinction learning and extinction retention

8.4.1 IU and SCR magnitude

Individual difference in IU were marginally related to SCR magnitude during the extinction learning phase in the expected direction [Stimulus x Time x IU, $F(1, 247.39) = 2.71$, $p = .09$, see Figure 5]. Participants with lower IU scores showed increased SCR to the CS+ compared to the CS- during early extinction learning, $p = .06$. In contrast, higher IU scores were not associated with a significant difference in SCR magnitude towards the CS+ and CS- during early extinction learning, $p = .31$. But, during late extinction learning, higher IU was associated with significant differential SCR between the CS+ and CS-, $p = .01$.

Individual differences in IU were also related to SCR magnitude towards the CS+ and CS- in the expected direction during extinction retention [Stimulus x IU, $F(1, 326.62) = 3.18$, $p = .07$, see Figure 5]. Lower IU scores were not associated with differential SCR between the CS+ and CS- during early extinction learning, $p = .33$. During late extinction retention, there was a marginal differential SCR between the CS+ and CS- for participants with lower IU scores, $p = .06$. During early extinction retention, higher IU was associated with increased SCR towards the CS+ versus the CS-, $p < .001$, and this effect remained present during late extinction retention, $p = .04$. However, higher IU was associated with significant reduction in SCR towards the CS+ across the extinction retention phase, $p = .02$.

3.4.2 IU and self-reported reactions to Conditioned Stimuli during extinction learning and extinction retention

8.4.2.1 IU and Anxiety ratings

There was not a significant relationship between individual differences in IU and anxiety ratings towards the CS+ and CS- across the extinction learning phase,

[Stimulus x Time x IU, $F(1, 185.38) = 0.54$, $p = .46$]. There were no other significant interactions between IU, stimulus and time, max $F = 2.00$.

There was not a significant relationship between individual difference in IU and anxiety ratings towards the CS+ and CS- across the extinction retention phase, [Stimulus x Time x IU, $F(1, 204.38) = 0.17$, $p = .68$]. There were no other significant interactions between IU, stimulus and time, max $F = 0.47$.

8.4.2.2 IU and US Expectancy ratings

There was not a significant relationship between individual differences in IU and US expectancy ratings towards the CS+ and CS- across the extinction learning phase, [Stimulus x Time x IU, $F(1, 219.46) = 1.26$, $p = .26$]. There were no other significant interactions between IU, stimulus, and time, max $F = 1.26$.

There was a significant relationship between IU and US expectancy ratings during the extinction retention phase [Stimulus x IU, $F(1, 179.16) = 3.94$, $p = .049$. Parameter estimates demonstrated that higher IU was associated with increased US expectancy ratings towards the CS- before extinction retention, $p = .05$.

8.4.3 Trait anxiety and SCR

There was not a significant relationship between individual differences in trait anxiety and SCR during extinction learning, [Stimulus x Time x STAI-T, $F(1, 314.78) = 0.01$, $p = .91$]. There were no other significant interactions between trait anxiety and stimulus, and time in this analysis, max $F = 0.58$.

Further, there was not a significant relationship between individual differences in trait anxiety and SCR during extinction retention, [Stimulus x Time x STAI-T, $F(1,$

325.60) = 0.04, *p* = .85]. There were no other significant interactions between trait anxiety, stimulus, and time, max *F* = 0.88.

3.4.4 Trait anxiety and self-reported reactions to Conditioned Stimuli during extinction learning and extinction retention

8.4.4.1 Trait anxiety and Anxiety ratings

There was not a significant relationship between individual differences in trait anxiety and anxiety ratings towards the CS+ and CS- across the extinction learning phase, [Stimulus x Time x STAI-T, *F*(1, 185.59) = 0.63, *p* =.43]. There were no other significant interactions between trait anxiety, stimulus, and time, max *F* = 2.05.

There was also not a significant relationship between individual differences in trait anxiety and anxiety ratings towards the CS+ and CS- across the extinction retention phase, [Stimulus x Time x STAI-T, *F*(1, 204.28) = 0.43, *p* =.52]. There were no other significant interactions between trait anxiety, stimulus, and time in this analysis, max *F* = 0.43.

8.4.4.2 Trait anxiety and US Expectancy ratings

There was an interaction at trend between individual difference in trait anxiety and US expectancy ratings towards the CS+ and CS- across the extinction learning phase, [Stimulus x Time x STAI-T, *F*(1, 220.06) = 3.26, *p* =.07]. However, parameter estimates did not demonstrate that there was a significant difference in US expectancy ratings towards the CS+ or CS- at high trait anxiety compared to low trait anxiety before extinction learning or after the extinction learning phase, *p* > .12.

There was not a significant relationship between trait anxiety and US expectancy ratings during extinction retention, [Stimulus x Time x STAI-T, *F*(1,

179.19) = 0.96, p =.33]. There were no other significant interactions between trait anxiety, stimulus and time in this analysis, max F = 0.96.

8.5 Aim 3: The effect of social anxiety on extinction learning and extinction retention when controlling for IU and trait anxiety

As we did not find any significant effects of social anxiety on SCR magnitude or self-reported reactions to the CS during extinction learning or retention when carrying out analyses to examine Aim 1, we did not conduct follow up analyses to investigate whether effects of social anxiety were robust when controlling for IU and trait anxiety scores within a single model.

9. Experiment 2 Discussion

The findings from experiment 2 further suggest that individual differences in social anxiety do not affect extinction learning or retention, despite the use of socially relevant CS and US within a threat conditioning task. Further, as with experiment 1, we did not replicate findings of previous literature that demonstrates a relationship between trait anxiety levels and compromised extinction processes. Instead, the SCR magnitude results from experiment 2 support the findings from experiment 1 and suggest that compromised and delayed extinction are more likely to be characteristic of high levels of IU and not social anxiety.

10. General Discussion

We employed a social conditioning paradigm using a socially relevant CS and US to study the effect of social anxiety on extinction learning and retention in two separate experiments. Contrary to hypotheses, we did not find evidence that social anxiety is associated with delayed or impaired safety learning during threat extinction or retention. Instead, the findings across both studies suggest that IU-related

mechanisms disrupt threat extinction processes, even during a socially relevant conditioning task.

The results of both experiments are somewhat surprising, given that previous research has demonstrated a relationship between social anxiety and IU (e.g. Boelen & Reijntjes, 2009; Carleton et al., 2010). The original purpose of including IU and trait anxiety in this research was to assess the specificity of any effect of social anxiety on extinction learning and retention, over and above other individual difference measures found to be associated with compromised threat extinction processes. Instead, we observed effects of IU, but no effect of social anxiety on extinction learning or retention across both psychophysiological and self-report measures. Hence, the results support and extend previous findings regarding the role of IU in extinction processes but do not provide support for the hypothesis that deficits in safety learning and retention are associated with elevated levels of social anxiety. This research therefore suggests that safety-learning processes might be modified by features related to broader negative affect (i.e., IU) that underly anxiety, rather than by features specific to individual anxiety subtypes.

Despite this conclusion, prior findings have demonstrated that relapse after exposure therapy is common for individuals with anxiety disorders, including social anxiety (Craske et al., 2014; Graham & Milad, 2011; Hofmann & Smits, 2008). Across the numerous maintenance models of SAD, it is proposed that individuals with social anxiety engage in maladaptive cognitive and behavioural processes including anticipatory and post-event processing, avoidance and safety behaviours, performance deficits, and attentional biases (including self-focused and external threat-focused)(Clark & Wells, 1995; Heimberg et al., 2014; Hofmann, 2007; Rapee & Heimberg, 1997; Wong & Rapee, 2016). Such cognitive and behavioural

processes may be particularly resistant to modification during exposure therapy because, in contrast to other fears and phobias, social cognitions are mostly inaccessible to disconfirmation (Craske, 2003; Foa et al, 1996), in that individuals with social anxiety are reliant on estimates of what they believe others think of their behaviour. Therefore, given that exposure alone cannot challenge these cognitions, our results align with the idea that relapse after exposure treatment might be explained, not by deficits in safety learning, but by the enduring negative cognitions that underpin social anxiety.

Further, the pattern of results might be explained to some extent by the nature of the research context. It is unlikely that threat extinction paradigms, carried out in the laboratory, elicit levels of social threat comparable to social evaluation experienced in the real world. Hence, it is probable that laboratory tasks, such as those used in the current experiments, are unable to capture compromised cognitive processes specifically related to social anxiety. Threat extinction tasks are, however, uncertain by nature. The omission of the US could, in some individuals, trigger a sense of future threat uncertainty (Dunsmoor et al., 2015). For this reason, threat extinction paradigms might be more aligned with IU, rather than social anxiety-related mechanisms.

Several limitations need to be considered when interpreting the results of the reported studies. In line with the previous literature that has highlighted the strengths of using socially relevant stimuli when investigating the role of social anxiety on conditioning processes (Lissek et al., 2008; Pejic et al., 2013), we presented participants with neutral facial expressions as CS and negative verbal feedback as US. Alongside the verbal feedback, we employed a second socially irrelevant US and administered an electric shock in both experiments. The shock was included due

to previous research that indicated that habituation to derogatory statements during acquisition resulted in the conditioned response not lasting into the extinction phase (Lissek et al., 2008). Considering that the primary focus across both experiments was on the effect of social anxiety on threat extinction and retention, all participants, regardless of their level of social anxiety, were required to maintain a conditioned response towards the CS+ into the extinction learning phase. Therefore, we chose to include the shock alongside the statements to ensure that the US was sufficiently threatening. However, participants rated the electric shock as significantly more anxiety-provoking than the verbal feedback across both studies. It is possible that the electric shock was perceived as largely aversive and masked the ability of the socially relevant verbal feedback to elicit the expected effects for individuals with higher social anxiety specifically.

Further, both experiments recruited young, female, university students which may limit the generalisability of the results. Females were specifically recruited in these experiments due to social anxiety being more prevalent in females compared to males (Remes et al., 2016; Sosic et al., 2008), as well as the female identities and voices used as CS and US within both tasks. It is important to note, therefore, that the findings from these experiments can only be interpreted in relation to females and we cannot draw any conclusions about the relationship between individual differences in IU and extinction processes in males. However, while the majority of previous studies that have examined the role of IU in threat extinction have recruited samples of both males and females, the number of female participants generally outweighs the number of male participants in these studies (Dunsmoor et al., 2015; Lucas et al., 2018; Morriss et al., 2015; 2016; 2020; Morriss & van Reekum, 2019). It could, therefore, be argued that across the literature, the effect of IU on extinction

processes is being driven by women. Future work should examine whether there are gender differences in the effect of IU on extinction learning and retention.

The majority of previous research investigating the effect of social anxiety on extinction processes has examined between-group differences. However, we adopted a dimensional approach in this research. We measured social anxiety as a continuous variable due to the push, backed up by strong empirical support, towards a dimensional approach to psychopathology (Shear et al., 2007). While we recruited appropriately large sample sizes in both experiments to support the use of this approach, we cannot rule out that effects of social anxiety may have been observed had we compared clinically diagnosed group with a control group.

In conclusion, individual differences in IU, and not social anxiety, predicted threat expression during extinction learning and extinction retention within two separate social conditioning experiments. These results provide insight into how IU modulates threat and safe associations in the presence and absence of threat; however, social anxiety does not seem to be related to compromised safety learning in this context. The current work supports the need to disentangle transdiagnostic and disorder-specific vulnerabilities in emotional disorders (Shihata et al., 2016), and further work in this area should examine how compromised cognitive processes, specifically related to social anxiety, affect the efficacy of exposure therapy for the treatment of SAD.

Acknowledgements

The authors wish to thank participants for taking part in these experiments, Josephine Arthur and Natalie Manningham for their assistance collecting data. This research was funded by an Economic and Social Research Council (ESRC) Doctoral Studentship awarded to Shannon Wake (1791652). Furthermore, this research was

supported by: (1) a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (27567) and an ESRC New Investigator Grant (ES/R01145/1) awarded to Jayne Morriss.

Author Contributions

S.W., C.v.R., T.J. & HD designed the studies. SW collected the data, conducted the analysis and interpretation and wrote the manuscript draft. JM helped with physio data extraction protocols/scripts and statistics for analysis. J.M., C.v.R., and HD contributed to interpretation, critical manuscript revision and approval of final manuscript. T.J. edited and approved the final manuscript.

Data Transparency and Conflicts of Interest Statements

The dataset reported here is not part of any published or currently in press works. The authors have no competing interests to declare.

References

Ahrens, L. M., Mühlberger, A., Pauli, P., & Wieser, M. J. (2014). Impaired visuocortical discrimination learning of socially conditioned stimuli in social anxiety. *Social Cognitive and Affective Neuroscience*, *10*(7), 929–937.

Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T., & Wieser, M. J. (2016). Fear conditioning and stimulus generalisation in patients with social anxiety disorder. *Journal of Anxiety Disorders*, *44*, 36–46.

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

Barrett, J., & Armony, J. (2009). Influence of trait anxiety on brain activity during the acquisition and extinction of aversive conditioning. *Psychological Medicine, 39*(2), 255-265.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300.

Ben-Shakhar, G. Standardization within individuals: A simple method to neutralise individual differences in skin conductance. *Psychophysiology*, 22, 292–299 (1985).

Boelen, P. A., & Reijntjes, A. (2009). Intolerance of uncertainty and social anxiety. Journal of Anxiety Disorders, 23, 130-135.

Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: questions and answers for the DSM-V. Depression and anxiety, 27(2), 168-189.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin,* 114(1):80-99.

Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: Psychometric properties of the English version. *Behaviour Research and Therapy*, *40*(8), 931–945.

Carleton, R. N., et al. (2010). "It's not just the judgements - it's that I don't know": Intolerance of uncertainty as a predictor of social anxiety. Journal of Anxiety Disorders, 24, 189-195.

Carleton, R. N. (2016a). Fear of the unknown: One fear to rule them all? *Journal of Anxiety Disorders,* 41, 5–21.

Carleton, R. N. (2016b). Into the unknown: A review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders,* 39, 30–43.

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), Social Phobia: Diagnosis, Assessment, and Treatment (69-93). New York, NY: Guilford Press.

Connor, K., Davidson, J., Churchill, L., Sherwood, A., Weisler, R., & Foa, E. (2000). Psychometric properties of the Social Phobia Inventory (SPIN): New self-rating scale. *British Journal of Psychiatry, 176*(4), 379-386. doi:10.1192/bjp.176.4.379

Craske, M. G. (2003). *Origins of Phobias and Anxiety Disorders: Why More Women Than Men?* California, USA: Elsevier.

Craske, M. G., & Rachman, S. J. (1987). Return of fear: Perceived skill and heart rate responsivity. *British Journal Clinical Psychiatry*, 26:187-200.

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximising exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23.

Dawson, M. E., Schell, A. M., Hazlett, E. A., Nuechterlein, K. H., & Filion, D. L. (2000). On the clinical and cognitive meaning of impaired sensorimotor gating in schizophrenia. *Psychiatry Research*, *96*(3), 187–197.

Delgado, M. R., Nearing, K. I., LeDoux, J. E., & Phelps, E. A. (2008). Neural Circuitry Underlying the Regulation of Conditioned Fear and Its Relation to Extinction. *Neuron*, *59*(5), 829–838.

Dugas, M. J., Buhr, K., & Ladouceur, R. (2004). The role of intolerance of uncertainty in etiology and maintenance of generalised anxiety disorder. In R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), Generalised anxiety disorder: Advances in research and practice (pp. 143–163). New York, NY: Guilford Press

Dunsmoor, J. E., Campese, V. D., Ceceli, A. O., LeDoux, J. E., & Phelps, E. A. (2015). Novelty-Facilitated Extinction: Providing a Novel Outcome in Place of an Expected Threat Diminshes Recovery of Defensive Responses. *Biological Psychiatry*, *78*(3), 203–209.

Foa, E., Franklin, M., Perry, K., & Herbert, J. (1996). Cognitive biases in generalised social phobia. *Journal of abnormal psychology*, *105*(3), 433-439.

Foa, E. B., Steketee, G., & Rothbaum, B. O. (*1989*). Behavioral/cognitive conceptualisations of post-traumatic stress disorder. *Behavior Therapy,* 20(2), 155–176.

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17*(6), 791-802.

Gazendam, F. J., Kamphuis, J. H., & Kindt, M. (2013). Deficient safety learning characterises high trait anxious individuals. *Biological Psychology, 92*(2), 342–352.

Grady, A. K., Bowen, K. H., Hyde, A. T., Totsch, S. K., & Knight, D. C. (2016). Effect of continuous and partial reinforcement on the acquisition and extinction of human conditioned fear. *Behavioral Neuroscience*, 130(36).

Graham, B. M., & Milad, M. R. (2011). The study of fear extinction: implications for anxiety disorders. *The American journal of psychiatry*, *168*(12), 1255–1265.

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience,* 14, 488–501.

Haaker, J., Lonsdorf, T. B., Schümann, D., Menz, M., Brassen, S., Bunzeck, N., Gamer, M., & Kalisch, R. (2015). Deficient inhibitory processing in trait anxiety: Evidence from context-dependent fear learning, extinction recall and renewal. *Biological Psychology, 111,* 65 -72.

Heimberg, R. G., Brozovich, F. A., & Rapee, R. M. (2014). A Cognitive-Behavioural Model of Social Anxiety Disorder. In *Social Anxiety* (Third Edit).

Hermann, C., Ziegler, S., Birbaumer, N., & Flor, H. (2002). Psychophysiological and subjective indicators of aversive pavlovian conditioning in generalised social phobia. *Biological Psychiatry*, *52*(4), 328–337.

Hirsch, C. R., & Clark, D. M. (2004). Information-processing bias in social phobia. *Clinical Psychology Review*, *24*(7), 799–825.

Hofmann, S.G., Smits, J.A., 2008. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomised placebo-controlled trials. *Journal Clinical Psychiatry,* 69(4).

Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: A comprehensive model and its treatment implications. *Cognitive Behaviour Therapy*, *36*(4), 193–209.

Jenkins, W. O., & Stanley, J. C., Jr. (1950). Partial reinforcement: A review and critique. *Psychological Bulletin*, 47, 193.

Jorstad-Stein, E.C., & Heimberg, R.G. (2009). Social phobia: an update on treatment. *Psychiatric Clinics of North America,* 32(3), 641-663.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1-16.

Leonard, D. W. (1975). Partial reinforcement effects in classical aversive conditioning in rabbits and human beings. *Journal of Comparative and Physiological Psychology,* 88, 596.

Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S., & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *American Journal of Psychiatry*, *165*(1), 124–132.

Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans - Biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience and Biobehavioral Reviews,* 80, 703-728.

Lucas, K., Luck, C. C., & Lipp, O. V. (2018). Behaviour Research and Therapy Novelty-facilitated extinction and the reinstatement of conditional human fear. *Behaviour Research and Therapy*, *109*, 68–74.

Ma, D.S., Correll, J. & Wittenbrink, B. Behav Res (2015) 47: 1122. The Chicago face database: A free stimulus set of faces and norming data.

Macdonald, B., Wake, S., & Johnstone, T. (2020). Selective Extinction Through Cognitive Evaluation: Linking Emotion Regulation and Extinction. *European Journal Neuroscience.* Accepted Author Manuscript.

Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: ten years of progress. *Annual review of psychology*, *63*, 129–151.

Morriss, J. (2019). What do I do now? Intolerance of uncertainty is associated with discrete patterns of anticipatory physiological responding to different contexts. Psychophysiology, 56(9).

Morriss, J., Chapman, C., Tomlinson, S., & van Reekum, C. M. (2018). Escape the bear and fall to the lion: The impact of avoidance availability on threat acquisition and extinction. *Biological Psychology*, *138*(August), 73–80.

Morriss, J., Hoare, S., & Reekum, C. M. van Reekum. (2018). It's time: A commentary on fear extinction in the human brain using fMRI. *Neuroscience and Biobehavioral Reviews*, *94*(May), 321–322.

Morriss, J., Christakou, A., & van Reekum, C. M. (2015). Intolerance of uncertainty predicts fear extinction in amygdala-ventromedial prefrontal cortical circuitry. *Biology of Mood and Anxiety Disorders*, *5*(1).

Morriss, J., Christakou, A., & van Reekum, C. M. (2016). Nothing is safe: Intolerance of uncertainty is associated with compromised fear extinction learning. *Biological Psychology*, *121*, 187–193.

Morriss, J., Christakou, A., & van Reekum, C. M. (2019). Multimodal evidence for delayed fear extinction in adolescence and young adulthood. *Scientific Reports*, *9*(7748).

Morriss, J., Macdonald, B., & Van Reekum, C. M. (2016). What is going on around here? Intolerance of uncertainty predicts threat generalisation. *PLoS ONE*, *11*(5).

Morriss, J., Saldarini, F., & Van Reekum, C. M. (2019). The role of threat level and intolerance of uncertainty in extinction. International Journal of Psychophysiology, 142, 1-9.

Morriss, J., & van Reekum, C., M. (2019). I feel safe when I know: Contingency instruction promotes threat extinction in high intolerance of uncertainty individuals. *Behavior Research and Therapy,* 116, 111-118.

Morriss, J., Wake, S., Lindner, M., McSorley, E., & Dodd, H. (2020). How many times do I need to see to believe? The impact of intolerance of uncertainty and exposure experience on safety-learning and retention in young adults. *International Journal of Psychophysiology*.

Olsson, A., Carmona, S., Downey, G., Bolger, N., & Ochsner, K. N. (2013). Learning biases underlying individual differences in sensitivity to social rejection. *Emotion*, *13*(4), 616–621.

Pavlov, I. (1927). Conditioned reflexes. New York, NY, US: Dover.

Pejic, T., Hermann, A., Vaitl, D., & Stark, R. (2013). Social anxiety modulates amygdala activation during social conditioning. *Social Cognitive and Affective Neuroscience*, *8*(3), 267–276.

Ponniah, K., & Hollon, S.D., (2008). Empirically supported psychological interventions for social phobia in adults: a qualitative review of randomised controlled trials. *Psychological Medicine,* 38, 3-14.

Rabinak, C. A., Mori, S., Lyons, M., Milad, M. R., & Phan, K. L. (2016). Acquisition of CS-US contingencies during Pavlovian fear conditioning and extinction in social

anxiety disorder and posttraumatic stress disorder. *Journal of Affective Disorders*, *207*(June 2016), 76–85.

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, *35*(8), 741–756.

Raymond, J. G., Steele, J. D., & Seriès, P. (2017). Modeling Trait Anxiety: From Computational Processes to Personality. *Frontiers in Psychiatry*, *8*, 1-1.

Reichenberger, J., Porsch, S., Wittmann, J., Zimmermann, V., & Shiban, Y. (2017). Social fear conditioning paradigm in virtual reality: Social vs. electrical aversive conditioning. *Frontiers in Psychology*, *8*(NOV), 1–15.

Remes, O., Brayne, C., van, d. L. R., & Lafortune, L. (2016). A systematic review of reviews on the prevalence of anxiety disorders in adult populations. *Brain and Behavior*.

Sachs, G., Anderer, P., Doby, D., Saletu, B., & Dantendorfer, K. (2003). Impaired conditional discrimination learning in social phobia. Neuropsychobiology, *47*, 66–72.

Shear, M. K., Bjelland, I., Beesdo, K., Gloster, A. T., & Wittchen, H. U. (2007). Supplementary dimensional assessment in anxiety disorders. International Journal of Methods in Psychiatric Research, 16(S1), S52-S64.

Shiban, Y., Reichenberger, J., Neumann, I. D., & Mühlberger, A. (2015). Social conditioning and extinction paradigm: A translational study in virtual reality. *Frontiers in Psychology*, *6*.

Shihata, S., McEvoy, P. M., Mullan, B. A., & Carleton, R. N. (2016). Intolerance of uncertainty in emotional disorders: What uncertainties remain? *Journal of Anxiety Disorders*, *41*, 115–124.

Smith, I., Landesman, S., & Failure, M. T. (2009). Relapse of successfully treated anxiety and fear: theoretical issues and recommendations for clinical practice. *Australian and New Zealand Journal of Psychiatry*, *43*(2), 89–100.

Sosic, Z., Gieler, U., & Stangier, U. (2008). Screening for social phobia in medical in- and outpatients with the German version of the Social Phobia Inventory (SPIN). *Journal of Anxiety Disorders*, *22*(5), 849-859.

Spence, S. H., & Rapee, R. M. (2016). The etiology of social anxiety disorder: An evidence-based model. *Behaviour Research and Therapy*, *86*, 50–67.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stein, M. B., Torgrud, L. J., & Walker, J. R. (2000). Social phobia symptoms, subtypes, and severity: findings from a community survey. Archives of general psychiatry, 57(11), 1046-1052.

Taylor, C. T., & Alden, L. E. (2010). Safety behaviors and judgmental biases in social anxiety dis- order. Behaviour Research and Therapy, 48, 226–237.

Tinoco-González, D., Fullana, M. A., Torrents-Rodas, D., Bonillo, A., Vervliet, B., Pailhez, G., … Torrubia, R. (2015). Conditioned Subjective Responses to Socially Relevant Stimuli in Social Anxiety Disorder and Subclinical Social Anxiety. *Clinical Psychology and Psychotherapy*, *22*(3), 221–231.

Vervliet, B., Craske, M. G., and Hermans, D. (2013). Fear extinction and relapse: state of the art. Annu. Rev. Clin. Psychol. 9, 215–248.

Vervliet, B., & Indekeu, E. (2015). Low-Cost Avoidance Behaviors are Resistant to Fear Extinction in Humans. *Frontiers in Behavioral Neuroscience*, *9*, 1–12.

Wells, A., & Papageorgiou, C. (1998). Social phobia: effects of external attention on anxiety, negative beliefs, and perspective taking. Behavior Therapy, *29,* 357–370.

Wiggert, N., Wilhelm, F. H., Boger, S., Georgii, C., Klimesch, W., & Blechert, J. (2017). Social Pavlovian conditioning: Short- and long-term effects and the role of anxiety and depressive symptoms. *Social Cognitive and Affective Neuroscience*, *12*(2), 329–339.

Wong, Q. J. J., & Rapee, R. M. (2016). The aetiology and maintenance of social anxiety disorder: A synthesis of complimentary theoretical models and formulation of a new integrated model. *Journal of Affective Disorders*, *203*, 84–100.

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.