J Psychopathol Behav Assess (2012) 34:323–331
DOI 10.1007/s10862-012-9284-7

Trait Anxiety and Fear Responses to Safety Cues: Stimulus Generalization or Sensitization?

Anneke D. M. Haddad · David Pritchett · Shmuel Lissek · Jennifer Y. F. Lau

Published online: 22 April 2012

Abstract Abnormal fear responding to threat cues may contribute to the aetiology and maintenance of persistent fears and pathological anxiety. Chronic anxiety may also involve abnormal fear responding to 'safety' cues, which do not signal danger. Yet investigations of fear responding to acquired safety cues are scarce and the basis of such responding remains unclear. Moreover, previous studies do not distinguish between stimulus generalization (an associative mechanism based on perceptual similarity between threat and safety cues) and sensitization (a non-associative mechanism whereby fear responses to *any* novel, intense, or fear-related stimulus are temporarily elevated). This study investigated responses to acquired safety cues in volunteers with varying trait anxiety, using a novel fear conditioning paradigm designed to distinguish between effects of trait anxiety on generalization and sensitization. The paradigm used three conditioned stimuli: a threat cue (CS+) and two safety cues (CS−), one perceptually similar to the CS+ and one perceptually dissimilar. Conditioned fear to these cues was indexed by fear potentiation of the startle blink reflex, skin conductance responses, and self-report. To examine how trait anxiety moderated responses to safety cues, participants were divided into high and low trait anxiety subgroups. Startle, skin conductance, and self-reported fear measures indicated that generalization of fear occurred for the safety cue which resembled the threat cue, but not for the perceptually dissimilar safety cue, consistent with the stimulus generalization hypothesis. There was some evidence that stimulus generalization was exaggerated in anxious individuals. The current study sheds light on the mechanism by which fear responses to safety cues arise in healthy individuals, and offers some insight into the influence of this mechanism in chronic anxiety.

Keywords Fear conditioning · Trait anxiety · Generalization · Sensitization

A. D. M. Haddad (✉) · D. Pritchett · J. Y. F. Lau
Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK
e-mail: anneke.haddad@psy.ox.ac.uk

S. Lissek
Department of Psychology, University of Minnesota, Minneapolis, MN, USA

The ability to detect threat cues is crucial. When a cue reliably signals imminent danger, it will come to elicit escape behaviours so that the danger is avoided. Yet the ability to detect 'safety' cues is also crucial, because some stimuli that resemble threat cues do not actually indicate danger. For example, many spiders are completely harmless to humans, even though they look like venomous spiders. Similarly, panic attacks are not intrinsically dangerous despite their superficial similarity to heart attacks. Finally, engaging in social interaction is normally a neutral or positive experience, even though some interactions may resemble past experiences that were embarrassing or humiliating. Anxious responding and avoiding such 'safety' cues as if they were threat cues could hinder survival through, for example, expending energy on unnecessary escape behaviours or failing to benefit from opportunities or resources presented by safety cues. Thus an inability to distinguish accurately between safety cues and threat cues could result in considerable, enduring fears and avoidance of cues which do not actually signal danger. This transfer of fear from a threat cue to a safety cue may have crucial implications for understanding the chronic fears and worries that characterise not only many anxiety disorders, but also persistent fears and worries that are more widespread in the general

population, and which are also impairing and costly. Yet despite the importance of stimulus generalization, and despite a long history of studies examining stimulus generalization in animals (e.g., Honig and Urcuioli 1981) and in humans (e.g., Grant and Schiller 1953; Hovland 1937), studies examining how this process relates to anxiety levels in humans are scarce.

Fear responses to threat and safety cues are typically tested with discriminative fear conditioning paradigms in which participants are presented with two neutral cues: a threat cue, or a reinforced conditioned stimulus (CS+), which is paired with the aversive unconditioned stimulus (UCS); and a safety cue (CS−), which is never paired with the UCS. Although safety cues are commonly employed in this manner as a control condition, few studies have explored individual differences in fear responses to safety cues and questions remain as to the factors that strengthen or diminish such responses. One important factor is the perceptual similarity of the safety cue to the threat cue. To investigate this, a recent study used a discriminative fear conditioning paradigm to assess fear responses to multiple safety cues which varied in their physical resemblance to the CS+ (Lissek et al. 2008). The threat and safety cues were small and large circles (counterbalanced) and following conditioning, fear (as measured by fear-potentiated startle) of circles with intermediate diameters was assessed. This yielded a curvilinear (quadratic) relationship between fear-potentiated startle response to a given stimulus and its perceptual similarity to the threat cue, suggesting that physical resemblance is a crucial factor in eliciting fear responses to safety cues. A number of other studies have also yielded evidence for generalization on the basis of perceptual similarity (Dunsmoor et al. 2009; Vervliet et al. 2005, 2006, 2010). Moreover, generalization based on physical characteristics appears to be stronger amongst safety cues that are more fear-relevant (Dunsmoor et al. 2009).

Another emerging question of interest is the degree to which anomalies in stimulus generalization can explain clinical anxiety disorders and persistent fears and worries. Assessing fear responses to multiple safety cues using discriminative fear conditioning has shown that individuals with clinical anxiety disorders show elevated generalization of fear from threat to safety cues (Lau et al. 2008; Lissek et al. 2010). In one of these studies, anxious individuals reported greater fear across the CS+ and CS− compared to non-anxious individuals (Lau et al. 2008). One explanation for these results is poorer discrimination between threat and safe cues amongst anxious individuals. In the second study, anxious group differences in psychophysiological fear emerged as different patterns of response to threat and safety cues in each group (Lissek et al. 2010). Specifically, whereas healthy controls showed precipitous, quadratic declines in fear responding to stimuli of increasing dissimilarity to the CS+, individuals with panic disorder had more gradual, linear declines indicating more transfer of conditioned fear to stimuli approximating the CS+.

These early findings require further replication, but also raise crucial questions. Firstly it is still unclear whether elevated fear responses to safety cues do in fact arise through stimulus generalization (i.e., an associative mechanism based on perceptual similarity between the CS+ and the CS−). An alternative explanation is that they arise through sensitization: a non-associative mechanism whereby exposure to an aversive stimulus results in a temporary, general elevation of fear responses to any novel, intense, or fear-related stimulus. Secondly, despite the emerging corpus of data in healthy individuals, previous studies have not examined individual differences in anxiety. Yet understanding how fear conditioning and generalization are moderated by individual differences in trait anxiety is important because, as noted earlier, trait anxiety is associated with persistent fears and worries which can be disabling and costly.

The primary aim of the present study was to distinguish between the stimulus generalization and sensitization hypotheses of fear responding to safety cues, using a novel discriminative fear conditioning paradigm in individuals with a range of levels of persistent fears and worries, as assessed by a measure of trait anxiety. One way to distinguish between sensitization and generalization explanations is to compare responses to safety cues which are perceptually similar to the threat cue with responses to safety cues which are perceptually dissimilar. According to the stimulus generalization hypothesis, fear of a safety cue should be proportional to its perceptual similarity to the threat cue. According to the sensitization hypothesis, participants should be indiscriminately fearful of all safety cues, irrespective of their similarity to the threat cue. Thus in our paradigm we examined responses to a threat cue and two safety cues. In contrast to the paradigm used by Lissek and colleagues (2008, 2009), we used photographs of human faces as the threat cue and UCS. Unlike geometric shapes, faces are highly salient stimuli with strong biological and social significance for humans. The UCS was a photograph of a fearful face (of the same person as shown in the CS+ photograph), presented in conjunction with a piercing scream. Neither of the two safety cues were ever paired with the scream. One of our safety cues was perceptually similar to the threat cue (CS+), comprising a photograph of a second face, while the other was perceptually dissimilar, comprising a grey oval. In contrast to a prior study conducted by Dunsmoor and colleagues (2009), our perceptually-similar CS− represented a different individual, rather than the same individual wearing a more fear-intense expression. Again, we did this to assess mechanisms that may underlie the emergence of real-life persistent fears and worries (e.g., worrying about befriending a

new acquaintance because he resembles someone who is a bully).

In line with previous studies, we expected that all participants would acquire a conditioned fear response to the threat cue (CS+). In accordance with the stimulus generalization hypothesis and contrary to the sensitization hypothesis, we predicted that participants would develop greater conditioned fear of the perceptually similar safety cue (i.e., the other neutral face) than of the perceptually dissimilar safety cue (i.e., the oval). We hypothesized that these differences would not be evident prior to learning but would persist even after extinction. In addressing these questions, we assessed whether differential learning emerged gradually across time by investigating stimulus differences in each block. We also made specific predictions about the effect of anxiety. On the basis of stimulus generalization accounts, we expected that anxious participants' fear responses to safety cues which were similar to the threat cue would be elevated, but those to dissimilar safety cues would not. Based on prior studies, we anticipated that this could occur as between-group differences in fear responses to threat and safety cues, and/or more subtly as different patterns of fear responding to threat and safety cues in high versus low anxious individuals.

Method

Participants

The participants comprised 50 adults (26 males, 24 females; M_{age}=30.7, SD=13.6) recruited from the local community. Participants were literate in English. Information sheets distributed during recruitment asked potential participants not to take part if they had any history of psychiatric illness or a learning difficulty. The study was approved by the local ethics committee, and all participants gave written informed consent to participation. All participants were paid £5 for taking part.

Trait Anxiety

The trait scale of the Spielberger State-Trait Anxiety Inventory (STAI, Spielberger et al. 1983) was used to assess common anxiety symptoms. Participants were divided into low and high trait anxiety subgroups, based on a median split on this scale (high anxiety group: M=45.61, SD=7.43; low anxiety group: M=31.92, SD=3.78).

Psychophysiological Apparatus

Stimulation and recording were controlled by a commercial system (Contact Precision Instruments, London, UK). Electromyography (EMG) recordings of the eye-blink startle reflex were made with two 4 mm Ag-AgCl electrodes placed under the left eye. Skin conductance responses (SCR) were recorded with two 6 mm Ag-AgCl electrodes filled with KY Jelly and placed on the medial phalange of the first and second fingers of the non-dominant hand. Amplifier bandwidth was set to 25–500 Hz and data were sampled at 1,000 Hz. The air-puff startle probe was a 40 ms puff of medical grade compressed air, delivered to the centre of the forehead through a polythene tube (2 m long, 32 mm inside diameter), affixed approximately 1 cm from the skin by way of a headpiece worn by the participant. Using this headpiece allowed for head movements while maintaining constant placement of the air-puff. A visor was positioned between the polythene tube and the participant's eyes to prevent the air-puff from reaching the cornea. A solenoid valve with an AC switch controlled delivery of the air-puffs. Prior to testing, air pressure was set at 0.3, 0.5, or 0.7 bar (measured at the level of the regulator). Pressure was set at the minimal level required to elicit six blinks during a test block of eight successive startle probe presentations.

Conditioning Paradigm

A classical, discriminative fear conditioning paradigm based on a previous task developed by our group was employed (Lau et al. 2008). This comprised pre-conditioning, conditioning, and extinction phases. Photographs of female faces with neutral expressions served as the threat cue (CS+) and similar safety cue (similar CS−), with one of the two faces selected at random to serve as the threat cue for each participant (see Lau et al. 2008 for the photographs). The dissimilar safety cue (dissimilar CS−) was a grey oval of similar proportions to the female faces. The UCS was a 3 s presentation of the CS+ face displaying a fearful expression, presented simultaneously with a 95 Db shrieking female scream lasting 1 s.

CSs were presented in a fixed pseudorandom order, with no more than two trials of the same type occurring consecutively. Pre-conditioning consisted of two trials of each CS type, all occurring in the absence of the UCS. Conditioning (36 trials) comprised 12 trials of each CS type, with nine CS+ of the presentations co-terminating with UCS delivery (75 % reinforcement schedule). This 75 % reinforcement schedule for the CS+ was used because partial reinforcement prevents habituation to the UCS (Mackintosh 1974). Extinction (27 trials) consisted of nine trials of each CS type, all presented in the absence of the UCS. Each trial comprised presentation of the given CS for 8 s, with the startle probe occurring 4 or 5 s after CS onset; on reinforced trials, the UCS was presented for 3 s immediately after the CS+. These trials were intermixed with inter-trial-interval (ITI) assessments, which provided a baseline measure of

startle. The ITI varied between 11 s and 13 s, resulting in inter-probe intervals of 18–25 s.

Procedure

After giving informed consent, participants completed the STAI-T. The EMG electrodes, SCR electrodes, and startle headpiece were then fitted, and air-puff pressure was set at an appropriate level (see apparatus and materials). Next, six startle probes (with an inter-probe interval of 18–25 s) were delivered to habituate the startle reflex, before the fear conditioning task began. The task consisted of seven blocks presented in a single session. Pre-conditioning was completed in a single block, while conditioning and extinction were each divided into three blocks (i.e., conditioning comprised 3 blocks of 12 trials, with each block containing 4 trials of each CS type; extinction comprised 3 blocks of 9 trials, with each block containing 3 trials of each CS type). Upon completion of each block, participants rated their fear of each CS in a booklet using a 9-point Likert scale (anchored with the labels *not at all scar*y and *extremely scary*).

Data Processing and Analysis

Startle EMG was rectified and smoothed (20 ms moving window average). The onset latency window for the blink reflex was 20–100 ms following startle probe onset. Trials in which blinks began before or after this window were discarded; these blinks were not deemed to represent genuine startle responses to the airpuff probe. Excessively noisy trials were also discarded. Peak blink amplitude within 150 ms of startle probe onset was determined. The mean EMG level for the 50 ms preceding the onset latency window (i.e. baseline) was subtracted from the peak amplitude, generating the critical outcome measure for each trial. Raw EMG magnitudes were standardized using within-subject T-score conversions. Participants who had valid blinks for fewer than 50 % of trials ($n=3$) were excluded from the EMG analyses. For each participant, mean standardized blink amplitude was calculated for each stimulus class (CS+, similar CS−, dissimilar CS−, ITI) in each phase (pre-conditioning, conditioning, extinction). Mean fear-ratings for the three conditioned stimuli (CS+, similar CS−, dissimilar CS−) were also calculated for each phase.

Skin conductance responses were measured as the maximal deflection occurring within a time window 1–4 s after the onset of each CS. Responses were detected automatically using PsyLab and checked manually for artefacts. Participants with no non-zero responses ($n=10$) were excluded from further analysis, leaving 40 participants in the sample for this outcome measure. For each participant, SCR values were square-root transformed and then range-corrected (by dividing by their largest response amplitude) in order to account for inter-individual differences in variability unrelated to psychological processes. Preliminary analyses indicated that there were differences in SCR responses to the three trial types at baseline; however, examination of the trial-by-trial data indicated that this was an artefact arising because SCRs for the first trial after habituation (which was always a dissimilar CS− trial) were significantly larger than those for all subsequent pre-conditioning trials (all p values <0.001). This trial was therefore excluded from further analyses.

We analyzed the effects of stimulus and anxiety group in each phase separately; in addition, in order to examine the effects of learning we added block as a factor to these analyses for the conditioning and extinction phases. All statistical analyses were performed using SPSS 17.0 and an alpha level of 0.05 was used throughout. Greenhouse-Geisser corrected values were used where the assumption of sphericity was violated.

Prior to the main EMG analysis, startle magnitude during habituation trials was compared between the high and low anxiety groups. There was no significant main effect of anxiety group on startle magnitude, nor was there a significant correlation between STAI-T score and startle magnitude during habituation. Thus any differences between the anxiety groups were not due to general elevation of fear-potentiated startle prior to the conditioning task.

On average, participants rated the blonde face as significantly scarier than the brunette during the pre-conditioning phase, $t(49)=4.32$, $p<0.001$. It is unclear why this occurred. However, we do not believe this affected our results, because: (1) this difference did not persist following the conditioning phase, $t(49)=1.58$, *ns*; (2) there was no difference between the task versions (blonde as CS+ *vs.* brunette as CS+) in terms of post-conditioning fear ratings of the CS+, $t(48)=1.31$, *ns*, or the similar CS−, $t(48)=1.54$, *ns*; (3) the proportion of participants who saw the blonde face as the CS+ did not vary between anxiety groups: low anxiety: 11/24; high anxiety: 13/26; $\chi^2(1)=0.77$, *ns*; (4) during pre-conditioning, the average blink amplitudes elicited by the two faces were very similar, blonde: $M=51.4$, $SD=8.67$; brunette: $M=53.46$, $SD=11.28$; $t(46)=1.12$, *ns*; and (5) during pre-conditioning, the average SCR amplitudes elicited by the two faces were very similar, blonde: $M=21.87$, $SD=27.73$; brunette: $M=19.12$, $SD=21.64$; $t(39)=0.59$, *ns*.

Results

Self-Reported Fear Ratings

Self-reported fear ratings to each stimulus for individuals in the low and high anxious groups are presented in Fig. 1, for pre-conditioning, conditioning, and extinction phases.

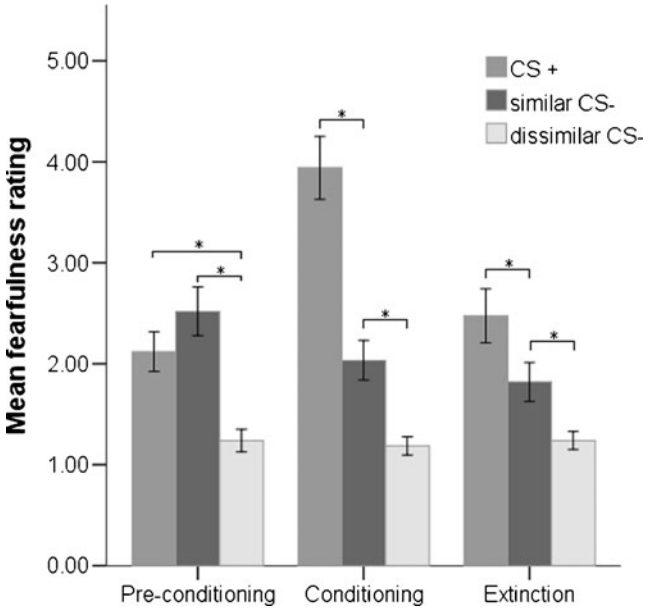

Fig. 1 Self-reported fear to the CS+, similar CS−, and dissimilar CS−, in pre-conditioning, conditioning, and extinction. Values represent means ± SEM

During **pre-conditioning**, there was a significant main effect of CS-type on fear ratings, $F(2,96)=20.02$, $p<0.001$, partial $\eta^2=0.294$. Participants reported greater fear of each of the two face stimuli (CS+ [$M=2.12$, $SD=1.39$]; similar CS− [$M=2.52$, $SD=1.70$]) relative to the oval (dissimilar CS− [$M=1.24$, $SD=0.80$]), both p values <0.001. The difference in self-reported fear between the CS+ and the similar CS− (i.e., between the two faces) did not reach significance, $t(49)=1.91$, $p=0.063$. Moreover, there was a significant main effect of anxiety: participants in the high anxiety group gave higher fear ratings ($M=2.26$, $SD=1.02$) than those in the low anxiety group ($M=1.64$, $SD=1.02$), $F(1,48)=4.56$, $p=0.038$, partial $\eta^2=0.087$ to all stimuli.

For the self-report data from the **conditioning phase**, the 3 (block) ×3 (CS-type) ×2 (anxiety group) ANOVA yielded only a significant main effect of CS-type, $F(1.52,72.84)=54.66$, $p<0.001$, partial $\eta^2=0.532$, and a main effect of anxiety, $F(1,48)=6.50$, $p=0.014$, partial $\eta^2=0.119$. The main effect of CS-type reflected greater self-reported fear of the CS+ ($M=3.94$, $SD=2.20$) compared to both the similar CS− ($M=2.03$, $SD=1.39$), $t(49)=6.31$, $p<0.001$, and the dissimilar CS− ($M=1.19$, $SD=0.65$), $t(49)=9.06$, $p<0.001$. In turn, the similar CS− received significantly higher fear ratings than the dissimilar CS−, $t(49)=4.59$, $p<0.001$. The main effect of anxiety reflected the fact that, as with pre-conditioning, high anxious individuals gave higher fear ratings ($M=2.74$, $SD=1.03$) than low anxious individuals ($M=2.00$, $SD=1.03$) across CS-types.

A significant main effect of CS-type was also present during **extinction**, $F(1.71,82.22)=14.69$, $p<0.001$, partial $\eta^2=0.234$. Again there was greater fear of the CS+ ($M=2.47$, $SD=1.89$) relative to both the similar CS− ($M=1.82$, $SD=1.37$), $t(49)=2.88$, $p=0.006$, and the dissimilar CS− ($M=1.24$, $SD=0.64$), $t(49)=4.73$, $p<0.001$, and greater fear of the similar CS− relative to the dissimilar CS−, $t(49)=3.14$, $p=0.003$. The difference in fear ratings between the anxiety groups, collapsed across CS-types, was at trend level (high anxiety: $M=2.10$, $SD=1.02$; low anxiety: $M=1.57$, $SD=1.02$), $t(1,48)=3.44$, $p=0.070$, partial $\eta^2=0.067$.

Psychophysiological Responses

Figure 2a shows EMG responses to each stimulus class (CS+, similar CS−, dissimilar CS−) for each phase collapsed across anxiety group. Data for conditioning and extinction phases are presented by block. In **pre-conditioning** startle responses, no significant main or interaction effects of CS-type or anxiety group were present.

For the EMG data from the **conditioning phase**, a 3 (block) ×3 (stimulus) ×2 (anxiety group) ANOVA yielded significant main effects of block, $F(1.723, 70.639)=44.028$, $p<0.001$, partial $\eta^2=0.518$ and stimulus, $F(2, 82)=25.662$, $p<0.001$, partial $\eta^2=0.385$. These effects were modified by an interaction between these factors, $F(4, 164)=3.105$, $p=0.017$, partial $\eta^2=0.070$. We therefore separated subsequent analyses by block to examine responses to CS-types in each anxiety group during 'early', 'middle' and 'late' stages of conditioning. We used a 3 (stimulus) ×2 (anxiety group) ANOVA within each block. For **conditioning block 1**, there was a trend level effect of anxiety group, $F(1, 42)=3.816$, $p=0.057$, partial $\eta^2=0.083$, which arose because collapsed across stimulus types, the high anxious individuals tended to have larger EMG scores (M=56.64, SD=3.31) than the low anxious individuals (M=54.36, SD=4.35). Critically, there was no differential response to the three stimuli ($p=0.160$). For **conditioning block 2**, the ANOVA yielded a significant main effect of stimulus, $F(2, 88)=16.73$, $p<0.001$, partial $\eta^2=0.275$, which was modified by an interaction with anxiety group, $F(2, 88)=3.73$, $p=0.028$, partial $\eta^2=0.078$. Post-hoc follow-up t-tests revealed that this interaction arose because, within the low anxious group, participants' responses discriminated between the CS+ (M=55.37, SD=4.34) and the similar CS− (M=51.15, SD=6.41), $t(23)=3.19$, $p=0.004$, Cohen's $d=0.77$, as well as between the CS+ and the dissimilar CS− (M=49.10, SD=4.90), $t(23)=4.98$, $p<0.001$, Cohen's $d=1.36$, but did not discriminate between the similar and dissimilar CS−, $p=0.106$. In contrast, within the high anxious group, participants' responses did not differ between the CS+ (M=52.45, SD=4.64) and the similar CS− (M=53.11, SD=5.76), $p=0.557$. Instead the mean response to the dissimilar CS− (M=48.55, SD=4.68) was significantly different to both the

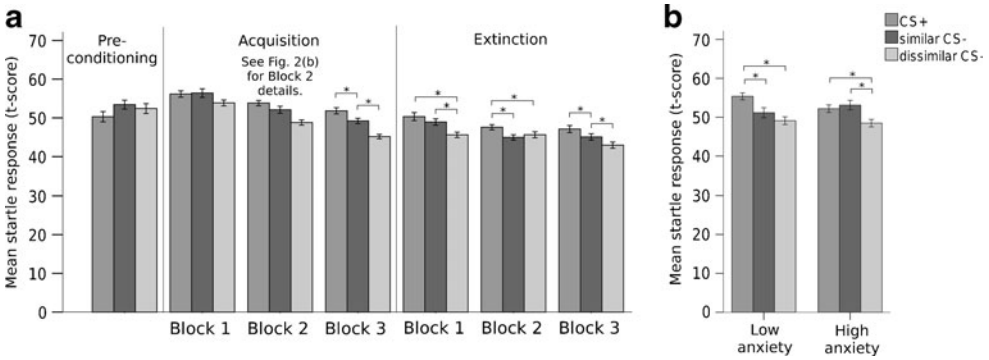

Fig. 2 EMG of fear potentiated startle responses to the CS+, similar CS−, and dissimilar CS− (**a**), during pre-conditioning, conditioning, and extinction (showing the effect of block in conditioning and extinction; see text for details); and **b** during block 2 of conditioning in the high and low anxiety groups. Values represent means ± SEM

CS+, $t(22)=3.59$, $p=0.002$, Cohen's $d=0.83$, and the similar CS−, $t(21)=2.97$, $p=0.007$, Cohen's $d=0.87$ (see Fig. 2b). Comparing the *difference* in responses to the CS+ and the similar CS− between the high and low anxious groups for this block revealed a significant difference, $t(45)=2.917$, $p=0.005$, with a large effect size of Cohen's $d=0.85$. For **conditioning block 3**, there was a significant main effect of stimulus, $F(2, 88)=24.55$, $p<0.001$, partial $\eta^2=0.358$, but no significant effects involving anxiety. Post-hoc paired t-tests showed that, collapsed across anxiety group, the participants had higher EMG responses to the CS+ (M=51.82, SD=5.45) than to the similar CS− (M=49.16, SD=4.72), $t(45)=2.909$, $p=0.006$, Cohen's $d=0.52$, which in turn received higher responses than the dissimilar CS− (M=45.17, SD=3.94), $t(46)=4.367$, $p<0.001$, Cohen's $d=0.93$.

For the EMG data from the **extinction** phase, the 3 (block) ×3 (stimulus) ×2 (anxiety group) ANOVA yielded significant main effects of block, $F(2, 88)=11.099$, $p<0.001$, partial $\eta^2=0.201$ and stimulus, $F(1.585, 69.732)=11.610$, $p<0.001$, partial $\eta^2=0.209$, as well as a trend-level intereaction between them $F(4, 176)=2.306$, $p=0.060$, partial $\eta^2=0.050$. We therefore separated subsequent analyses by block, using a 3 (stimulus) ×2 (anxiety group) ANOVA within each block. For **extinction block 1**, there was a significant main effect of stimulus, $F(1.688, 75.941)=8.772$, $p=0.001$, partial $\eta^2=0.163$, but no significant effects involving anxiety. The effect arose because collapsed across anxiety level, participants' responses did not discriminate between the CS+ (M=50.32, SD=7.14) and the similar CS− (M=48.93, SD=5.63), $p=0.162$, but did discriminate between the dissimilar CS− (M=45.61, SD=4.83) and both the CS+, $t(46)=3.45$, $p=0.001$, Cohen's $d=0.77$, and the similar CS−, $t(46)=3.11$, $p=0.003$, Cohen's $d=0.63$. For **extinction block 2**, there was again a significant main effect of stimulus, $F(2, 90)=3.784$, $p=0.026$, partial $\eta^2=0.078$, but no significant effects involving anxiety. Post-hoc tests revealed that participants' responses discriminated between the CS+ (M=47.58, SD=4.70) and the similar CS− (M=44.96, SD=4.71), $t(46)=2.723$, $p=0.009$, Cohen's $d=0.56$ as well as, at trend level, between the CS+ and the dissimilar CS− (45.64, SD=5.33), $t(46)=1.817$, $p=0.076$, Cohen's $d=0.39$, but did not discriminate between the similar and dissimilar CS−, $p=0.455$. For **extinction block 3**, there was once again a significant main effect of stimulus, $F(2, 88)=7.552$, $p=0.001$, partial $\eta^2=0.146$, but again no significant effects involving anxiety. In this block, post-hoc pairwise comparisons were all significant at least at trend level, with mean responses for the CS+ (M=47.07, SD=6.24) being at trend level higher than those for the similar CS− (M=45.07, SD=5.53), $t(45)=1.772$, $p=0.083$, Cohen's $d=0.34$, which in turn were higher than those for the dissimilar CS− (M=42.96, SD=5.51), $t(45)=2.330$, $p=024$, Cohen's $d=0.38$.

Figure 3 shows skin conductance responses (SCR) to each stimulus class (CS+, similar CS−, dissimilar CS−) for each phase collapsed across anxiety group. Again, data from the conditioning and extinction phases are presented separately by block to describe any time-related changes in the discrimination between CS-types. At **pre-conditioning**, there was no significant effect of CS-type or anxiety.

For the SCR data from the **conditioning phase**, the 3 (block) ×3 (stimulus) ×2 (anxiety group) ANOVA yielded significant main effects of block, $F(1.742, 66.209)=$,10.933 $p<0.001$, partial $\eta^2=0.233$, and stimulus $F(1.733, 65.843)=13.306$, $p<0.001$, partial $\eta^2=0.259$, as well as an interaction between them, $F(3.111, 118.214)=3.025$, $p=0.031$, partial $\eta^2=0.074$. We therefore separated subsequent analyses by block to examine the rates of learning within the two anxiety groups, using a 3 (stimulus) ×2 (anxiety group) ANOVA within each block. These three analyses showed that within each of the three conditioning blocks, a significant main effect of stimulus was present (p values≤0.012); none of these blocks showed effects involving anxiety. Post-hoc paired t-tests revealed that in **conditioning block 1**, SCRs

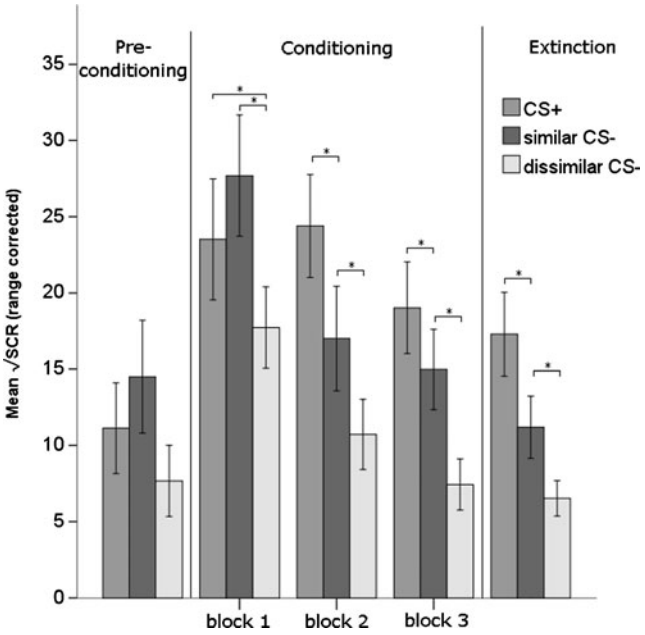

Fig. 3 Skin conductance responses to the CS+, similar CS−, and dissimilar CS−, during pre-conditioning, conditioning, and extinction (showing the effect of block in conditioning only; see text for details). Values represent means ± SEM

to the CS+ (M=23.51, SD=25.11) and similar CS− (M=27.69, SD=25.16) did not differ significantly (p=0.167), whereas both were different (at least at trend level) to SCRs to the dissimilar CS− (M=17.73, SD=16.91), CS+ > dissimilar CS−: t(39)=1.883, p=0.067, Cohen's d=0.27; similar CS−> dissimilar CS−: t(37)=2.67, p=0.011, Cohen's d=0.27. By **conditioning block 2** participants' SCRs showed differential responses to the three stimuli: responses to the CS+ (M=24.40, SD=21.36) were significantly higher than those to the similar CS− (M=17.01, SD=21.69), t(37)=3.01, p=0.005, Cohen's d=0.34, which in turn were higher, at trend level, than responses to the dissimilar CS− (M=10.72, SD=14.60), t(37)=1.920, p=0.062, Cohen's d=0.34. Similarly in **conditioning block 3**, SCRs differentiated all three stimuli: responses to the CS+ (M=19.02, SD=19.01) were higher, at trend level, than those to the similar CS− (M=14.98, SD=16.69), t(39)=1.72, p=0.094, Cohen's d=0.23, which in turn were higher than responses to the dissimilar CS− (M=7.43, SD=10.58), t(39)=3.24, p=0.002. Cohen's d=0.54.

At **extinction**, the 3 (block) ×3 (stimulus) ×2 (anxiety group) ANOVA yielded only a significant main effect of stimulus F(2, 72)=15.12, p<0.001, partial η^2=0.285. Post-hoc paired t-tests collapsed across block and anxiety group revealed that this was because due to the fact that participants' responses to all three CS types differed: responses to the CS+ (M=17.30, SD=17.42) were significantly higher than those to the similar CS− (M=11.19, SD=12.94), t(39)=3.14, p=0.003, Cohen's d=0.40, which in turn were significantly higher than responses to the dissimilar CS− (M=6.53, SD=7.38), t(39)=3.00, p=0.005, Cohen's d=0.44.

Discussion

In the present study, we developed a novel discriminative fear conditioning paradigm to elucidate the mechanism by which fear of social threat signals (those that predict danger) is extended to social safety signals (those that have never predicted danger). In contrast to several previous studies, we employed two safety cues, one of which was perceptually similar to the threat cue and one of which was dissimilar. This enabled us to distinguish between two proposed mechanisms for the emergence of fear of safety cues. Stimulus generalization, an associative mechanism, ought to have resulted in participants developing greater fear of a perceptually similar safety cue compared to a perceptually dissimilar safety cue. On the other hand, sensitization, an indiscriminate, non-associative mechanism, ought to have resulted in participants becoming equally fearful of both safety cues. After conditioning, individuals' subjective and psychophysiological responses (fear-potentiated startle and skin conductance) to the two safety cues did indeed vary depending on their similarity to the face threat cue. In general, fear responses to the perceptually similar CS− (another face) were stronger than those to the perceptually dissimilar CS− (an oval). These results are clearly consistent with the stimulus generalization hypothesis but inconsistent with sensitization and are comparable with several other studies (Dunsmoor et al. 2009; Vervliet et al. 2005, 2006, 2010 Lissek et al. 2008; Dunsmoor et al. 2009). Interestingly, the gradient of fear responses did not appear instantaneously, but rather emerged over time: differential responding was not observed in Block 1, but was clearly apparent by Block 3 in both the EMG and the SCR data. The differences persisted through the extinction phase.

Post-hoc analyses also yielded interesting patterns associated with anxiety in the EMG data. During the first block of conditioning, there was no evidence for differential fear responding to the three conditioned stimuli in either the high or the low anxious groups. In block 2, the low anxious group demonstrated higher fear-potentiated startle to the threat cue compared to the two safety cues. In contrast, fear-potentiated startle responses in the high anxious group failed to distinguish between the threat cue and perceptually similar safety cue during this block, although they did distinguish the dissimilar safety cue. During Block 3 both high and low anxiety groups discriminated between CS+ and similar CS− at approximately equal rates (and thus there was no anxiety group × stimulus interaction for

Block 3). This pattern suggests that high trait anxious individuals may have delayed discrimination learning or prolonged generalization of fear responses to perceptually similar safety cues, similar to that seen in panic disorder (Lissek et al. 2009). However, these differences were not apparent in the skin conductance data. Finally, as expected, main effects of anxiety characterized self-reported fear ratings.

Persistent anxiety is characterised by exaggerated and persistent fears to situations that resemble threats but are in fact safe. These features are suggestive of an inability to discriminate safety cues from threat cues. Our data suggest that these effects may be explained by generalization of fear among physically similar cues, such as our threat cue and our perceptually similar safety cue. Moreover, our approach extends investigations of stimulus generalization by using stimuli with clear biological and social salience. Although our findings of differential patterns of responding in the high and low trait anxious group must be considered preliminary, they resonate with those reported by other empirical studies in anxious patients and with a recent meta-analysis of simple fear conditioning studies (Lissek et al. 2005, 2010, 2009). This concluded that, compared to healthy controls, anxiety patients demonstrate increased psychophysiological arousal (e.g., skin conductance response, fear-potentiated startle) during the acquisition of conditioned fear, as well as increased conditioned responding not only to threat cues but also to safety cues. Moreover, experimental evidence in patients with panic disorder demonstrates that, unlike the quadratic relationship in healthy controls, they show a linear relationship between fear-potentiated startle response to a given stimulus and its perceptual similarity to the threat cue (Lissek et al. 2010). This results in a pattern whereby, compared to healthy controls, panic patients' fear-potentiated startle responses remain elevated for stimuli that are relatively dissimilar to the threat cue. Our data lend some support to the notion that chronic high anxiety may be also characterised by prolonged generalization of fear responses to safety cues which share perceptual features with a threat cue: amongst highly trait anxious participants, the perceptually similar safety cue elicited a fear-potentiated startle response in block 2 of conditioning that was indistinguishable from responses to the threat cue. This was not true of low trait anxious participants, whose fear-potentiated startle responses distinguished between the threat cue and similar safety cue by block 2.

What might be the basis for these exaggerated fear responses to the similar safety cue but not to the dissimilar safety cue in high anxious individuals? One explanation is that, because the similar safety cue shares some but not all the features of the threat cue, after participants have learnt the CS+ −UCS association, the similar CS- may be conceptualised as a 'degraded' threat cue. In block 2 of conditioning, this degraded threat cue may continue to convey sufficient 'danger' information to activate a fear response in high trait anxious individuals, but not in low trait anxious individuals. In other words, high trait anxious individuals may have a lower threshold for threat appraisal that is exceeded by the degraded threat content of the similar safety cue. By contrast, the partial threat information conveyed by the similar safety cue may not activate the fear system for those low in anxiety because it is insufficiently threatening, and only the threat cue itself is sufficiently threatening to exceed their threshold for threat appraisal (c.f., Lissek et al. 2010). Of course, it could be argued that the dissimilar safety cue shared some features with the CS+, in that it was a visual stimulus with a similar shape to the faces. In this view, a more definitive test of the sensitization hypothesis would involve using a more dissimilar safety cue —perhaps of a different modality. Finally, our findings cannot shed light on the question of whether high anxiety levels lead to overgeneralization or whether a propensity to overgeneralize increases the perception of the world as a threatening, unsafe place, thereby increasing an individual's tendency to be anxious. Future work could address this question using prospective designs or possibly by experimental manipulation of state anxiety levels or the tendency to generalize.

It is not clear why the differential pattern of results between high and low anxious individuals was apparent for fear potentiated startle magnitude but not for skin conductance responses or self report. We would tentatively suggest several factors that may be relevant. First, skin conductance responses may be less specific, reflecting general arousal rather than specifically fearful responses. Second, the two measures reflect physiological processes occurring at different times within each trial: skin conductance measures were taken 1–4 s following CS onset, whereas the startle probe occurred 4 or 5 s after CS onset. Future studies could examine whether the timecourse of the fear response differs between high and low anxious individuals by, for example, probing startle responses at a wider range of timepoints following CS onset. Finally, although self-report may have been subject to demand responding, there may have been a genuine dissociation between subjective appraisal of the similar safety cue and non-conscious physiological arousal when it was displayed. Tentatively, the discrepancy between our physiological and self-report measures might suggest that generalization processes are subtle, occurring with minimal awareness.

These results are subject to a number of limitations. First, the mean STAI score in the high anxious group was only moderately high (45.61; SD=7.43). STAI scores vary considerably in clinical samples; for example, Fisher and Durham (1999) report a mean pre-treatment STAI score of 57.00 (SD=9.45) across six treatment studies of Generalized

Anxiety Disorder, which is rather higher than in our high anxious group. On the other hand, Lissek and colleagues (2010) report that their sample of patients with Panic Disorder had a mean STAI score of 39.89 (SD=10.80)—somewhat lower than in our high anxious group. Thus differences between the high anxious and low anxious groups may not apply to comparisons of clinically anxious and non-anxious groups. As noted earlier, however, our findings do converge with studies in clinically anxious individuals (Lissek et al. 2005). In addition, high trait anxiety has been shown to be a risk factor for anxiety disorders and thus our data may well reflect conditioning-based vulnerability markers for clinical anxiety. However, future studies of persistent fears and worries may benefit from selecting individuals with extreme trait anxiety scores. Second, there were unexpected differences between the two versions of the task (blonde face as CS+ vs brunette face as CS+), although these differences were only apparent in the self-report data at pre-conditioning. Third, the UCS was a human scream, which may have had greater 'belongingness' to the face CSs than to the oval (Hamm et al. 1989). Future work using this paradigm could investigate whether the effects found here are specific to faces by comparing responses to this paradigm with one where the CS+ is a geometric shape (such as the grey oval), the similar CS− is a similar shape and the dissimilar CS− is a neutral face. Fourth, the complexity of the design and the use of multiple outcome measures mean that there were multiple contrasts within a small sample; it is possible, therefore, that our power to detect between-group differences was limited. However, for the contrast where we did report an interaction involving anxiety (i.e., EMG responses in block 2 of acquisition) the effect size was large.

In summary, we developed and piloted a novel paradigm for assessing online learning of threat and safety cues. Using this paradigm, we demonstrated that not all safety cues are equal in their elicitation of fear. Subjective and physiological fear responses were greater when the safety cue resembled the threat cue than when the safety cue was perceptually dissimilar to the threat cue. This is consistent with a stimulus generalization account for the persistence of fear responses to safety cues. We also found some suggestion that even in a non-clinical population, highly anxious individuals show such strong stimulus generalization that they show persistent responding to a perceptually similar safety cue as if it was a threat cue. This suggests that elevated fear of safety cues in individuals with persistent fears and worries may be mediated by stimulus generalization mechanisms.

Acknowledgements The authors wish to thank Elizabeth Michael for help with scoring the skin conductance data.

References

Dunsmoor, J. E., Mitroff, S. R., & LaBar, K. S. (2009). Generalization of conditioned fear along a dimension of increasing fear intensity. *Learning & Memory, 16*, 460–469.

Fisher, P. L., & Durham, R. C. (1999). Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychological Medicine, 29*, 1425–1434.

Grant, D. A., & Schiller, J. J. (1953). Generalization of the conditioned galvanic skin response to visual stimul. *Journal of Experimental Psychology, 46*, 309–313.

Hamm, A. O., Vaitl, D., & Lang, P. J. (1989). Fear conditioning, meaning, and belongingness: a selective association analysis. *Journal of Abnormal Psychology, 98*, 395–406.

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): twenty-five years of research on stimulus generalization. *Journal of Experimental Analysis of Behavior, 36*, 405–445.

Hovland, C. I. (1937). The generalization of conditioned responses. IV. The effects of varying amounts of reinforcement upon the degree of generalization of conditioned responses. *Journal of Experimental Psychology. General, 21*, 261–276.

Lau, J. Y. F., Lissek, S., Nelson, E. E., Lee, Y., Roberson-Nay, R., Poeth, K., et al. (2008). Fear conditioning in adolescents with anxiety disorders: results from a novel experimental paradigm. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*, 94–102.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., et al. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour Research and Therapy, 43*, 1391–1424.

Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., et al. (2008). Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. *Behaviour Research and Therapy, 46*, 678–687.

Lissek, S., Rabin, S. J., McDowell, D. J., Dvir, S., Bradford, D. E., Geraci, M., et al. (2009). Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. *Behaviour Research and Therapy, 47*, 111–118.

Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., et al. (2010). Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. *The American Journal of Psychiatry, 167*, 47–55.

Mackintosh, N. J. (1974) The psychology of animal learning. London: Academic Press

Spielberger, C. D., Gorsuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, A. G. (1983). *Manual for the state-trait anxiety inventory (Form Y)*. Palo Alto: Consulting Psychologists Press, Inc.

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

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2006). Generalization gradients for acquisition and extinction in human contingency learning. *Experimental Psychology, 53*, 132–142.

Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010). Fear generalization in humans: impact of verbal instructions. *Behaviour Research and Therapy, 48*, 38–43.