Behaviour Research and Therapy 135 (2020) 103755

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Generalization of extinction with a generalization stimulus is determined by learnt threat beliefs

Alex H.K. Wong *, Valentina M. Glück, Juliane M. Boschet, Paula Engelke

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

ARTICLE INFO

Keywords:
Generalization of extinction
Threat belief
Expectancy violation
Rule-based generalization
Fear generalization
Extinction

ABSTRACT

Expectancy violation refers to the mismatch between an expected and the actual outcome. Maximizing expectancy violation is crucial for exposure-based treatment. Since the original stimulus of fear acquisition (CS+) is rarely available, stimuli that resemble the CS+ (generalization stimuli; GSs) are presented during treatment. A given GS may evoke either strong or weak generalized fear depending on an individual's threat beliefs. Presenting this GS in extinction would then evoke different levels of expectancy violation, which determines the strength of the subsequent generalization of extinction to other stimuli, including the CS+. After differential fear conditioning, participants exhibited discrete generalization gradients depending on their inferred relational rules (Linear vs Similarity). Crucially, the Linear group showed strong generalized fear to the GS used in extinction. This strong expectancy violation led to enhanced extinction learning and subsequently to strong generalization of extinction as characterized by a flat generalization gradient, and reduced conditioned fear to the CS+. In contrast, the Similarity group showed weak generalized fear to the same GS in extinction, and limited generalization of extinction. These results corroborate the importance of expectancy violation in exposure-based treatment, and suggest that exposure sessions designed to evoke strong threat beliefs may lead to better treatment outcome.

Exposure-based therapy, a pillar of cognitive behavioural therapy, is considered one of the most effective treatments for anxiety disorders (Carpenter et al., 2018; Hofmann & Smits, 2008). Exposure-based therapy involves repeatedly presenting fear-provoking stimuli without any aversive outcome, thereby disconfirming an individual's threat belief to these fear-related stimuli. Disconfirmation of threat beliefs involves expectancy violation through a mismatch between an expected outcome and the actual outcome (Rescorla & Wagner, 1972). Expectancy violation has been proposed to be a major factor influencing exposure treatment outcome (Craske et al., 2014, 2018), hence suggesting a positive association between challenging patients' threat beliefs and treatment outcome.

In the laboratory, exposure-based therapy is modelled via fear extinction within a Pavlovian conditioning framework. Fear extinction refers to the repeated presentation of a conditioned stimulus (CS+) alone that previously signalled the occurrence of an aversive unconditioned stimulus (US). As a result of the CS - no US presentation, conditioned fear to the CS+ gradually decreases across extinction trials. The fear extinction model is proposed to be a valid laboratory model for exposure-based therapy (Scheveneels, Boddez, Vervliet, & Hermans, 2016). However, in exposure-based therapy, the exact same stimuli or situations at the time of fear acquisition (i.e., the original CS+) are unlikely to be reproduced. Therefore, cues that resemble the original fear-related stimuli are presented during exposure-based treatment. This use of extinction stimuli that are similar, but not identical to the original CS+, is equivalent to presenting a generalization stimulus (GS) that perceptually or conceptually resembles the CS+ during extinction in a laboratory model. Empirical studies have shown that extinction learning to a GS generalized weakly to the CS+ in test (e.g., Barry, Griffith, Vervliet & Hermans, 2016; Vervliet & Geens, 2014; Vervliet, Vansteenwegen, Baeyens, Hermans, & Eelen, 2005; Vervliet, Vansteenwegen, & Eelen, 2004; Vervoort, Vervliet, Bennett & Baeyens, 2004; Wong & Lovibond, 2020; Zbozinek & Craske, 2018), or to another novel GS in test (e.g., Vervliet et al., 2004; Vervoort, Vervliet, Bennett, & Baeyens, 2014; Wong & Lovibond, 2020; Zbozinek & Craske, 2018), as indicated by little to no decrease in conditioned fear to the respective test stimulus.

The weak generalization of GS extinction potentially results from

---

* Correspondence author.
E-mail address: hon.wong@uni-wuerzburg.de (A.H.K. Wong).

https://doi.org/10.1016/j.brat.2020.103755
Received 14 August 2020; Received in revised form 29 September 2020; Accepted 7 October 2020

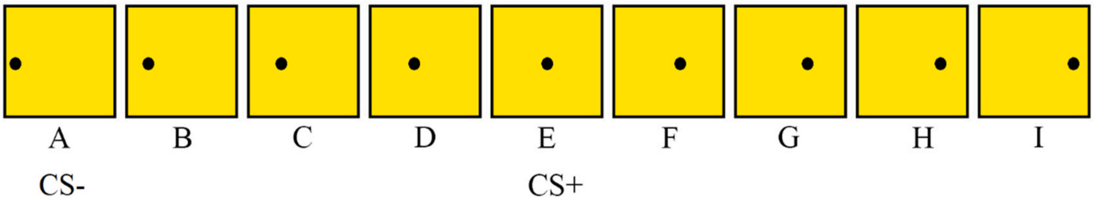

Fig. 1. Stimulus dimension; Stimulus E served as the CS + whereas stimulus A served as the CS-. The stimulus labels (A–I) were not presented to participants. The color version of this figure can be seen in the online article.

weaker activation of conditioned fear to GSs than to a CS+, due to the CS+ not sharing all its excitatory elements with the GS (i.e., generalization decrement; Mackintosh, 1974). Therefore, these unshared excitatory elements do not gain inhibitory strength during GS extinction. As a result, stimuli that consist of these unshared elements (the CS+ or other novel GSs) can still evoke fear after GS extinction. This renders a weak generalization of GS extinction effect (for details, see Blough, 1975; Pittig, Treanor, LeBeau, & Craske, 2018). Struyf, Hermans, and Vervliet (2018) thus proposed that presenting a GS that activates at least as much conditioned fear as the CS+ in extinction may produce a strong generalization of GS extinction effect. In support of this notion, the use of an asymmetrical stimulus dimension (i.e., intensity of one end of the dimension is higher than the other end) often resulted in stronger responding to a GS that was more intense than the CS+ (intensity generalization; Hull, 1949; Razran, 1949; but see Pierrel & Sherman, 1960; Zielinski & Jakubowska, 1977). For instance, responding was stronger to a tone GS with a higher decibel level than a tone CS+ (Huff, Sherman, & Cohn, 1975). Therefore, to examine whether a GS more intense than a CS+ (hence evokes more conditioned fear than a CS+ does) would lead to strong generalization of GS extinction, Struyf et al. (2018) used an intensity stimulus dimension of facial stimuli with increasingly fearful expressions. During differential Pavlovian fear acquisition, the stimulus with the least fearful expression served as the safety stimulus (CS-) whereas the stimulus in the middle of the dimension (i.e., moderately fearful expression) served as the CS+. In a subsequent extinction phase, one group received the GS with the highest intensity along the dimension (i.e., the most fearful expression at the opposite end of CS- along the stimulus dimension; peak GS), one group received the CS+, and another group received a GS intermediate of the CS+ and the CS- (i.e., intensity of fearful expression weaker than the CS+; weak GS). Both the peak GS group and the CS+ group showed stronger conditioned fear to the extinction stimulus than the weak GS group in early extinction trials. This not only supports the notion that a GS more intense than the CS+ evokes at least as much conditioned fear as the CS+, but also suggests that the extinction stimulus in the peak GS group and the CS+ group led to higher levels of expectancy violation than the weak GS group. In a final test phase, all stimuli along the dimension were presented without an US. Both the peak GS group and the CS+ group showed a flatter generalization of extinction gradient than the weak GS group. More specifically, while both the peak GS group and CS+ group showed similarly low level of fear responding to the CS+, the peak GS group showed less fear responding to the peak GS than the CS+ group. This suggests that extinction learning to a GS that evoked stronger fear than the CS+ led to greater generalization of extinction across the test stimuli.

The authors interpreted this pattern via a typicality asymmetry account, in which generalization from highly typical exemplars (e.g., bear) to other members of the same category (e.g., mammals) is stronger, whereas generalization from atypical exemplars (e.g., armadillo) to other category members is weaker (Dunsmoor & Murphy, 2014; Scheveneels, Boddez, Patrick, & Hermans, 2017). Participants may have learnt that a fearful expression signalled a shock. Thereby, the stimulus with the strongest fearful expression (the peak GS) became a highly typical exemplar of the category of fearful faces. Presenting it in extinction maximized expectancy violation and hence resulted in strong generalization of GS extinction.

The strong generalization of GS extinction found by Struyf et al. (2018) could, however, be interpreted alternatively with the formation of threat beliefs through rule-based generalization. Recent evidence demonstrated a diverse range of generalization gradients, and the shape of these gradients corresponded with participants' self-reported rules (Ahmed & Lovibond, 2019; Lee, Hayes, & Lovibond, 2018; Wong & Lovibond, 2017, 2018). For instance, participants that responded based on a linear rule (e.g., the further away the stimulus is positioned from the CS+ along the stimulus dimension in the opposite direction of the CS-, the higher the likelihood of a shock), exhibited a linear gradient. Individuals who responded based on a similarity rule (e.g., the more perceptually similar the stimulus is to the CS+, the higher the likelihood of a shock), exhibited a peaked, bell-shaped gradient with the highest fear responding to the CS+, with decreasing responding as stimuli became more dissimilar to the CS+. As a result, participants from these two rule groups would hold different threat beliefs. Specifically, they would show different generalized fear to the GS at the extreme end of the stimulus dimension in the direction opposite of the CS-: Participants who responded on a linear rule basis showed strong generalized fear to this GS, while those who responded based on a similarity rule showed weak generalized fear to the same GS. Evidence also suggests that participants are biased to respond on a linear-rule basis along an intensity stimulus dimension (e.g., Ahmed & Lovibond, 2019). Considering that Struyf et al. (2018) used such an intensity stimulus dimension, most participants might have been encouraged to respond based on a linear rule (i.e., the more fearful the facial expression, the higher the likelihood of a shock), hence forming strong generalized fear to the peak GS. Therefore, presenting the peak GS in extinction may have had produced a high level of expectancy violation, resulting in a strong generalization of GS extinction. However, given that rule-based generalization was not assessed, this alternative explanation could not be tested.

The current study therefore sought to systematically examine how different threat beliefs formed through rule-based generalization determine the magnitude of generalization of GS extinction. We used a differential conditioning procedure, with the stimulus intermediate of a non-intensity based stimulus dimension serving as the CS+ and the stimulus at the extreme left end of the dimension serving as the CS- (Fear acquisition). The full range of stimuli along the dimension was then presented to assess the generalization gradients of fear acquisition (Fear generalization test). Next, participants were presented with the GS at the extreme right end of the dimension in extinction (GS extinction). In the final test phase, generalization of GS extinction was assessed by presenting selected GSs to the participants (Post-extinction generalization test). Based on previous studies (e.g., Lee et al., 2018; Wong & Lovibond, 2017), we expected a majority of participants would form either a similarity- or a linear-based rule in the current design. This allowed us to compare how threat beliefs formed by qualitatively different generalization rules influence the effectiveness of generalization of GS extinction. We predicted that participants who responded with a linear rule would show strong generalized fear to the GS in extinction, as the GS was placed at the opposite end on the continuum to the CS-. This would lead to a high level of expectancy violation in the extinction phase, resulting in strong generalization of GS extinction. In contrast, participants who responded with a similarity rule would show weak generalized fear to the same GS in extinction; this would evoke little if any expectancy violation, thus resulting in limited generalization of extinction.

1. Method

1.1. Participants

Individuals from the University of Würzburg and the general community were recruited as participants and received either partial course credit or 10€ for participation. We followed the sample size in previous rule-based generalization studies (e.g., Lee et al., 2018; Wong & Lovibond, 2017), which was approximately 50 participants in total. To ensure appropriate sample sizes in both the similarity-based group (Similarity group) and the linear-based group (Linear group), we used a recruitment strategy to recruit an initial sample of 50 participants. We then classified participants according to their reported rules (see Results for details) and checked if there were at least 20 participants in each of the two rule groups. If not, we continued recruiting until both groups reached at least 20 participants. This led to a total recruitment of 63 participants.[1] The Ethics Committee of the Institute of Psychology at the University of Würzburg approved the experimental procedure.

1.2. Apparatus and materials

The stimulus dimension consisted of 9 yellow squares with black outline (9.5 cm × 9.5 cm) containing a black dot that varied in position from the left to the right. These stimuli were labelled A (with the dot at the left-most) to I (with the dot at the right-most), by manipulating the dot position by 0.8 cm from one stimulus to the next (see Fig. 1). Stimulus E and A always served as the CS+ and CS-, respectively. The position of the CS- was not counterbalanced given that the stimulus dimension was symmetrical and intensity bias was minimized (see Lee et al., 2018; Wong & Lovibond, 2017). All stimuli were individually presented in the centre of a white screen.

A computer equipped with Presentation software (Neurobehavioral Systems Inc., Berkeley, CA, Version 20.1) presented all visual stimuli and recorded the US expectancy ratings. Another computer with BrainVision Recorder (Brain Products GmbH, Gliching, Germany) recorded the skin conductance data via two Ag/AgCl electrodes at a sampling rate of 1000 Hz. A DS7A Digitimer stimulator generated an electric shock US. The electric shock consisted of 125 pulses separated by 5 ms (i.e., US duration of 625 ms), delivered through a bar electrode attached to participants' wrist.

1.3. Procedure

After providing written informed consent, participants were asked to fill in the German version of DASS-21 (Lovibond & Lovibond, 1995; Nilges & Essau, 2015). The DASS-21 is a short version of the original DASS (Depression Anxiety Stress Scale), which validly measures and discriminates three different constructs: depression, anxiety and stress (Antony, Bieing, Cox, Enns & Swinson, 1998; Brown, Chorpita, Korotitsch, & Barlow, 1997; Henry & Crawford, 2005; Lovibond, 1998). Next, the skin conductance and electric shock electrodes were attached to the participants' non-dominant hand. Skin conductance electrodes filled with isotonic gel were attached to the hypothenar muscles on the palm. Participants were then led through an electric shock US intensity calibration, in which the shock US intensity was gradually increased starting from 0.2 mA until participants reached a level of shock US intensity that was perceived as 'definitely unpleasant but not painful'. The ensuing experiment consisted of four phases: fear acquisition, fear generalization test, GS extinction and post-extinction generalization test (see Table 1).

Fear acquisition (shock electrodes connected). Participants were informed that some figures would be presented on the screen, which might or might not be followed by an electric shock. They were asked to learn the relationship between the figures and the administration of electric shock (cf. Mertens, Boddez, Krypotos, & Engelhard, under review). Participants were instructed to indicate their shock expectancy ratings using a visual analogue scale (VAS) that appeared at the bottom of the screen during CS presentations. The expectancy VAS ranged from 0 to 100 in which 0 indicates certain no shock and 100 indicates certain shock. The fear acquisition phase was divided into two blocks: in each block, six trials of CS+ and six trials of CS- were presented, leading to twelve trials per block. The CS+ was reinforced at 75% and the CS- was never reinforced. The CS+ was not fully reinforced to allow room for stronger fear responding to stimuli beyond the CS+ in the opposite direction of the CS-, for instance a potential rule-based linear gradient. The CSs were presented alongside the shock expectancy VAS for 8s. The electric shock was delivered immediately after the offset of reinforced CS+ trials for 625 ms. The presentation order was pseudo-randomized, so that the same trial type was not presented more than twice in a row, and the first and last CS+ presentations were always reinforced. The inter-trial interval (ITI) varied between 15 and 18s and was applied to all the following phases.

Fear generalization test (shock electrodes disconnected). The experiment was paused and the experimenter disconnected the shock electrodes. Participants were informed that the electrodes were disconnected to examine if the disconnection would induce a change in physiological responding, therefore setting up the cover story for disconnecting the shock electrodes. Given that no shock could be possibly administered anymore, participants were asked to continue providing their shock expectancy ratings, assuming hypothetically that it was still possible for them to receive a shock. This procedure was used to prevent the occurrence of extinction learning during the test phase, which reduces the probability of participants modifying their response strategy due to extinction learning (e.g., Wong & Lovibond, 2017). Therefore, this phase allowed participants to reveal rule-based generalization without any interference from extinction learning. All nine stimuli along the dimension (stimuli A to I) were presented once each in a randomized order. Thus, the CSs and seven GSs of varying levels of similarity to the CS+ were presented. Each stimulus was presented for 8s and participants were prompted to rate their shock expectancies via the expectancy VAS.

GS extinction (shock electrodes reconnected). The experiment was paused again and the experimenter reconnected the shock electrodes. Participants were told that it was again possible to receive an electric shock. Stimulus I was presented for 6 trials for 8s alongside the VAS without reinforcement.

Post-extinction generalization test (shock electrodes connected). This phase continued seamlessly from the previous phase. Stimuli A, C, E, G and I were presented once each in a randomized order alongside the shock expectancy VAS for 8s. We only presented 5 selected stimuli instead of the whole dimension to minimize the effect of ongoing extinction in this phase. No electric shock was delivered in this phase.

After the completion of the conditioning task, participants were asked to fill in a 2-page questionnaire (see Supplementary Materials). On the first page, the expectancy ratings to stimuli A and I that the individual participant had made during the fear generalization phase (when the shock electrodes were disconnected) were presented. Participants were prompted to explain why they had made these ratings, and to write

---

[1] Seventeen participants were recruited after the COVID-19 restrictions had been lifted in Germany (mid-June 2020). For safety measures, participants cleaned their hands with soap followed by alcohol wipes, in addition to wearing a facial mask throughout the experiment. These safety measures did not affect the recording of SCR (see Supplementary Materials).

Table 1

A to I indicate the different stimuli along the stimulus dimension; + indicates shock US presentation; - indicates shock US omission; numbers in brackets indicate the number of trials of that type in each phase. The shock electrodes were disconnected in the fear generalization test.

Fear acquisition	Fear generalization test	GS extinction	Post-extinction generalization test
E+ (9)	[A-I]- (1)	I- (6)	A- (1)
E− (3)			C- (1)
A- (12)			E− (1)
			G- (1)
			I- (1)

down in detail any rules or strategies of responding they used. The second page was handed out to the participant only after the first page was completed, and consisted of five statements. Each statement described a potential relationship between the stimuli and the electric shock in terms of relational rules (linear right, linear left, similarity, no rule and other). Participants had to rate how much they considered each statement to be true on a scale of 0–100, with 0 being false and 100 being true. If participants perceived that none of the given statements matched their responding during the conditioning task, they were asked to write down their own description of the relationship between the stimuli and the shock US in the ‘Other’ section.

1.4. Scoring and analysis

Analysis of skin conductance data was based only on data collected when the shock electrodes were connected (i.e., during fear acquisition, GS extinction, and post-extinction generalization test). Given that a shock was expected to occur during these phases, anticipatory fear indexed by skin conductance could be measured. First, a 1 Hz high-pass filter and a 50 Hz notch filter were applied to the skin conductance data. Next, we calculated the SCR by finding the difference between the maximum response and the corresponding trough in the interval of 1s after stimulus onset to stimulus offset (see Pineles, Orr, & Orr, 2009). This was first done by identifying the maximum and minimum responses in each stimulus interval automatically by BrainVision Analyzer. We then checked if the detected minimum response was located at the corresponding trough of the maximum response (i.e., a stimulus interval with multiple peaks); if not, we manually adjusted it accordingly. The SCRs were then square root transformed to reduce skewness (Boucsein et al., 2012).

Both expectancy ratings and SCR data were analysed by a set of planned contrasts, using a multivariate repeated measures model (O’Brien & Kaiser, 1985). Planned contrasts compare specific levels of factors to test for a priori hypotheses, allowing for more targeted tests than an omnibus ANOVA (Bird, 2004). Planned contrasts were used to assess acquisition, extinction and generalization gradients, and to compare these between the different rule groups for US expectancy ratings and SCRs. For fear acquisition, three orthogonal repeated measures contrasts were used. First, the averaged responding to the CS+ was compared to the CS- (main effect of CS type). Second, the second block was compared to the first block to examine whether there were any changes in responding across acquisition (main effect of Block). Third, the interaction of these two contrasts examined whether the changes in responding across blocks differed between the CSs (CS type × Block interaction). For GS extinction, a linear trend repeated measures contrast was used to assess whether responding to stimulus I decreased across extinction trials. For both fear generalization test and post-extinction generalization test, two orthogonal polynomial trends were used: A linear trend repeated measures contrast was used to capture linear gradients across the stimulus dimension, and a quadratic trend repeated measures contrast was used to assess peaked, bell-shaped gradients. With an additional repeated measures contrast, we compared responding to the CS+ and the CS- in the post-extinction generalization test, to examine whether differential responding to the CSs persisted after GS extinction. We used between-group contrasts to compare the difference in responding between the rule groups. Finally, all interactions between the group and repeated measures contrasts were calculated to assess rule group differences in all phases. For the post-extinction generalization test, the additional contrast comparing responding to the CSs and the two polynomial contrasts were non-orthogonal. Therefore, we applied a Bonferroni correction for these three non-orthogonal repeated measures contrasts ($\alpha/3 = 0.05/3 = 0.0167$) to control for family-wise-error rate. The Bonferroni corrected alpha levels are indicated as *‘p (.05/3)’*. An additional exploratory analysis was carried out to examine whether the generalization gradients for expectancy ratings differed in shape before and after GS extinction, by examining the polynomial trends across stimuli presented in both fear generalization test and post-extinction generalization test (stimuli A, C, E, G & I), and whether this change differed between rule groups. All analyses were carried out using R (R Core Team, 2020) with Afex (Singmann, Bolker, Westfall, Aust & Ben-Shacher, 2020) and Phia (De Rosario-Martinez, 2015) packages.

2. Results

Statistical analyses were restricted to participants who had acquired differential US expectancy ratings. This was defined by a difference of at least 50 points between the averaged expectancy ratings to the last four CS+ trials and the averaged expectancy ratings to the last four CS- trials in the fear acquisition phase, same as in our prior study (Wong & Lovibond, 2017). Two participants were excluded on this basis. Two additional participants were excluded, one due to technical issues that prohibited the recording of skin conductance, and one for requesting to reduce shock US intensity during the experiment, leaving a total of 59 participants.

2.1. Categorization of rule-based generalization

Three raters, who were blind to the US expectancy and SCR data categorized participants into different rule groups based on participants’ responses to the post-experimental questionnaire. The raters first classified participants’ self-reported rules according to the open-ended question on the first page of the questionnaire. If the reported rules were unclear, the raters would then consult the close-ended section of the questionnaire and categorized participants according to the rule the participants most strongly endorsed. All raters reached a substantial level of consensus as indicated by Fleiss’ Kappa ($\kappa = 0.70$, $p < .001$). Discrepancies in rule classification were then resolved via discussion (see more details in the Supplementary Materials). Thirty-four participants were classified as having adopted a similarity rule, whereby they expected that stimuli that were more perceptually dissimilar to the CS+ were less likely to signal shock (Similarity group). Twenty participants were classified as having adopted a linear rule, whereby they expected that the more the dot position was to the right, the more likely it would signal shock (Linear group). Five participants reported not able to infer any clear rules (No-rule group). Considering the small sample size of the No-rule group (and therefore the lack of statistical power), participants from the No-rule group were excluded from statistical analyses. The final sample thereby comprised 54 participants (38 females) with a mean age of 26.1 years ($SD = 5.4$) and a mean electric shock intensity of 1.1 mA ($SD = 0.5$). Table 2 shows the demographic and DASS-21 data. No significant group differences emerged.

2.2. Fear acquisition

Fig. 2A and 2B show the mean US expectancy ratings in each rule group during fear acquisition (data plotted trial by trial can be seen in the Supplementary Materials). Overall, participants showed higher expectancy ratings to the CS+ compared with the CS-, yielding a

Table 2
Demographic and DASS-21 data: Means (and standard deviations).

	Linear group (*n* = 20)	Similarity group (*n* = 34)	*F* or χ^2	*p*
Age	26.6 (5.7)	25.8 (5.3)	0.29	.593
Sex - Females	14 (70%)	24 (71%)	<0.01	.964
US intensity	1.1 mA (0.5)	1.0 mA (0.5)	0.09	.765
DASS 21-Anxiety	2.1 (3.2)	3.7 (5.4)	1.46	.232
DASS 21-Depression	4.1 (5.4)	5.0 (5.3)	0.38	.542
DASS 21-Stress	6.1 (5.9)	8.9 (6.9)	2.37	.130

significant main effect of CS type, $F(1,52) = 1072.0, p < .001, \eta_p^2 = 0.95$. Differential US expectancy ratings to the CSs emerged across acquisition, confirmed by a significant interaction between CS type and Block, $F(1,52) = 182.25$, $p < .001$, $\eta_p^2 = 0.78$. Importantly, no interactions involving groups reached significance (highest $F = 1.40$, $p = .243$, $\eta_p^2 = 0.03$).

Fig. 2C and 2D show the square root SCRs in each rule group during fear acquisition. Participants showed stronger SCRs to the CS+ than to the CS-, supported by a main effect of CS type, $F(1,52) = 49.46$, $p < .001$, $\eta_p^2 = 0.49$. However, unlike in the US expectancy measure, there was a lack of increasing differentiation in conditioned fear to the CSs across acquisition, $F(1,52) = 0.20$, $p = .654$, $\eta_p^2 < 0.01$. Instead, there was a general decrease in responding to the CSs over blocks, $F(1,52) = 5.41$, $p = .024$, $\eta_p^2 = 0.09$, presumably due to habituation to the stimuli. Unexpectedly, the Linear group showed larger differential responding to the CSs than the Similarity group, yielding a significant interaction between CS type and Group, $F(1,52) = 4.57$, $p = .037$, $\eta_p^2 = 0.08$. This group difference was largely driven by stronger fear responding to the CS+ in the Linear group compared with the Similarity group, whereas both groups showed a similar level of responding to the CS-. No other effects reached significance (highest $F = 0.77$, $p = .384$, $\eta_p^2 = 0.02$).

In sum, both groups successfully acquired differential responding to the CSs in both US expectancy ratings and SCR. While there were no predisposing group differences in the acquisition of US expectancy ratings, the Linear group showed enhanced acquisition of skin conductance responding to the CS+.

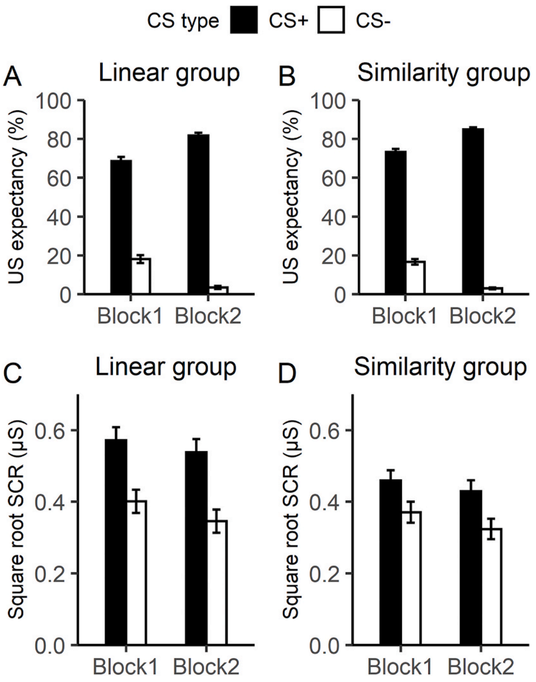

Fig. 2. Mean US expectancy ratings (A & B) and square root skin conductance responding (C & D) across fear acquisition blocks. Error bars indicate the standard error of the mean.

2.3. *Fear generalization test*

Fig. 3A shows the fear generalization gradients of US expectancy in the Similarity and the Linear group. The Similarity group was characterized by a peaked gradient with decreasing expectancy ratings to stimuli more different from the CS+ along the stimulus dimension, whereas the Linear group was characterized by a S-shaped gradient with low ratings to stimuli on the left of CS+ and high ratings to stimuli right of the CS+. The Similarity group showed a more peaked gradient than the Linear group, yielding a significant interaction between Group and quadratic trend across stimuli, $F(1,52) = 22.42$, $p < .001$, $\eta_p^2 = 0.30$. In contrast, the Linear group showed a more linear gradient than the Similarity group, confirmed by a significant interaction between Group and linear trend across stimuli, $F(1,52) = 62.55$, $p < .001$, $\eta_p^2 = 0.55$. Collectively, distinct fear generalization gradients were exhibited according to the different rules adopted by participants.[2]

2.4. *GS extinction*

Fig. 4A shows the US expectancy to the extinction stimulus (stimulus I) during GS extinction. Averaged across groups, expectancy ratings to the extinction stimulus decreased across trials, supported by a main effect of linear trend across trials, $F(1,52) = 133.02, p < .001, \eta_p^2 = 0.72$. Crucially, the Linear group showed higher ratings than the Similarity group to the extinction stimulus in early extinction trials, followed by a more rapid and greater decrease in ratings to it. This pattern was supported by a significant interaction between Group and linear trend across trials, $F(1,52) = 36.54$, $p < .001$, $\eta_p^2 = 0.41$. Follow-up analyses showed that both groups showed extinction learning to the GS (lowest $F = 28.55$, $p < .001$, $\eta_p^2 = 0.46$).

The SCR data showed a similar pattern (Fig. 4B). Averaged across groups, responding to the extinction stimulus decreased across trials, yielding a significant linear trend across trials, $F(1,52) = 20.79$, $p < .001$, $\eta_p^2 = 0.29$. The Linear group descriptively showed stronger SCRs to the extinction stimulus early in extinction compared to the Similarity group. However, there was no statistical evidence to suggest stronger extinction learning to the extinction stimulus in the Linear group, as the interaction between Group and linear trend across trials did not reach significance, $F(1,52) = 0.07$, $p = .788$, $\eta_p^2 < 0.01$.

2.5. *Post-extinction generalization test*

Fig. 3B shows the US expectancy to the selected test stimuli along the stimulus dimension. Averaged across groups, we observed a peaked gradient with the highest ratings to the CS+, yielding a significant main effect of quadratic trend along the stimulus dimension, $F(1,52) = 202.53$, $p\ (.05/3) < .001$, $\eta_p^2 = 0.80$. No effects involving linear trend across stimuli reached significance (highest $F = 3.12$, $p\ [.05/3] = .083$, $\eta_p^2 = 0.06$). Crucially, the Linear group showed a flatter gradient than the Similarity group, confirmed by a significant interaction between Group and quadratic trend across stimuli, $F(1,52) = 28.84$, $p(.05/3) < .001$, $\eta_p^2 = 0.36$. Notably, the flat gradient in the Linear group was characterized by lower ratings to the CS+ compared with the Similarity group. This effect was indicated by smaller differential ratings to the CSs in the Linear group, supported by a significant interaction between Group and differential ratings to the CSs (i.e., ratings for stimulus E compared to stimulus A), $F(1,52) = 34.05$, $p\ (.05/3) < .001$, $\eta_p^2 = 0.40$. However, follow-up analyses revealed that Linear group still showed significantly

---

[2] Without the possibility of receiving an electric shock, no anticipatory fear responses could be assessed by the SCR measurement. Nonetheless, we included the SCR data collected during fear generalization test in the Supplementary Materials. No significant effects were observed.

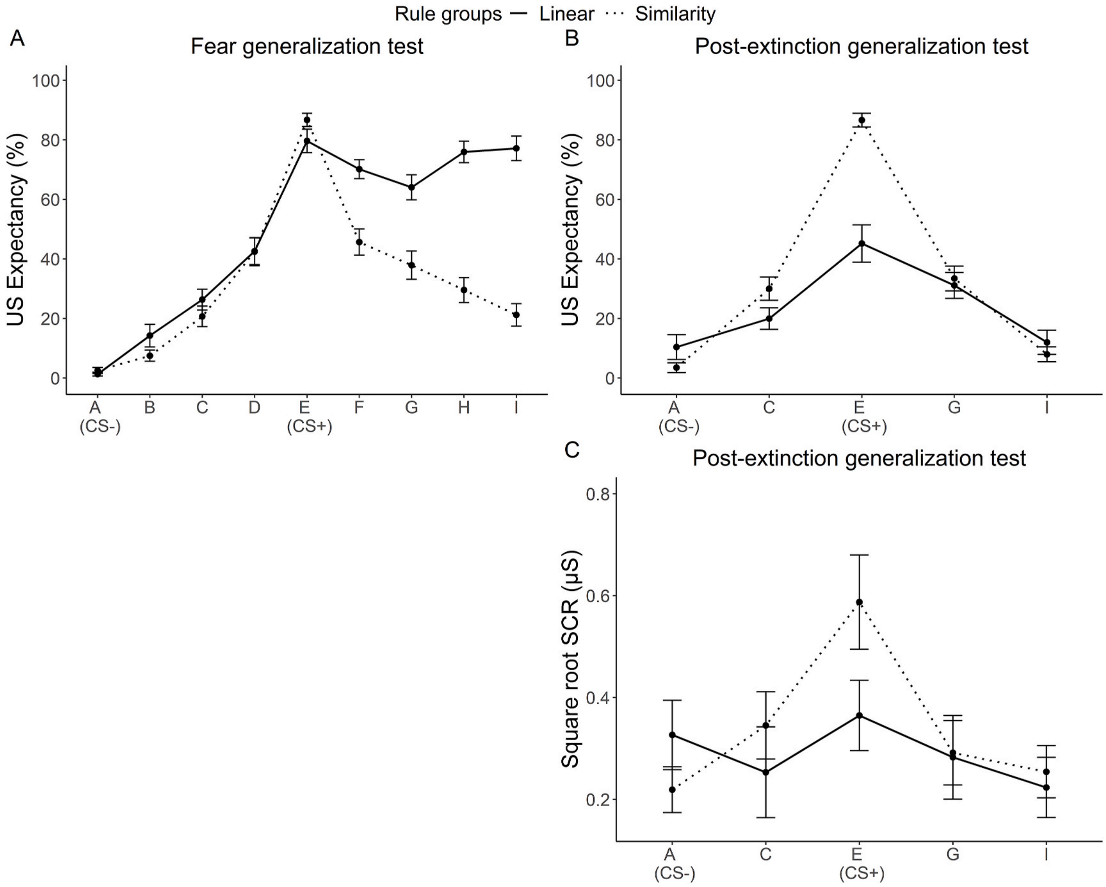

Fig. 3. (A) Mean US expectancy ratings made during fear generalization test. (B) Mean US expectancy ratings and (C) square root skin conductance responding in the post-extinction generalization test. Error bar indicates the standard error of the mean.

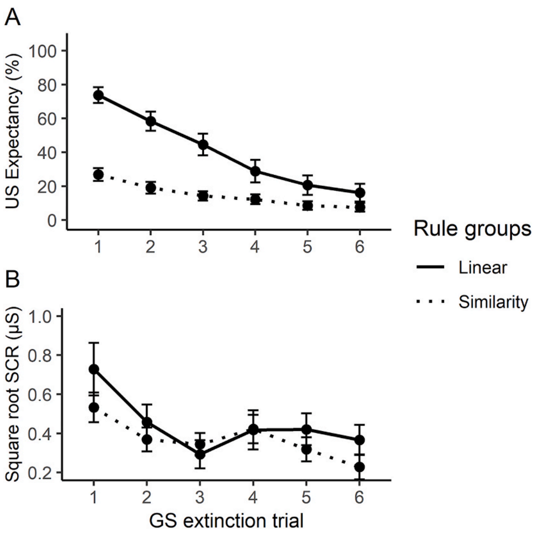

Fig. 4. Mean US expectancy ratings (A) and square root skin conductance responding (B) in the GS extinction phase. Error bars indicate standard error of the mean.

higher expectancy ratings for the CS+ than for the CS-, $F = 13.25$, p (*.05/3*) $< .001$, $\eta_p^2 = 0.41$.

The SCR data showed a similar peaked gradient (Fig. 3C), yielding a significant quadratic trend across stimulus dimension across groups, $F(1,52) = 12.0$, p *(.05/3)* $< .001$, $\eta_p^2 = 0.19$. No effects involving linear trend across the stimulus dimension reached significance (highest $F = 1.03$, p *[.05/3]* $= .314$, $\eta_p^2 = 0.02$). The Linear group showed a flatter gradient than the Similarity group, however, the interaction between Group and quadratic trend across stimuli did not reach significance, $F(1,52) = 5.53$, p *(.05/3)* $= .023$, $\eta_p^2 = 0.10$. Importantly, the Linear group showed less fear responding to the CS+ than the Similarity group, supported by a significant interaction between Group and differential responding to the CSs, $F(1,52) = 6.50$, p *(.05/3)* $= .014$, $\eta_p^2 = 0.11$. Follow-up analyses confirmed that the Similarity group exhibited significant differential responding to the CSs, $F(1,33) = 15.99$, p *(.05/3)* $< .001$, $\eta_p^2 = 0.33$. In contrast, there was no evidence for a persistence of differential responding to the CSs in the Linear group, $F(1,19) = 0.38$, p *(.05/3)* $= .545$, $\eta_p^2 = 0.02$.

In sum, the Linear group showed a flatter generalization gradient of GS extinction than the Similarity group. Although this pattern was significant in the expectancy ratings, it remained descriptive in the SCR measure. Furthermore, the Linear group showed less fear responding to the CS+ than the Similarity group in both expectancy ratings and SCR measure.

2.6. *Comparison between the gradients before and after GS extinction*

The linearity of the generalization gradient decreased after GS extinction as compared with before GS extinction, presumably due to a decrease in expectancy ratings to the CS+ and stimuli right of the CS+. This decrease was more significant in the Linear group than the Similarity group. This pattern was supported by a 3-way interaction between phase, linear trend across stimuli and groups, $F(1,52) = 46.71$, $p < .001$, $\eta_p^2 = 0.41$. Follow-up analyses confirmed that both groups showed a significant decrease in their generalization gradients' linearity (lowest $F = 17.67$, $p < .001$, $\eta_p^2 = 0.35$). This pattern suggests that both groups showed generalization of GS extinction, and provides further evidence that this effect was larger in the Linear group than in the Similarity group. No interactions involving quadratic trend reached significance (highest $F = 2.56$, $p = .116$, $\eta_p^2 = 0.05$).

3. Discussion

Using a differential conditioning procedure, the current study examined whether different learnt threat beliefs led to different magnitude of generalized fear to the same GS, thus presenting this GS in extinction would evoke different levels of expectancy violation. Importantly, we examined whether this difference in expectancy violation would determine the strength of generalization of GS extinction along a non-intensity based stimulus dimension. We identified rule groups that reported using either a linear or a similarity rule during generalization and examined group differences in generalization gradients before and after GS extinction.

The generalization gradients observed in the fear generalization test were highly consistent with participants' reported rules. Specifically, participants reporting a similarity rule showed a peaked generalization gradient, with the highest US expectancy ratings for the CS+ and a gradual decrease in expectancies to stimuli more different from the CS+. In contrast, participants who reported a linear rule showed increasing US expectancies from the CS- and reached an asymptotic level at the CS+, while expectancies for stimuli right of the CS+ slightly decreased and then increased back to the asymptotic level again. These patterns replicated the discrete generalization gradients between rule groups observed in past studies (Ahmed & Lovibond, 2019; Lee et al., 2018; Wong & Lovibond, 2017, 2018). In fact, rule-based generalization shares some similarities with studies of feature generalization (Ahmed & Lovibond, 2015a; Vervliet, Kindt, Vansteenwegen, & Hermans, 2010; Vervliet & Geens, 2014). In these feature generalization studies, participants learnt that one feature of the CS+ (e.g., shape or color) was the predictor of an aversive US through either verbal instructions (Ahmed & Lovibond, 2015a; Vervliet et al., 2010) or direct experience (Ahmed & Lovibond, 2015b; Vervliet & Geens, 2014) before or during fear acquisition. In a subsequent generalization phase, participants showed more generalized fear to a GS that shared the predictive feature with the CS+ (e.g. the same color) than a GS that shared a non-predictive feature with the CS+ (e.g., the same shape). That is, although the two GSs were equally similar to the CS+, participants exhibited selective fear generalization to the stimulus that shared the learnt predictive feature with the CS+, therefore resulting in participants holding different threat beliefs (e.g., novel stimuli that shared the same color with the CS+ predict an aversive outcome). The results from our study are thus in line with these findings and further corroborate the notion that threat beliefs formed through either feature learning of the CSs or rule formation, guide fear generalization, indicating the importance of identifying individual threat beliefs.

Consistent with our prediction, generalization rules differentially affected fear extinction learning. During GS extinction, where the GS at the far right on the stimulus dimension served as the extinction stimulus, the Linear group showed strong generalized fear to it, as evidenced by strong US expectancy ratings to the GS in both fear generalization test and GS extinction. The absence of an US therefore evoked a large amount of expectancy violation, resulting in strong extinction learning as indexed by a rapid decrease in US expectancy (Craske et al., 2014; Rescorla & Wagner, 1972). In contrast, the Similarity group showed weak generalized fear to the extinction stimulus, as shown by the low level of expectancy ratings to it in both fear generalization test and GS extinction. The absence of a shock US was thus largely consistent with participants' threat belief, hence little if any extinction learning took place. However, this pattern was only observed in the US expectancy ratings. A potential explanation for the lack of group differences in the extinction of SCR is presumably due to its high inter-individual variability (Lykken & Venables, 1971).

A major finding was that the Linear group showed a relatively flat generalization gradient of US expectancy ratings in the post-extinction generalization test. Even more critically, this flat gradient was characterized by a low level of fear responding to the CS+ in both US expectancies and SCR measures. In contrast, the Similarity group showed a sharp peaked gradient comparable to the gradient before GS extinction, with strong fear responding to the CS+ compared to the Linear group. Combining these patterns with the findings observed in the GS extinction phase, the results strongly suggest that the magnitude of extinction outcome depends on the violation of one's threat expectancy to the extinction stimulus and thereby determines the degree of subsequent generalization of GS extinction. The current findings are highly consistent with the notion that the degree of expectancy violation strongly influences extinction learning (Craske et al., 2014, 2018), and further illustrate how expectancy violation to a GS generalizes to other fear-related stimuli. Crucially, as the same GS was used in extinction across all participants, the observed group difference in the generalization gradient of GS extinction is due to the GS evoking expectancy violation to different extents. An alternative explanation for the current findings is that different degrees of expectancy violation led to the preservation or extinction of one's initial threat belief. The GS in extinction was largely consistent with the Similarity group's initial threat belief (i.e., the further away the dot is from the center of the box, the less likely shock occurs), therefore this threat belief was largely preserved. This was indicated by the similar peaked gradients before and after GS extinction in the Similarity group. In contrast, the GS in extinction violated the Linear group's initial threat belief (i.e., the further the dot is to the right, the more likely shock occurs), thus leading to the violation of this threat belief. Nonetheless, the current study could not distinguish whether the findings were purely driven by the generalization of GS extinction, the violation of one's initial threat belief, or potentially a combination of both. Future studies could assess to what extent participants continue to embrace their initial threat beliefs after GS extinction.

Although the Linear group exhibited strong generalization of GS extinction, threat expectancy to the CS+ still persisted (i.e., differential expectancy ratings to the CS+ and the CS- remained after GS extinction). This apparent persistence of threat expectancy to the CS+ could be a by-product of the stimulus dimension in use. Lee et al. (2018) suggested that the dot-in-a-box stimulus dimension has a clearly defined mid-point. Combined with the stimulus at the mid-point of the dimension serving as a CS+ might have encouraged the classification of the stimulus dimension into three categories: left of middle (stimulus A to D), middle (stimulus E) and right of middle (stimulus F to I). As a result, extinction learning to the GS might strongly generalize to stimuli in the same category (stimuli right of the middle) but weaker to stimuli of other categories. This may have produced the seemingly persistent threat expectancy to the CS+ in the Linear group. Future studies could use a more continuous stimulus dimension with no clearly defined mid-point (e.g., blue-green dimension). This might minimize the categorization of stimulus dimension, and allows the investigation of whether the persistence of threat expectancy to the CS+ after strong expectancy violation is merely an artefact or not.

From a clinical perspective, the current findings underline that the selection of stimuli presented during exposure-based treatment should be individually tailored to optimize the generalization of extinction learning. Specifically, GSs that elicit strong fear during exposure sessions are preferable to allow for a high level of expectancy violation. This is especially crucial considering that the exact circumstances and stimuli of acquisition are very unlikely to be reproduced, as is often the case in clinical practice. Importantly, the same stimulus may provoke a high threat expectancy in one patient but not the other, depending on the individual threat beliefs. For instance, a dog phobic individual that believes that the bigger the size of a dog, the more aggressive it is may show stronger fear to a calm Rottweiler than to a frequently barking Chihuahua. However, another dog phobic individual that believes the frequency of barking is positively associated with aggressiveness may show more fear to a Chihuahua than to a Rottweiler. Therefore, presenting a Chihuahua during exposure-sessions may trigger stronger threat expectancy violation and therefore effectively reduces fear to dogs for one dog phobic patient but not necessarily the other. To further explore the clinical implications of learnt threat beliefs, future studies could examine the association between threat beliefs and behavioral avoidance (e.g., Pittig, Wong, Gluck, & Boschet, 2020; Preusser, Margraf, & Zlomuzica, 2017).

One limitation of the current study is that we had no control over the sample size of the rule groups because we could not manipulate the rules inferred by participants. The unequal sample sizes in the rule groups (n=20 in the Linear group vs n=34 in the Similarity group) may have reduced statistical power. However, we also perceive the self-generated rules as a strength of this study, considering that it parallels with clinically anxious individuals spontaneously forming their own threat beliefs. A second limitation was that the stimulus dimension in use might have encouraged the classification of stimuli into different categories, rendering the stimulus dimension non-continuous (see Lee et al., 2018). This might have then limited the generalization of GS extinction to other stimuli, for instance, the CS+. A third limitation is that the Linear group showed stronger acquisition of SCR to the CS+ than the Similarity group. However, this group difference is unlikely to have confounded the current findings considering that the Linear group still exhibited less SCR and lower US expectancy ratings to the CS+ than the Similarity group after GS extinction.

In conclusion, the current work replicated the discrete gradients formed by rule-based generalization (Ahmed & Lovibond, 2019; Lee et al., 2018; Wong & Lovibond, 2017), suggesting the formation of different threat beliefs. As a result, the same GS elicited different levels of generalized fear between rule groups. Presenting this GS in extinction accordingly evoked different levels of expectancy violation and thereby different extinction learning outcomes. Participants for whom expectancy violation was maximized during extinction learning (i.e., Linear group) showed greater generalization of GS extinction than participants with little to no expectancy violation during extinction (i.e., Similarity group). The present work emphasises the importance of identifying different threat beliefs in clinically anxious individuals, and suggests tackling the strongest threat belief to maximize the effectiveness of exposure-based treatments.

CRediT authorship contribution statement

Alex H.K. Wong: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Visualization. **Valentina M. Glück:** Formal analysis, Writing - review & editing. **Juliane M. Boschet:** Formal analysis, Writing - review & editing. **Paula Engelke:** Formal analysis, Writing - review & editing.

Acknowledgements

This study was supported in part by the *Overhead* funding *der Fakultät HW* to AHKW. The authors would like to thank Kyra Wagner, Benedikt Bohn and Jacqueline Nar for their help in data collection and data processing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2020.103755.

Conflict of interest

The authors declare no conflict of interest.

References

Ahmed, O., & Lovibond, P. F. (2015a). The impact of instructions on generalization of conditioned fear in humans. *Behavior Therapy, 46*(5), 597–603. https://doi.org/10.1016/j.beth.2014.12.007.

Ahmed, O., & Lovibond, P. F. (2015b). The impact of previously learned feature-relevance on generalisation of conditioned fear in humans. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 59–65. https://doi.org/10.1016/j.jbtep.2014.08.001.

Ahmed, O., & Lovibond, P. F. (2019). Rule-based processes in generalisation and peak shift in human fear conditioning. *Quarterly Journal of Experimental Psychology, 72*(2), 118–131. https://doi.org/10.1177/1747021818766461.

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176–181. https://doi.org/10.1037/1040-3590.10.2.176.

Barry, T. J., Griffith, J. W., Vervliet, B., & Hermans, D. (2016). The role of stimulus specificity and attention in the generalization of extinction. *Journal of Experimental Psychopathology, 7*(1), 143–152. https://doi.org/10.5127/jep.048615.

Bird, K. D. (2004). *Analysis of variance via confidence intervals*. London: Sage Publications.

Blough, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 1*(1), 3–21. https://doi.org/10.1037/0097-7403.1.1.3.

Boucsein, W., Fowles, D. C., Grimnes, S., et al. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology, 49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x.

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the depression anxiety stress scales (DASS) in clinical samples. *Behaviour Research and Therapy, 35*, 79–89. https://doi.org/10.1016/S0005-7967(96)00068-X.

Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A., & Hofmann, S. G. (2018). Cognitive behavioural therapy for anxiety and related disorders: A meta- analysis of randomized placebo-controlled trials. *Depression and Anxiety, 35*(6), 502–514. https://doi.org/10.1002/da.22728.

Craske, M. G., Hermans, D., & Vervliet, B. (2018). *State-of-the-art and future directions for extinction as a translational model for fear and anxiety. 373*. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, Article 20170025. https://doi.org/10.1098/rstb.2017.0025, 1742, pii.

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

De Rosario-Martinez, H. (2015). Phia: Post-hoc interaction analysis. R package version 0.2- 1 https://CRAN.R-project.org/package=phia.

Dunsmoor, J. E., & Murphy, G. L. (2014). Stimulus typicality determines how broadly fear is generalized. *Psychological Science, 1–6*. https://doi.org/10.1177/0956797614535401.

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227–239. https://doi.org/10.1348/014466505X29657.

Hofmann, S. G., & Smits, J. A. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *Journal of Clinical Psychiatry, 69*(4), 621–632. https://doi.org/10.4088/jcp.v69n0415.

Huff, R. C., Sherman, J. E., & Cohn, M. (1975). Some effects of response-independent reinforcement in auditory generalization gradients. *Journal of the Experimental Analysis of Behavior, 23*, 81–86. https://doi.org/10.1901/jeab.1975.23-81.

Hull, C. L. (1949). Stimulus intensity dynamism (V) and stimulus generalization. *Psychological Review, 56*, 67–76. https://doi.org/10.1037/h0058051.

Lee, J. C., Hayes, B. K., & Lovibond, P. F. (2018). Peak shift and rules in human generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000558. Advance online publication.

Lovibond, P. F. (1998). Long-term stability of depression, anxiety, and stress syndromes. *Journal of Abnormal Psychology, 107*, 520–526. https://doi.org/10.1037/0021-843X.107.3.520.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney Psychology Foundation.

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

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

Mertens, G., Boddez, Y., Krypotos, A.-M., & Engelhard, I. M.. Human fear conditioning depends on stimulus contingency instructions. Neurobehavioral Systems, Inc., Berkeley, CA https://psyarxiv.com/by964/.

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

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin, 97*, 316–333.

Pierrel, R., & Sherman, J. G. (1960). Generalization of auditory intensity following discrimination training. *Journal of the Experimental Analysis of Behavior, 3*, 313–322. https://doi.org/10.1901/jeab.1960.3-313.

Pineles, S. L., Orr, M., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology, 46*(5), 984–995. https://doi.org/10.1111/j.1469-8986.2009.00852.x.

Pittig, A., Treanor, M., LeBeau, R. T., & Craske, M. G. (2018). The role of associative fear and avoidance learning in anxiety disorders: Gaps and directions for future research. *Neuroscience & Biobehavioral Reviews, 88*, 117–140. https://doi.org/10.1016/j.neubiorev.2018.03.015.

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

Preusser, F., Margraf, J., & Zlomuzica, A. (2017). Generalization of extinguished fear to untreated fear stimuli after exposure. *Neuropsychopharmacology, 42*, 2545–2552. https://doi.org/10.1038/npp.2017.119.

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Razran, G. H. S. (1949). Stimulus generalisation of conditioned responses. *Psychological Bulletin, 46*, 337–365. https://doi.org/10.1037/h0060507.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Prokasy (Ed.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Croft.

Scheveneels, S., Boddez, Y., Patrick, M., & Hermans, D. (2017). One for all: The effect of extinction stimulus typicality on return of fear. *Journal of Behavior Therapy and Experimental Psychiatry, 57*, 37–44. https://doi.org/10.1016/j.jbtep.2017.03.002.

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2016). The validity of laboratory- based treatment research: Bridging the gap between fear extinction and exposure treatment. *Behaviour Research and Therapy, 86*, 87–94. https://doi.org/10.1016/j.brat.2016.08.015.

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *Afex: Analysis of factorial experiments. R package cersion 0*, 27-2 https://CRAN.R-project.org/package=afex.

Struyf, D., Hermans, D., & Vervliet, B. (2018). Maximizing the generalization of fear extinction: Exposures to a peak generalization stimulus. *Behaviour Research and Therapy, 111*, 1–8. https://doi.org/10.1016/j.brat.2018.09.005.

Vervliet, B., & Geens, M. (2014). Fear generalization in humans: Impact of feature learning on conditioning and extinction. *Neurobiology of Learning and Memory, 113*, 143–148. https://doi.org/10.1016/j.nlm.2013.10.002.

Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010). Fear generalization in humans: Impact of verbal instructions. *Behaviour Research and Therapy, 48*(1), 38–43. https://doi.org/10.1016/j.brat.2009.09.005.

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*(3), 357–371. https://doi.org/10.1016/j.brat.2004.02.005.

Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory, 11*, 555–558. http://www.learnmem.org/cgi/doi/10.1101/lm.77404.

Vervoort, E., Vervliet, B., Bennett, M., & Baeyens, F. (2014). Generalization of human fear acquisition and extinction within a novel arbitrary stimulus category. *PLoS One, 9*(5), Article e96569. https://doi.org/10.1371/journal.pone.0096569.

Wong, A. H. K., & Lovibond, P. F. (2017). Rule-based generalization in single-cue and differential fear conditioning in humans. *Biological Psychology, 129*, 111–120. https://doi.org/10.1016/j.biopsycho.2017.08.056.

Wong, A. H. K., & Lovibond, P. F. (2018). Excessive generalisation of conditioned fear in trait anxious individuals under ambiguity. *Behaviour Research and Therapy, 107*, 53–63. https://doi.org/10.1016/j.brat.2018.05.012.

Wong, A. H. K., & Lovibond, P. F. (2020). Generalization of extinction of a generalization stimulus in fear learning. *Behaviour Research and Therapy, 125*, Article 103535. https://doi.org/10.1016/j.brat.2019.10353.

Zbozinek, T. D., & Craske, M. G. (2018). Pavlovian extinction of fear with the original conditional stimulus, a generalization stimulus, or multiple generalization stimuli. *Behaviour Research and Therapy, 107*, 64–75. https://doi.org/10.1016/j.brat.2018.05.009.

Zielinski, K., & Jakubowska, E. (1977). Auditory intensity generalization after CER differentiation training. *Acta Neurobiologiae Experimentalis, 37*, 191–205.