Behaviour Research and Therapy 184 (2025) 104649

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Fear generalization in individuals with subclinical symptoms of panic disorder

Angelos-Miltiadis Krypotos [a,b,*], Gaëtan Mertens [c], Despoina Matziarli [a], Irene Klugkist [d], Iris M. Engelhard [a]

[a] Department of Clinical Psychology, Utrecht University, Netherlands
[b] Research group of Healthy Psychology, KU Leuven, Belgium
[c] Department of Medical and Clinical Psychology, Tilburg University, Netherlands
[d] Department of Methodology and Statistics, Utrecht University, Netherlands

ARTICLE INFO

Keywords:
Panic disorder
Fear generalization
Therapy
Fear conditioning
Experimental psychopathology

ABSTRACT

Panic disorder (PD) is a debilitating mental health condition, characterized by a preoccupation with the occurrence of panic attacks. Previous research has found that PD patients display increased fear generalization, which entails inflated fear responses to ambiguous stimuli (e.g., intermediate size circles) following fear conditioning wherein a neutral stimulus (e.g., large circle) gets paired with an aversive stimulus (e.g., electric shocks), whereas another neutral stimulus (e.g., small circle) is not paired with this aversive stimulus. The overgeneralization of fear to ambiguous stimuli may be a causal mechanism in the development of panic symptoms. However, this finding requires replication, particularly among subclinical groups to establish temporal priority of fear overgeneralization prior to the development of PD symptoms. This study examines whether fear generalization levels differ between individuals with high and low levels of some PD symptoms. Participants (N = 110) underwent fear conditioning and generalization, measuring physiological and self-report fear responses. Successful fear acquisition and generalization were observed. However, fear generalization did not significantly differ between groups with high and low PD symptomatology. These findings suggest that generalization observed in clinical populations *might* result from psychopathology rather than causing it. Using both clinical and subclinical samples in experimental psychopathology research is therefore important.

Panic disorder (PD) is a debilitating, prevalent, and socially impactful mental health disorder. It is characterized, among other things, by excessive panic attacks and by preoccupation with the occurrence of future attacks. Its lifetime prevalence is estimated between 1.7% and 2.7% (de Jonge et al., 2016; Goodwin et al., 2005; Kessler et al., 2012; Olaya et al., 2018). In addition, PD has been related to tremendous societal costs, with expenses channeled towards the relevant therapies and other indirect expenses (e.g., costs relating to loss in productivity; (Gustavsson et al., 2011). Given the severity of PD, experimental psychologists and therapists have long investigated the causal factors contributing to its symptomatology pathogenesis.

The typical way for investigating the acquisition and maintenance of panic symptomatology in humans is *differential fear conditioning* procedures (Cooper et al., 2022). Such procedures involve the pairing of an initially neutral stimulus (the conditioned stimulus or CS+; e.g., the picture of a large circle) with an aversive stimulus (unconditioned stimulus or US; e.g., electric stimulation on the hand), whereas another neutral stimulus (CS-; e.g., the picture of a small circle) is never paired with a US. Such pairings typically lead to the acquisition of enhanced fear responses (conditioned responses or CRs; e.g., increased startle reflex or fear ratings) towards the CS + compared to the CS-. Importantly, CRs are not limited to the CSs but can spread also to stimuli resembling it (Generalization Stimuli or GSs; e.g., a circle that is a bit smaller than the CS+). This shows that CRs can be elicited not only by stimuli that are directly associated with the original aversive event, but also by stimuli that share formal or perceptual similarities with the CS (Cooper et al., 2022; Dymond et al., 2015; Fraunfelter et al., 2022; Honig & Urcuioli, 1981). Importantly, fear generalization may be a mechanism for why fear responses can generalize to neutral stimuli or situations even in the absence of direct conditioning. For instance, if someone witnesses a car accident and develops a fear of driving, they may also experience fear around other vehicles or transportation-related stimuli, suggesting that

* Corresponding author. Department of Clinical Psychology, Utrecht University, Netherlands.
E-mail address: a.m.krypotos@uu.nl (A.-M. Krypotos).

https://doi.org/10.1016/j.brat.2024.104649
Received 11 December 2023; Received in revised form 2 October 2024; Accepted 5 November 2024
Available online 26 November 2024

the fear response system generalized the fear response from the initial fear-inducing stimulus (the car accident) to other similar stimuli (other vehicles or transportation-related stimuli).

Fear generalization is an adaptive feature, as it undoes the need for learning about potentially dangerous situations. However, researchers consider the *over*generalization of conditioned fear responses as a potential risk factor contributing to the development and maintenance of PD (Dymond et al., 2015; Fraunfelter et al., 2022; Lissek et al., 2010; Mineka & Zinbarg, 2006). Etiological models of PD have suggested that fear conditioning is a mechanism by which previously neutral stimuli, such as contexts or bodily sensations (e.g., the gym or heart palpitations), become associated with the traumatic experience of a panic attack (Bouton et al., 2001). However, given time, fear responses may expand not only in the conditioned stimuli but also to stimuli or situations which resemble the fear stimuli (Fraunfelter et al., 2022). As a result, encountering these stimuli may trigger further panic attacks. Furthermore, the triggers may spread to similar cues (e.g., a crowded train or increased heartrate due to exercise), expanding the range of stimuli that can provoke panic attacks and as such exacerbate panic symptomatology. This showcases the importance of fear generalization, as impairments in functioning seen in panic disorders could be prevented if fear responses were limited to the initial fear stimuli and did not expand to similar stimuli or situations. (Fraunfelter et al., 2022). This last characteristic of panic symptomatology suggests that fear generalization may be a key causal mechanism explaining the onset of PD. As such, detecting excessive fear generalization in individuals with PD can aid in exploring potential risk factors implicated in the development of this disorder.

Lissek et al. (2010) provided evidence in this direction, by testing the relation between PD and de novo fear generalization. Specifically, a group of individuals diagnosed with PD and a healthy comparison group underwent a fear conditioning and fear generalization procedure. During fear conditioning, half of the participants learnt to associate a large circle (the CS+) with the presentation of an electric stimulation (the US), and a small circle (the CS-) with the absence of the electric stimulation. The other half experienced the reversed contingencies. During the generalization phase, all participants encountered the CSs again, together with circles that varied in size between the CS+ and the CS- (the GSs). CRs were measured via the measuring of the startle reflex potentiation and having participants rate the level of risk for encountering the US. Results showed that participants with PD showed reduced linear departure of the generalization gradient compared to comparison participants, indicating stronger CRs towards more GSs. The participants with PD were also faster in responding to risk ratings for stimuli that were similar to the conditioned threat stimulus compared to comparison group. Lissek et al.'s (2010) findings had a major impact on the explanation of PD, but also on anxiety disorders in general, and to date, their study is one of the key textbook examples connecting generalization with psychopathology (Beck & Sloan, 2012; Craighead et al., 2013; Emmelkamp & Ehring, 2014). Furthermore, it constituted the key pillar for other studies that provided evidence that generalization patterns similar to those observed in their study are also present in other disorders such as generalized anxiety disorder (Cha et al., 2014; Lissek et al., 2014) and post-traumatic stress disorder (Lissek & van Meurs, 2015). Importantly, these findings are also used as a basis for the development of relevant clinical interventions for panic disorder and anxiety disorders in general. For instance, it has been argued that a successful intervention for preventing the exacerbation of fear beliefs is by teaching patients to better discriminate between cues that predict a genuine panic attack from cues that are unreliable predictors of such events (Dymond et al., 2015).

A systematic review and meta-analysis by (Cooper et al., 2022) reported that although studies on fear generalization in PD are sparse, the results indicate great importance in two different dimensions of similarity (physical resemblance, contextual overlap), and therefore encourage further research. Please note, however, that there are also previous efforts (e.g. Fullana et al., 2016, that have failed to replicate the key findings of Lissek et al. (2010)). One remaining field of inquiry is whether fear generalization is a genuine risk factor for the development of PD, or is merely part of the PD symptomatology and the result of other factors (e.g., negative emotionality, neuroticism, etc.) that give rise to PD. In the former case, tests on fear generalization could then be used for identifying individuals at risk for PD. Furthermore, only if fear generalization is a true causal factor, interventions based on preventing generalization (e.g., discrimination training) would be useful for preventing PD symptoms. However, if fear generalization is just part of the broader symptomatology of PD, then using fear generalization as a predictor and as a target for treatment is less promising.

To shed light on this issue, we decided to test if differences in fear generalization emerge in a sample with some Panic Disorder symptoms. Specifically, individuals with varying levels of PD symptomatology were invited to participate in this study. The procedure of the experiment followed closely the design of Lissek et al. (2010), wherein participants underwent a fear acquisition and fear generalization protocol. We measured physiological (i.e., startle), self-reported responses (i.e., fear ratings, risk ratings, reaction times, and anxiety ratings) and reactions times to assess fear generalization. The main hypothesis of the study was that individuals that report higher scores in PD symptomatology would report higher fear generalization compared to individuals with lower PD scores.

1. Methods

1.1. Participants

A total of 121 participants were recruited. However, the data of 11 participants were unusable due to technical issues during testing, reducing the number of the final sample to 110 (69 males; 41 females; Mean age: 22.31; Standard deviation age: 4.29). This sample allowed us to detect a Cohen's f of 0.14 (small effect size), for a repeated Measures ANOVA with 2 groups, 6 measurements (the CS and the GSs classes, see below), a power of 0.8 and an alpha of 0.05. Of the 110 individuals, seven were of Asian ethnicity, one of African ethnicity, two of mixed ethnicity, 97 of white/European ethnicity, and three others. These participants were separated into higher and lower panic group based on their responses to the panic related questionnaires (see *Questionnaires subsection below*). The inclusion criteria of the experiment were: not having consumed any substances other than nicotine or caffeine within one day before the experiment, not a current (possible) pregnancy, no cardiovascular problems or disease, no vision problems (unless corrected), no hearing problems (unless corrected), no current or past psychiatric diagnosis, no electrical implants (e.g., pacemaker, ICD, neurostimulator), and fluency in either English or Dutch. Convenience sampling was applied for the participant recruitment and the study was advertised on social media platforms (Facebook, Instagram, WhatsApp), via posters, and the Sona Systems website. Participants were compensated either with 12 Euros or 1.5 study credits.

1.2. Preregistration

The study was preregistered via the Open Science Framework (https://osf.io/uzsaj).[1] We deviated from the preregistration as follows. Although our goal was to reach a maximum of 150 participants, as mentioned in the preregistration, this was not possible due to lab time

[1] Initially the study was a direct replication of Lissek et al. (2010) recruiting a clinical and a healthy sample. However, the study faced serious problems with data collection during the pandemic. Therefore, we had to eventually change our design given the shortage of suitable participants. The updated preregistration reflects those changes in the sample. We report deviations from the preregistration in the analysis section.

constrains. As such, and given the reduced sample, we decided to perform a median split for separating the groups to low and high symptomatology groups, instead of comparing the top 33% of the sample with the lowest 33% as stated in the preregistration. However, the main results after using both samples were largely the same. As such, we here report the analyses after separating the groups with a median split.

1.3. Materials and measurements

1.3.1. Stimuli

A 2-msec electrical stimulation delivered on the wrist of the non-dominant hand via Digitimer served as the US. The level of the US was determined via a staircase procedure (see Procedure). Ten circles of gradually increasing size depicted on a computer screen served as the conditioned (CS) and generalization stimuli (GSs; see Fig. 1). The smallest and largest in size circles represented either the CS + or the CS-. The assignment of the small or large circle as the CS + or CS- was randomly determined among participants.

1.3.2. Startle responses

Electromyographic measurement of the startle blink response was captured using a pair of 6-mm tin-cup electrodes positioned beneath the left eye (one in the outer corner of the eye, one under the eye) and a pair positioned in the middle of the forehead (about 2 cm below the hairline with 1 cm side by side). The sampling rate was set at 1000 Hz, with a bandwidth ranging from 30 to 500 Hz. The startle response was elicited by delivering a brief burst of white noise (40 ms in duration, with a sound intensity level of 102 dB) through closed headphones. This is in line with the published recommendations for measuring the startle response (Blumenthal et al., 2005).

1.3.3. Risk and anxiety ratings of stimuli

Risk ratings were accessed via the question "Level of risk?" and participants could answer in a 3-point scale (1 = no risk, 2 = moderate risk, and 3 = high risk). Participants also reported levels of anxiety evoked by each stimulus using 10-point Likert scales (1 = none, 5 = some, 10 = a lot).

1.3.4. Questionnaires

To measure panic disoder symptoms, the twin scales developed by Chambless and colleagues (1984), were used: the Agoraphobic Cognitions Questionnaire (ACQ) and the Body Sensations Questionnaire (BSQ). Chambless and colleagues (1984) have suggested that the two questionnaires provide independent information. ACQ is a self-reported 14-item scale assessing what the individual anticipates they would do when being anxious (e.g., *I am going to throw up*) using a 5-point Likert scale (1 = *thought never occurs*, 5 = *thought always occurs*). The items are conceptualized to assess six behavioral/social outcomes (e.g., *I am going to act foolish*) and eight physiological outcomes (e.g., *I will have a heart attack*). BSQ is a self-reported 17-item scale that evaluates fear of various physiological anxiety and panic reactions (e.g., *heart palpitations*) ranked from 1 (*not at all*) to 5 (*extremely*). Both the ACQ and the BSQ have been reported to have good internal consistency (Cronbach's $\alpha = 0.82$ and Cronbach's $\alpha = 0.89$ respectively;Arrindell, 1993) and a high test-retest reliability ($r = 0.79$ and $r = 0.79$ respectively;Arrindell, 1993). For this study the observed internal consistencies were Cronbach's $\alpha = 0.85$ for ACQ and Cronbach's $\alpha = 0.90$ for BSQ. To create a single variable of panic symptomatology, the total scores of both questionnaires were calculated by averaging the responses to the individual items composing that score and then averaging those scores. Higher scores indicate higher levels of subclinical PD symptomatology. Participants were separated into the higher and lower PD group based on a median split on the average scores.

For the measurement of depression symptomatology, the Beck Depression Inventory (BDI; Beck et al., 1961) was used. BDI is a self-reported questionnaire measuring depressive symptoms and consisting of 21 items. All items are measured using a 4-point Likert scale ranging in intensity (e.g., 0 = *I do not feel sad*, 1 = *I feel sad*, 2 = *I am sad all the time*, 3 = *I am so sad and unhappy that I can't stand it*). BDI has been found to have high internal consistency (Cronbach's $\alpha = 0.81$;Beck et al., 1988) and high construct validity (Beck et al., 1988). The sum of scores was calculated, with higher scores being indicative of a higher depressive symptomatology. For the present study, Cronbach's α was 0.81.

Trait anxiety was measured using the Y version of the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1977). The STAI-Y is composed of 20 items for the assessment of state anxiety and 20 for the assessment of trait anxiety. State anxiety items include: "*I am worried*", "*I feel calm*". Trait anxiety items include: "*I worry too much over something that really doesn't matter*", "*I am content*". All items are measured using a 4-point Likert scale (1 = *Almost never*, 5 = *Almost always*) with higher scores indicating a higher level of trait anxiety. STAI-Y has been found to have a great internal consistency ranging from $\alpha = 0.86$ to $\alpha = 0.95$ and a high test-retest reliability (scores ranging from $r = 0.69$ to $r = 0.89$; Spielberger et al., 1977). For the present study, the internal consistencies were Cronbach's $\alpha = 0.74$ and Cronbach's $\alpha = 0.75$ for STAI-S and STAI-T, respectively.

1.3.5. Procedure

Our study design was heavily based on Lissek et al. (2010). Participants were first provided with the information brochure, and the opportunity to inquire any questions or concerns. Subsequently, they were asked to sign the informed consent form.

Following that, they were connected to the Digitimer, which provided the electrical stimulation on their forearm. The level of the electrocutaneous stimulation was adjusted using a staircase procedure. Specifically, the level was increased gradually until the participant experienced the stimulation as highly uncomfortable but not painful.

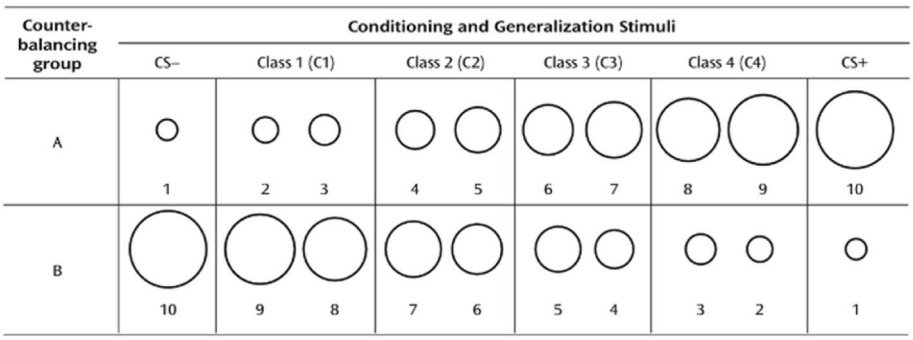

Fig. 1. Overview of the used Conditioned and Generalization Stimuli, separated per counterbalancing group.
Note. Overview of the stimuli. The figure is based on Lissek et al. (2010). In group A, the largest circle was the CS+ and the smallest circle was the CS-, while in group B this was reversed.

The reference sentence used as the highest anchor on a 1–10 scale was: "What is the most uncomfortable you would tolerate at a dentist's appointment?". The procedure was stopped when participants reached a level rated 6 on this scale or when they indicated that they felt uncomfortable going higher (when still at a level <6). Then, the four electrodes were placed on the participants' face for the measurement of the startle blink. Participants received further explanation about the experimental procedure and instructions on how to register their risk ratings. Furthermore, they were instructed that there would be a CS-US contingency in the task which they would have to discover.

The experimental task itself consisted of three phases: pre-acquisition, acquisition, and generalization. In the pre-acquisition phase, participants were presented with two circles (a big and a small one), with one of them being followed by an aversive sound (CS+) and the other not (CS-). Each CS was presented 6 times. The participants were also asked to indicate risk of the US for 50% of the trials. Participants did not receive any electrical stimulation throughout the pre-acquisition phase.

In the acquisition phase, participants were presented with 12 trials of the CS-, and 12 trials of the CS+, with nine CS + trials being followed by the US (75% reinforcement). As in the pre-acquisition phase, participants were asked to provide risk indications for 50% of the trials of the acquisition phase. Also, 50% of the trials included the startle probe presentation. Finally, in the generalization phase, participants were presented with the previous CSs, and additionally, with the intermediate sized circles (GSs). This phase included 12 trials of the CS- representation, along with 12 trials of the CS + representation (6 trials followed by the US; 50% reinforcement). Participants were also presented with 12 trials of each of the GSs, which were never reinforced by the US. During the generalization phase, risk indications were asked for 50% of the trials, with the other 50% of the trials including the startle probe presentation. At the end of the generalization phase, participants filled in the anxiety ratings.

After the experimental task, the electrodes and headphones were removed, and participants were asked to fill in the STAI-Y, BDI, BSQ, and ACQ. Finally, participants were debriefed, thanked, and compensated.

1.3.6. Data reduction

The filtration (28–500 Hz), rectification and smoothing (15.9 Hz low-pass filter) of startle electromyography (EMG) was carried out using BrainVisionAnalyzer 2.1. The time window for identifying the onset latency of the blink reflex ranged from 20 to 150 ms, and the highest voltage of the reflex response was assessed within 150 ms from the onset. The average baseline EMG level (between −30 and 20 ms after stimulus onset) was subtracted from peak levels to determine the startle blink amplitude.

For all dependent variables, apart from the anxiety ratings, the responses to the generalization stimuli were averaged per two GSs, as was done in previous studies (e.g., Lissek et al., 2010).

1.3.7. Data analyses

For the acquisition phase, a 2 × 2 (group [high vs low PD symptomatology] by stimulus [CS- vs. CS+]) analysis of variance (ANOVA) with repeated measures was run, with group as a between subject factor and stimulus as within subject factor. Additionally, generalization effects were analyzed using a 2 × 6 (group [high vs low PD symptomatology] by stimulus type [CS- vs. C1 vs. C2 vs. C3 vs. C4 vs. CS+]) ANOVA with repeated measures, with again group as between subject factor and stimulus as within subject factor. Quadratic trend analyses were conducted for testing the shape of generalization gradients (Dymond et al., 2015). The quadratic effect tests whether fear responses increase when GS physical resemblance moves further away from the CS- and towards the CS+. We, however, acknowledge that there are different ways to test for fear generalization – see Discussion. Alpha was set at 0.05 and was corrected in case of multiple testing using Holm's correction. Lastly, in case of violation of sphericity, we used the Greenhouse-Geisser correction. In order to be able to gauge evidence for the null hypothesis, relative to the alternative hypothesis, and in line with our previous work (e.g., Krypotos et al., 2020), we also employed Bayesian analyses using JASP (Love et al., 2019), using the default settings in JASP. We report Bayes Factors as BF_{10}. The higher the Bayes factors, the higher the evidence that the data came from the alternative, compared to the null, hypothesis, with a cutoff of 10 being considered as strong evidence. For tutorials on Bayesian statistics, we point to relevant resources (Krypotos, Blanken, et al., 2017; Krypotos, Klugkist, & Engelhard, 2017). Please note that our conclusions are mainly based on the results of the *p*-values and not the Bayes factors.

Following a reviewer's helpful suggestion, we decided to analyse the generalization data using a multilevel approach. This allows us to analyse data using a dimensional approach, rather than dichotomizing data. The limitations of dichotomizing data been related to the loss of resolution and power (Irwin & McClelland, 2003; Lonsdorf & Merz, 2017). Following a dimensional approach follows the introduced RDOC recommendation and seem to also be in line with recent suggestions viewing personal characteristics as a continuum rather than viewing them in a dichotomous manner.

For the multilevel analyses we followed the modeling approach of Vanbrabant et al., 2015. and we have fitted 6 different hierarchical models varying in complexity, with the final model including linear and quadratic parameters for the generalization stimuli parameter as well as the mean of the panic questionnaires. The models were.

Model 1. The simplest model, with a random intercept for participants and the fixed effect of stimulus.

Model 2. Adds a random slope for stimulus (allowing the effect of stimulus to vary across participants).

Model 4. Includes both stimulus and a quadratic term for the fixed effects, with random intercept and slope for stimulus.

Model 3. Adds the interaction between stimulus and panic questionnaire scores, plus random intercept and slope for stimulus.

Model 5. Includes stimulus, the quadratic term for stimulus, and random intercept, slope for stimulus, and slope for the quadratic term.

Model 6. The most complex model, including stimulus, its quadratic term, panic questionnaires, interactions between quest and both stimulus and the quadratic term for sitmulus, and random intercept and slope for stimulus.

Model fit was assessed using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Below, we report the results of the winning models.

We followed the same approach using a Bayesian approach with the brms package in R. Default priors were used for these models and model selection was done using the Leave-One-Out criterion. The full results of the Bayesian models can be found as a supplementary appendix.

2. Results

Fear potentiated startle. During the acquisition phase (Fig. 2A), results showed successful fear acquisition with higher startle responses for the CS + compared to the CS-, $F(1, 100) = 88.437$, $p < 0.001$, $\eta^2p = 0.275$, $BF_{10} > 1000$, an effect that did not differ per group, $F(1, 100) = 0.602$, $p = 0.440$, $\eta^2p = 0.006$, $BF_{10} = 0.308$. In the generalization phase, we observed a main effect of stimulus, $F(2.316, 226.949) = 62.594$, $p < 0.001$, $\eta^2p = 0.226$, $BF_{10} > 1000$, which did not differ per group, $F(2.316, 226.949) = 0.732$, $p = 0.732$, $\eta^2p = 0.004$, $BF_{10} = 0.010$. Follow-up post hoc tests showed no significant differences between CS+ and the other stimuli (all *ps* > 0.001) and a significant quadratic term ($p < 0.001$), but also no significant differences between CS- and the other stimuli (all *ps* = 1) and a significant quadratic term ($p < 0.001$). The multilevel analyses showed that the winning model, both for the frequentists and Bayesian results, includes quadratic terms for the stimulus

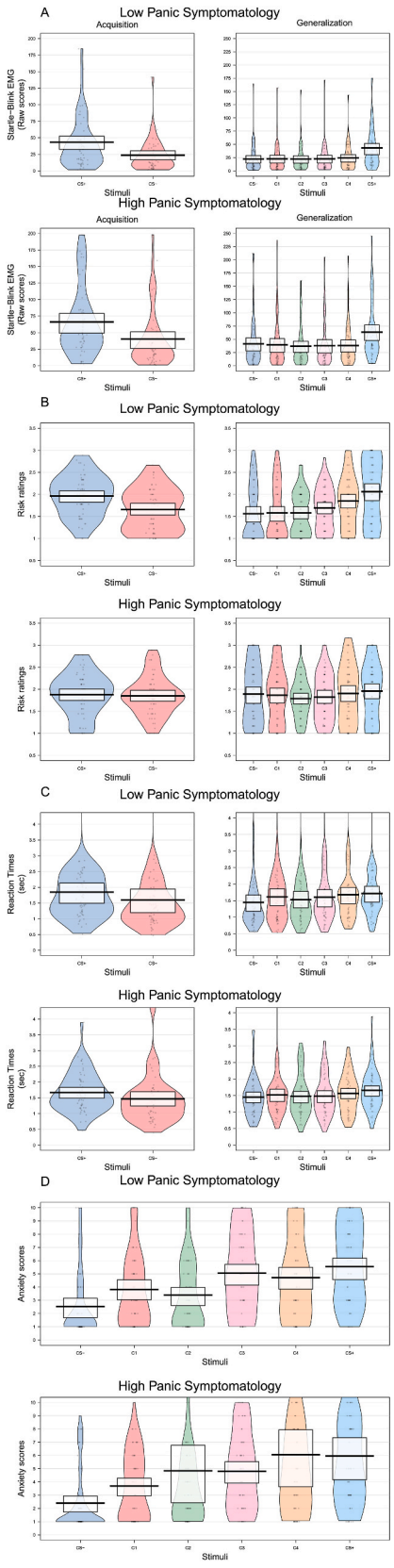

Fig. 2. Per group, Startle-blink EMG results for the acquisition and the generalization phase, Risk ratings for the acquisition and the generalization phase, Reaction time results for the acquisition and the generalization phase, and Anxiety scores for the generalization phase.
Note: CS: Conditioning Stimulus. C: Class.

parameter, which allows for the curvature over the stimulus dimension (Vanbrabant et al., 2015). The model (Model 5) that included the stimuli parameter and its quadratic term, along with random slopes for both termsprovided the best fit to the data (AIC = 5139.8, BIC = 5183.9), outperforming simpler models such as the random intercept-only model (Model 1; AIC = 5287.5, BIC = 5305.1) and models that did not include quadratic effects. Although a more complex model incorporating interactions with the panic questionnaires (Model 6) was also tested, it did not significantly improve fit (AIC = 5188.9, BIC = 5233.0). Model 5 as also the winning model for the Bayesian models. These results suggest successful acquisition and generalization of fear, with no significant group differences, in case of a median split, subclinical levels of panic, or the multilevel approach.

Risk ratings. Participants gave higher risk ratings (Fig. 2B) to the CS + than the CS-, $F(1, 107) = 10.477$, $p = 0.002$, $\eta^2p = 0.089$, $BF_{10} > 1000$, which did seem to differ per group, $F(1, 107) = 4.353$, $p = 0.039$, $\eta^2p = 0.039$, $BF_{10} = 0.180$. Specifically, post-hoc tests with the Holms correction showed that the lower PD symptomatology group reported higher CS differentiation ($t = 3.237$, $p = 0.010$), whereas the higher PD symptomatology group did not report such differences ($t = 0.302$, $p = 0.763$). During the generalization phase, we again found significant differences for stimuli, $F(1.615, 172.817) = 6.063$, $p = 0.005$, $\eta^2p = 0.054$, $BF_{10} = 374.863$, which did not seem to differ between groups $F(1.615, 172.817) = 2.528$, $p = 0.094$ $\eta^2p = 0.023$, $BF_{10} = 0.612$. Post-hoc tests showed significant differences between the CS+ and all GSs (all *ps* ≤ 0.005), except GS4 ($p = 0.502$), with again a significant quadratic term, $t = 2.837$, $p = 0.005$. Also, no significant differences between the CS- and all GSs (all *ps* > 0.239) were observed, although a significant quadratic term was present, $t = 2.837$, $p = 0.005$. We also repeated the generalization analyses using the differential scores for the CSs during the acquisition phase as covariate. Again, the main effect of CS/GS levels was significant, $F(0.758, 288.551) = 3.096$, $p = 0.032$, $\eta^2p = 0.028$, $BF_{10} = 364.164$, and there were no between group differences $F(0.758, 288.551) = 0.193$, $p = 0.884$, $\eta^2p = 0.002$, $BF_{10} = 0.612$. However, we observed between group differences when we compared the top and bottom 33% of the participants, $F(0.547, 0.172) = 3.181$, $p = 0.022$, $\eta^2p = 0.054$, $BF_{10} = 364.164$.

The multilevel results were similar to that of the fear potentiated startle data, showing that the winning model, both for the frequentists and the Bayesian results, is the one that included the stimuli parameter and its quadratic term, along with random slopes for both terms (Model 5; AIC = 717.04, BIC = 761.87). Again, scores in the questionnaires did not result in a better fit. Collectively, results showed successful acquisition and generalization for risk ratings, although acquisition did not seem to be present in the high PD symptomatology group.

Reaction times. Reaction times for providing the risk ratings were increased for the CS+ (Fig. 2C) compared to the CS-, $F(1, 108) = 13.746$, $p < 0.001$, $\eta^2p = 0.113$, $BF_{10} = 70.130$, which did not differ per group, $F(1, 108) = 0.174$, $p = 0.677$, $\eta^2p = 0.002$, $BF_{10} = 0.206$. In the generalization phase, there were significant CS and GSs differences, $F(4.298, 464.150) = 5.257$, $p < 0.001$, $\eta^2p = 0.046$, $BF_{10} = 90.491$, but no group differences, $F(4.298, 464.150) = 3.601$, $p = 0.792$, $\eta^2p = 0.004$, $BF_{10} = 0.382$. Post-hoc results showed significant differences between CS+ and CS- and GS2 (all *ps* < 0.001), but no significant differences between CS+ and GS1, GS3, or GS4 (all *ps* > 0.07). There were no significant differences between CS- and GSs (all *ps* > 0.239), except for CS- and GS4 (all *ps* > 0.011). Lastly, although the linear trend was significant ($t = 4.516$, $p < 0.001$) the quadratic term was not ($t = 0.891$, $p = 0.373$).

Regarding the multilevel models, there was a discrepancy between the frequentists and Bayesian results. Specifically, the model that integrated the stimuli parameter and its quadratic term, along with random slopes for both terms (Model 5), provided the best overall fit (AIC = 13105, BIC = 13150), although the enhancements over simpler models were minimal. The random intercept-only model (Model 1) exhibited a similar fit (AIC = 13101, BIC = 13118), while more complex models, including those with interactions with panic questionnaires (Model 6;

AIC = 13101, BIC = 13146), did not demonstrate a significant improvement in fit. However, for the Bayesian results, the winning model was Model 6, showing an influence of the panic scores.

Taken together, the results showed good CS differentiation using reaction times and some evidence for fear generalization. However, none of these two patterns differed per group in the median split analyses, and results were inconclusive for the multilevel models.

Anxiety ratings. The anxiety ratings (Fig. 2D) collected at the end of the experiment showed that participants evaluated the stimuli differently, $F(3.150, 340.246) = 13.381$, $p < 0.001$, $\eta_p^2 = 0.110$, $BF_{10} > 1000$, which did not differ per group, $F(33.150, 340.246) = 1.343$, $p = 0.245$, $\eta_p^2 = 0.012$, $BF_{10} = 0.049$. The linear term was significant, $t = 8.001$, $p < 0.001$. There were no significant differences between CS+ and GS3, GS4, CS- ($ps \leq 0.005$), between GS1 and GS4 and CS- ($ps \leq 0.005$), between GS2, = GS3, and CS- ($p < 0.001$), and GS4 and CS- ($p = 0.048$). There were no significant differences between CS- and the other GSs (all $ps > 0.165$).

The model that included the stimuli parameter and an additional linear term for its quadratic effect, along with random slopes (Model 5), had an AIC of 6181.5 and was outperformed by Model 6, which incorporated interactions with panic questionnaires and achieved a lower AIC of 6180.4. This indicates a better fit compared to the simpler random intercept-only model (Model 1; AIC = 6245.0, BIC = 6265.0). For the Bayesian results, model 4 was the winning model.

Taken together, similar to previous findings, we found significant differences between CS+ and CS- using anxiety ratings, and good evidence for anxiety generalization. The group differences did not reach significance, and there was inconclusive evidence of whether panic levels contribute to the effect.

3. Discussion

We set out to explore whether individuals with some panic disorder symptoms would also relate to differences in fear generalization, as was previously found for patients with a clinical diagnosis of PD. Across subjective and physiological measures, the pattern showed that participants successfully acquired the CS-US associations, with higher physiological and self-reported responses being higher for the CS + than the CS-. There was also evidence for fear generalization across measures, mostly due to significant quadratic trends across measures. However, and independent of the statistical model (i.e., repeated measures or multilevel modelling) or inference (i.e., frequentists or Bayesian) followed, there were no group differences between high- and low-PD symptomatology participants, apart from the risk ratings during the acquisition phase.

Our findings can be best interpreted in combination with previous findings within clinical populations. To illustrate, and as mentioned in the introduction, Lissek et al. (2010) demonstrated that PD patients show increased fear generalization compared to healthy controls. From this, it could be assumed that we should have also observed increased fear responses to the subclinical sample with the high PD symptomatology compared to the low one. We did not observe this as both groups reported similar fear levels. These results could mean that differences in fear generalization arise because of PD, rather than being a causal factor in the development of PD, and that the relationship between fear generalization and PD symptomatology does not follow a linear continuum, with higher fear generalization being gradually related to higher levels of PD symptomatology.

We need to stress that the differences between our findings and that of Lissek et al., 2010. In contrast to Lissek et al. (2010) we did not include a clinical sample to our study, making direct comparisons unwarranted. Whereas in our study, the sample consisted of healthy university students with varying levels of PD symptomatology, the sample of the Lissek et al. (2010) consisted of patients clearly meeting the criteria for a clinical diagnosis of PD. Nonetheless, the observation that fear generalization does not differ between high- and low subclinical levels of PD symptomatology does demonstrate the subtility of fear generalization as a risk factor or predictive marker for PD. As such, our study shows that in the field of experimental psychopathology, we need to study the full distribution of phenomena (here panic symptomatology) rather than focusing on specific samples (e.g., subclinical versus clinical samples). This suggestion is in line with recent dimensional conceptualizations of psychopathology (Kotov et al., 2022) and by following we may better tap into the mechanisms of psychopathology.

Our study had limitations. First, despite having a large sample, especially for psychophysiological research, we did not collect the targeted 150 individuals given the time constrains. This has resulted in changes in our analytic strategy (i.e., dividing the sample up into a high and low PD symptomatology group, instead of a high, intermediate, and low group), although similar results also arose with our initial statistical plan. Second, there seemed to be group differences in the reaction time variable in the acquisition phase, potentially indicating meaningful group differences in this phase. However, these differences were not present in the generalization phase where they would be expected, which suggests that this finding may be coincidental. A significant limitation of our study is the used questionnaires. Although they have been used before in similar studies, they do not directly tap on the worry or avoidance components of panic symptomatology according to DSM-5 (American Psychiatric Association, 2013). However, please note that avoidance has been highly correlated with threat cognitions (Rosebrock et al., 2022), so although avoidance was not directly measured, we did access key panic disorder symptoms. Lastly, we relied largely on *p*-values and not Bayes factors as often the evidence provided by Bayes factors was inconclusive, indicating a not large enough sample for these analyses. Future studies could replicate the study with large enough evidence to draw concrete conclusions also by using Bayes factors. Lastly, we should acknowledge that despite the use of different analytic models, there are still different types of analyses and data reduction strategies that could have been used to analyse the present data. Specifically, a recent review of the literature found eight ways in which generalization can be quantified (Stegmann et al., 2024). Four of the most used methods for quantifying generalization showed substantial overlap, but also some important differences. However, because our study aims to replicate previous work on the relationship between fear generalization and anxiety symptoms, we followed the operationalization of generalization used in previous comparable studies.

To conclude, despite the largely successful fear acquisition and fear generalization results in the current study, we did not find evidence that fear generalization is different based on subclinical levels of PD symptomatology. These findings seem to suggest that differences in fear generalization reported in prior research with clinical samples may be the consequence, rather than the cause, of anxious psychopathology.

CRediT authorship contribution statement

Angelos-Miltiadis Krypotos: Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Gaëtan Mertens:** Writing – review & editing, Project administration, Methodology, Funding acquisition. **Despoina Matziarli:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Formal analysis, Conceptualization. **Irene Klugkist:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Iris M. Engelhard:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization.

Author note

The research was supported by an NWO (Reg. # 401.18.056) grant awarded to Angelos-Miltiadis Krypotos, Iris Engelhard, Marcel van den

Hout, Irene Klugkist, Arne Leer, Gaëtan Mertens, Dieuwke Sevenster, and Muriel Hagenaars.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Rosalie de Niet for programming the computer task and assisting during the initial parts of the experiment, as well as Frauke Muesch, Gabriel McKeown, Maria Christina Trouptsidi, Ines Turcinhodzic, Malena Oexemann, Mirte Hesselink, Dorien Mons, Tobias Paarlberg, Laurens Schaap, Alisia van den Berg, Zwi Lamers, Suzanne van der Linde, and Jip Fraters for assistance during data collection. Lastly, we are thankful to Prof. Lissek for sharing his material with us as well as his unconditional support during the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2024.104649.

Data availability

I have shared all data and material online https://osf.io/kgeqt/

References

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

Arrindell, W. A. (1993). The fear of fear concept: stability, retest artefact and predictive power. *Behaviour Research and Therapy, 31*(2), 139–148.

Beck, J. G., & Sloan, D. M. (2012). *The Oxford Handbook of traumatic stress disorders*. Oxford University Press.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review, 8*(1), 77–100.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4, 561-571.

Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological Review, 108*(1), 4–32. https://doi.org/10.1037/0033-295x.108.1.4

Cha, J., Greenberg, T., Carlson, J. M., DeDora, D. J., Hajcak, G., & Mujica-Parodi, L. R. (2014). Circuit-wide Structural and functional measures predict Ventromedial Prefrontal Cortex fear generalization: Implications for generalized anxiety disorder. *Journal of Neuroscience, 34*(11), 4043–4053. https://doi.org/10.1523/JNEUROSCI.3372-13.2014

Cooper, S. E., van Dis, E. A. M., Hagenaars, M. A., Krypotos, A.-M., Nemeroff, C. B., Lissek, S., Engelhard, I. M., & Dunsmoor, J. E. (2022). A meta-analysis of conditioned fear generalization in anxiety-related disorders. *Neuropsychopharmacology, 47*(9). https://doi.org/10.1038/s41386-022-01332-2. Article 9.

Craighead, W. E., Miklowitz, D. J., & Craighead, L. W. (2013). *Psychopathology: History, diagnosis, and Empirical Foundations*. John Wiley & Sons.

de Jonge, P., Roest, A. M., Lim, C. C. W., Florescu, S. E., Bromet, E. J., Stein, D. J., Harris, M., Nakov, V., Caldas-de-Almeida, J. M., Levinson, D., Al-Hamzawi, A. O., Haro, J. M., Viana, M. C., Borges, G., O'Neill, S., de Girolamo, G., Demyttenaere, K., Gureje, O., Iwata, N., … Scott, K. M. (2016). Cross-national epidemiology of panic disorder and panic attacks in the world mental health surveys. *Depression and Anxiety, 33*(12), 1155–1177. https://doi.org/10.1002/da.22572

Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear generalization in humans: Systematic review and Implications for anxiety disorder research. *Behavior Therapy, 46*(5), 561–582. https://doi.org/10.1016/j.beth.2014.10.001

Emmelkamp, P., & Ehring, T. (2014). *The Wiley Handbook of anxiety disorders*. John Wiley & Sons.

Fraunfelter, L., Gerdes, A. B. M., & Alpers, G. W. (2022). Fear one, fear them all: A systematic review and meta-analysis of fear generalization in pathological anxiety. *Neuroscience & Biobehavioral Reviews, 139*, Article 104707. https://doi.org/10.1016/j.neubiorev.2022.104707

Fullana, M. A., Torrents-Rodas, D., & Torrubia, R. (2016). Behavioral fear generalization task data in panic disorder. *Hospital Clinic, Barcelona & Universitat Autònoma [Unpublished raw data]* https://osf.io/kgeqt/.

Goodwin, R. D., Faravelli, C., Rosi, S., Cosci, F., Truglia, E., de Graaf, R., & Wittchen, H. U. (2005). The epidemiology of panic disorder and agoraphobia in Europe. *European Neuropsychopharmacology, 15*(4), 435–443. https://doi.org/10.1016/j.euroneuro.2005.04.006

Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., Dodel, R., Ekman, M., Faravelli, C., Fratiglioni, L., Gannon, B., Jones, D. H., Jennum, P., Jordanova, A., Jönsson, L., Karampampa, K., Knapp, M., Kobelt, G., Kurth, T., … Olesen, J. (2011). Cost of disorders of the brain in Europe 2010. *European Neuropsychopharmacology, 21*(10), 718–779. https://doi.org/10.1016/j.euroneuro.2011.08.008

Honig, W. K., & Urcuioli, P. J. (1981). The Legacy of Guttman and Kalish (1956): 25 Years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36*(3), 405–445. https://doi.org/10.1901/jeab.1981.36-405

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research, 40*(3), 366–371. https://doi.org/10.1509/jmkr.40.3.366.19237

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research, 21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Kotov, R., Cicero, D. C., Conway, C. C., DeYoung, C. G., Dombrovski, A., Eaton, N. R., First, M. B., Forbes, M. K., Hyman, S. E., Jonas, K. G., Krueger, R. F., Latzman, R. D., Li, J. J., Nelson, B. D., Regier, D. A., Rodriguez-Seijas, C., Ruggero, C. J., Simms, L. J., Skodol, A. E., … Wright, A. G. C. (2022). The hierarchical Taxonomy of psychopathology (HiTOP) in psychiatric practice and research. *Psychological Medicine, 52*(9), 1666–1678. https://doi.org/10.1017/S0033291722001301

Krypotos, A.-M., Baas, J. M. P., & Engelhard, I. M. (2020). Reduction of conditioned avoidance via contingency reversal. *Cognition & Emotion, 34*(6), 1284–1290. https://doi.org/10.1080/02699931.2020.1727417

Krypotos, A.-M., Blanken, T. F., Arnaudova, I., Matzke, D., & Beckers, T. (2017). A Primer on Bayesian analysis for experimental Psychopathologists. *Journal of Experimental Psychopathology, 8*(2), 140–157. https://doi.org/10.5127/jep.057316

Krypotos, A.-M., Klugkist, I., & Engelhard, I. M. (2017). Bayesian hypothesis testing for human threat conditioning research: An introduction and the condir R package. *European Journal of Psychotraumatology, 8*(sup1), Article 1314782. https://doi.org/10.1080/20008198.2017.1314782

Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized anxiety disorder is associated with overgeneralization of Classically conditioned fear. *Biological Psychiatry, 75*(11), 909–915. https://doi.org/10.1016/j.biopsych.2013.07.025

Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., & Grillon, C. (2010). Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. *American Journal of Psychiatry, 167*(1), 47–55. https://doi.org/10.1176/appi.ajp.2009.09030410

Lissek, S., & van Meurs, B. (2015). Learning models of PTSD: Theoretical accounts and psychobiological evidence. *International Journal of Psychophysiology, 98*(3), 594–605. https://doi.org/10.1016/j.ijpsycho.2014.11.006. Part 2.

Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans - Biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience & Biobehavioral Reviews, 80*, 703–728. https://doi.org/10.1016/j.neubiorev.2017.07.007

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E.-J. (2019). Jasp: Graphical statistical Software for Common statistical designs. *Journal of Statistical Software, 88*, 1–17. https://doi.org/10.18637/jss.v088.i02

Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It's not what you thought it was. *American Psychologist, 61*(1), 10–26. https://doi.org/10.1037/0003-066X.61.1.10

Olaya, B., Moneta, M. V., Miret, M., Ayuso-Mateos, J. L., & Haro, J. M. (2018). Epidemiology of panic attacks, panic disorder and the moderating role of age: Results from a population-based study. *Journal of Affective Disorders, 241*, 627–633. https://doi.org/10.1016/j.jad.2018.08.069

Rosebrock, L., Lambe, S., Mulhall, S., Petit, A., Loe, B. S., Saidel, S., Pervez, M., Mitchell, J., Chauhan, N., Prouten, E., Chan, C., Aynsworth, C., Murphy, E., Jones, J., Powling, R., Chapman, K., Dudley, R., Morrison, A., O'Regan, E., … Freeman, D. (2022). Understanding agoraphobic avoidance: The development of the Oxford cognitions and Defences questionnaire (O-CDQ). *Behavioural and Cognitive Psychotherapy, 50*(3), 257–268. https://doi.org/10.1017/S1352465822000030

Spielberger, C. D., Gorsuch, R.n, L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1977). *Stait–Trait Anxiety Inventory for AdultsMind Garden*. Redwood City.

Stegmann, Y., Reutter, M., Hutterer, K., Hildebrandt, L., Deckert, J., Deserno, L., Lonsdorf, T., Domschke, K., Pauli, P., Reif, A., Rosenkranz, K., Schümann, D., Schiele, M., Zwanzger, P., Andreatta, M., & Gamer, M. (2024). How to measure fear generalization? Reliability and validity of Commonly used Indices. *OSF*. https://doi.org/10.31234/osf.io/fk26y

Vanbrabant, K., Boddez, Y., Verduyn, P., Mestdagh, M., Hermans, D., & Raes, F. (2015). A new approach for modeling generalization gradients: A case for hierarchical models. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00652