



# Reducing Generalization of Conditioned Fear: Beneficial Impact of Fear Relevance and Feedback in Discrimination Training

Katharina Herzog<sup>1</sup>, Marta Andreatta<sup>1,2†</sup>, Kristina Schneider<sup>1</sup>, Miriam A. Schiele<sup>3</sup>, Katharina Domschke<sup>3</sup>, Marcel Romanos<sup>4</sup>, Jürgen Deckert<sup>5†</sup> and Paul Pauli<sup>1\*†</sup>

<sup>1</sup> Department of Psychology (Biological Psychology, Clinical Psychology, and Psychotherapy), Center of Mental Health, University of Würzburg, Würzburg, Germany, <sup>2</sup> Department of Psychology, Educational Sciences, and Child Studies, Erasmus University of Rotterdam, Rotterdam, Netherlands, <sup>3</sup> Department of Psychiatry and Psychotherapy, Medical Center—University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany, <sup>4</sup> Department of Child and Adolescent Psychiatry, Psychosomatics, and Psychotherapy, Center of Mental Health, Würzburg, Germany, <sup>5</sup> Department of Psychiatry, Psychosomatics, and Psychotherapy, Center of Mental Health, Würzburg, Germany

## OPEN ACCESS

### Edited by:

Ali Khatibi,  
University of Birmingham,  
United Kingdom

### Reviewed by:

Jonas Zaman,  
KU Leuven, Belgium  
Francisco Esteves,  
Mid Sweden University, Sweden

### \*Correspondence:

Paul Pauli  
pauli@psychologie.uni-wuerzburg.de

### †ORCID:

Marta Andreatta  
orcid.org/0000-0002-1217-8266  
Jürgen Deckert  
orcid.org/0000-0003-1008-4650  
Paul Pauli  
orcid.org/0000-0003-0692-6720

### Specialty section:

This article was submitted to  
Cognitive Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 09 February 2021

**Accepted:** 14 April 2021

**Published:** 01 June 2021

### Citation:

Herzog K, Andreatta M,  
Schneider K, Schiele MA,  
Domschke K, Romanos M, Deckert J  
and Pauli P (2021) Reducing  
Generalization of Conditioned Fear:  
Beneficial Impact of Fear Relevance  
and Feedback in Discrimination  
Training. *Front. Psychol.* 12:665711.  
doi: 10.3389/fpsyg.2021.665711

Anxiety patients over-generalize fear, possibly because of an incapacity to discriminate threat and safety signals. Discrimination trainings are promising approaches for reducing such fear over-generalization. Here we investigated the efficacy of a fear-relevant vs. a fear-irrelevant discrimination training on fear generalization and whether the effects are increased with feedback during training. Eighty participants underwent two fear acquisition blocks, during which one face (conditioned stimulus, CS+), but not another face (CS−), was associated with a female scream (unconditioned stimulus, US). During two generalization blocks, both CSs plus four morphs (generalization stimuli, GS1–GS4) were presented. Between these generalization blocks, half of the participants underwent a fear-relevant discrimination training (discrimination between CS+ and the other faces) with or without feedback and the other half a fear-irrelevant discrimination training (discrimination between the width of lines) with or without feedback. US expectancy, arousal, valence ratings, and skin conductance responses (SCR) indicated successful fear acquisition. Importantly, fear-relevant vs. fear-irrelevant discrimination trainings and feedback vs. no feedback reduced generalization as reflected in US expectancy ratings independently from one another. No effects of training condition were found for arousal and valence ratings or SCR. In summary, this is a first indication that fear-relevant discrimination training and feedback can improve the discrimination between threat and safety signals in healthy individuals, at least for learning-related evaluations, but not evaluations of valence or (physiological) arousal.

**Keywords:** fear generalization, feedback, discrimination training, fear-relevant training, classical conditioning

## INTRODUCTION

The adaptive mechanism of fear generalization prevents the encounter with unknown threats by extending previous learning to new cues (Dunsmoor and Paz, 2015). Thus, fear responses are elicited not only by cues (conditioned stimulus, CS+) predicting threat (i.e., the unconditioned stimulus, US; LeDoux and Pine, 2016) but also by stimuli, which were never associated with

the US but share physical or semantical properties with the threat-associated cues (Lissek et al., 2008; Dunsmoor and Paz, 2015; Dymond et al., 2015). Notably, this mechanism is exaggerated in anxiety patients compared to healthy controls, leading to over-generalization of fear (Lissek et al., 2010, 2014). Similarly, high trait-anxiety individuals, i.e., individuals at risk for anxiety disorders (Raymond et al., 2017), tend to over-generalize conditioned fear (Baumann et al., 2017; Stegmann et al., 2019; but see Torrents-Rodas et al., 2013; for a meta-analysis, see Sep et al., 2019).

Generalization of conditioned fear seems to be related to an incapacity in perceptually discriminating the relevant stimuli (Holt et al., 2014; Struyf et al., 2018; Zaman et al., 2019). For example, Holt et al. (2014) first assessed the participants' ability to discriminate the to-be-conditioned face from other faces, i.e., they assessed the just noticeable differences (JND) between CS+ vs. CS- and their morphs (i.e., generalization stimuli, GSs). After differential conditioning, a generalization test revealed increased US expectancy for the GSs below and at JND, indicating generalization, but not for those above JND. Thus, generalization of conditioned fear did not occur to those GSs, which were reliably discriminated from the CS+. Accordingly, improving the discrimination of CS+ may be a promising approach to reduce fear generalization.

Following this idea, several kinds of discrimination training have been developed to prevent fear generalization (Ginat-Frolich et al., 2017, 2019; Lommen et al., 2017). Specifically, two studies applied a discrimination training of fear-irrelevant stimuli (Ginat-Frolich et al., 2017, 2019). Each discrimination training trial started with a 4-s presentation of one abstract shape (target) at the center of a screen. Then, the target plus a new similar shape were presented on the left and right of the screen, and the participants had to identify the target. This training group was compared to a non-discriminative control task group where the participants had to indicate on which side of the screen each shape appeared. As a result, the authors observed reduced fear generalization of US expectancy in the first compared to the latter group.

Another study applied a discrimination training with stimuli sharing one perceptual characteristic (i.e., color) with the CS+ (Lommen et al., 2017). The participants had to decide if two stimuli presented shortly after each other differed in size or color or not at all. The participants assigned to a control task had to indicate instead which of two words was related to each presented stimulus. In contrast, this discrimination training which occurred before fear acquisition did not reduce the fear generalization of US expectancy but of avoidance behavior, again compared to the non-discriminative control task. These studies suggest that discrimination trainings with either fear-irrelevant or partial fear-relevant stimuli reduce fear generalization. However, please note that these trainings occurred before (Lommen et al., 2017) or after (Ginat-Frolich et al., 2017, 2019) fear learning. Therefore, only the latter two studies reflect a treatment training, while the former mirrors a prevention training.

One limitation of the above-mentioned studies is that they have tested the effects of discrimination training before having examined the generalization gradient of their participants.

However, anxiety disorder patients start treatments after fear acquisition and with a long history of over-generalization. Thus, we examined the fear-reducing effects of a discrimination training after both fear acquisition and a first demonstration of fear generalization. Our study was designed in order to reduce fear generalization, which occurred due to problems with perceptually discriminating CS+ from safe similar stimuli. As there is evidence that perceptual learning cannot be easily transferred to an unpracticed stimulus set (Furmanski and Engel, 2000), we examined whether discrimination trainings with fear-relevant vs. fear-irrelevant stimuli show higher efficacy. Additionally, we examined whether feedback on discrimination performance further enhances training efficiency. The latter expectation is based on previous studies demonstrating that positive feedback (Sasaki et al., 2009) on participants' performance and, in particular, when associated with reward (Weil et al., 2010) is able to improve stimulus perception.

To examine these hypotheses, our participants first underwent differential fear conditioning followed by two generalization blocks separated by discrimination training. Importantly, participants either learned to discriminate fear-relevant (discrimination of CS+ from GSs and CS-) or fear-irrelevant (discrimination of a specific line width from other lines) stimuli either with or without performance feedback. We hypothesized that the fear-relevant training reduces fear generalization more effectively than the fear-irrelevant training, and that this effect is especially strong if feedback is applied.

## MATERIALS AND METHODS

### Participants

The participants were recruited *via* an internet platform of the University of Würzburg. The exclusion criteria were psychiatric and neurological disorders, intake of psychoactive medication, excessive consumption of alcohol or nicotine, and pregnancy. Moreover, only participants between the ages of 18 and 50 were included. They were screened by a telephone interview. In the end, 80 participants were randomly divided into four groups (for details, see **Table 1**). Before the experiment started, all the participants read and signed the informed consent. The study was approved by the Ethical Committee of the Medical Board of the University of Würzburg and was conducted in accordance with the ethical principles of the Helsinki Declaration.

### Stimulus Material

Two neutral face expressions of a brunette and a blond woman (03F\_NE\_C, 10F\_NE\_C, NimStim Face Stimulus Set, Tottenham et al., 2009) served as conditioned stimuli (CS). One face stimulus (CS+) was associated with the aversive unconditioned stimulus (US) in 83% of the trials, while the other face (CS-) was never associated with the US (see also "Procedure"). The faces were counter-balanced across participants.

A compound of a 95-dB female desperate scream (International Affective Digitized Sounds, IADS, FemScream2, No. 276; Bradley and Lang, 1999) and the woman's fearful face expression were presented at the offset of CS+ for 1.5 s as US.

**TABLE 1** | Descriptive statistics of the four groups.

	Relevant_DT_noFB	Relevant_DT_FB	Irrelevant_DT_noFB	irrelevant_DT_FB	Comparisons
<i>N</i>	20	20	20	20	
Gender (♀)	10	10	10	10	
Age (SD)	24.25 (4.27)	26.10 (8.42)	25.70 (7.40)	24.75 (5.37)	$F_{(1, 76)} = 0.91, p = 0.344$
Language	17 German	17 German	18 German	19 German	$\chi^2(3) = 2.16, p = 0.540$
Handedness	19 right-hand	19 right-hand	19 right-hand	17 right-hand	$\chi^2(1) = 0.72, p = 0.396$
STAI (SD)	34.30 (9.14)	32.70 (6.53)	34.90 (10.18)	35.25 (7.41)	$F_{(1, 76)} = 0.27, p = 0.607$
BDI (SD)	5.1 (5.91)	3.70 (4.46)	5.25 (6.21)	5.55 (6.35)	$F_{(1, 76)} = 0.43, p = 0.513$

Four gradual morphs of the CSs were created in 20% steps by means of the software Squirrelz Morph (for details, see Schiele et al., 2016, version 2.1, Xiberpix, Solihull, United Kingdom) and used as generalization stimuli (GS). The CSs and GSs were presented for 6 s each.

After each block (see “Procedure”), the participants indicated the arousal (“how much stress/tension/arousal was elicited by this stimulus?”) and valence (“how pleasant vs. unpleasant was the stimulus for you?”) of each face on Likert scales from 1 (“calm” or “very unpleasant”) to 9 (“intense” or “pleasant”). Then, US expectancy ratings (“how high is the probability that you will hear the scream by this stimulus?”) were asked using a Likert scale ranging from 0 (“very unlikely”) to 100 (“very likely”), except after the habituation phase. The faces were rated in a fixed order (brunette woman, blond woman, morphs ranging from brunette to blond woman) and were presented for 1 s each before a Likert scale appeared.

For sample description and assessment of emotional state, the participants were asked to fill in selected questionnaires (for details, see **Supplementary Material**).

## Procedure

In the laboratory, the participants were seated on a comfortable chair, and they filled in the questionnaires. All stimuli were presented using Presentation software, version 16.0 (Neurobehavioral Systems, Inc., Albany, CA). The participants were instructed to passively view pictures and that they will occasionally hear an unpleasant loud sound, but the CS-US contingency was not disclosed.

The experimental procedure (**Figure 1A**) was based on previous studies (Lau et al., 2008; Schiele et al., 2016). Stimuli were presented in a pseudo-randomized order so that the same stimulus appeared not more than twice in a row. During inter-trial-intervals (ITI, time between two stimulus presentations), a white fixation cross was displayed in the center of the screen for 9–12 s randomly. During the habituation phase, both the CS+ and the CS− were presented four times each without any US.

The acquisition phase and generalization phase were divided into two identical blocks containing six presentations of each stimulus. In the acquisition blocks, US was delivered at CS+ offset in five trials out of six and never after CS−. During the generalization blocks, i.e., generalization tests, all faces were presented, and in three CS+ trials the US was delivered to prevent the extinction of conditioned fear. Between the

generalization blocks, the participants underwent discrimination trainings (see below).

After each block, the participants rated the arousal and valence of the faces and indicated their US expectancy (except after the habituation phase). Throughout the whole experiment, skin conductance responses (SCRs) were recorded. After the experiment, the participants completed the state questionnaires again.

## Discrimination Trainings

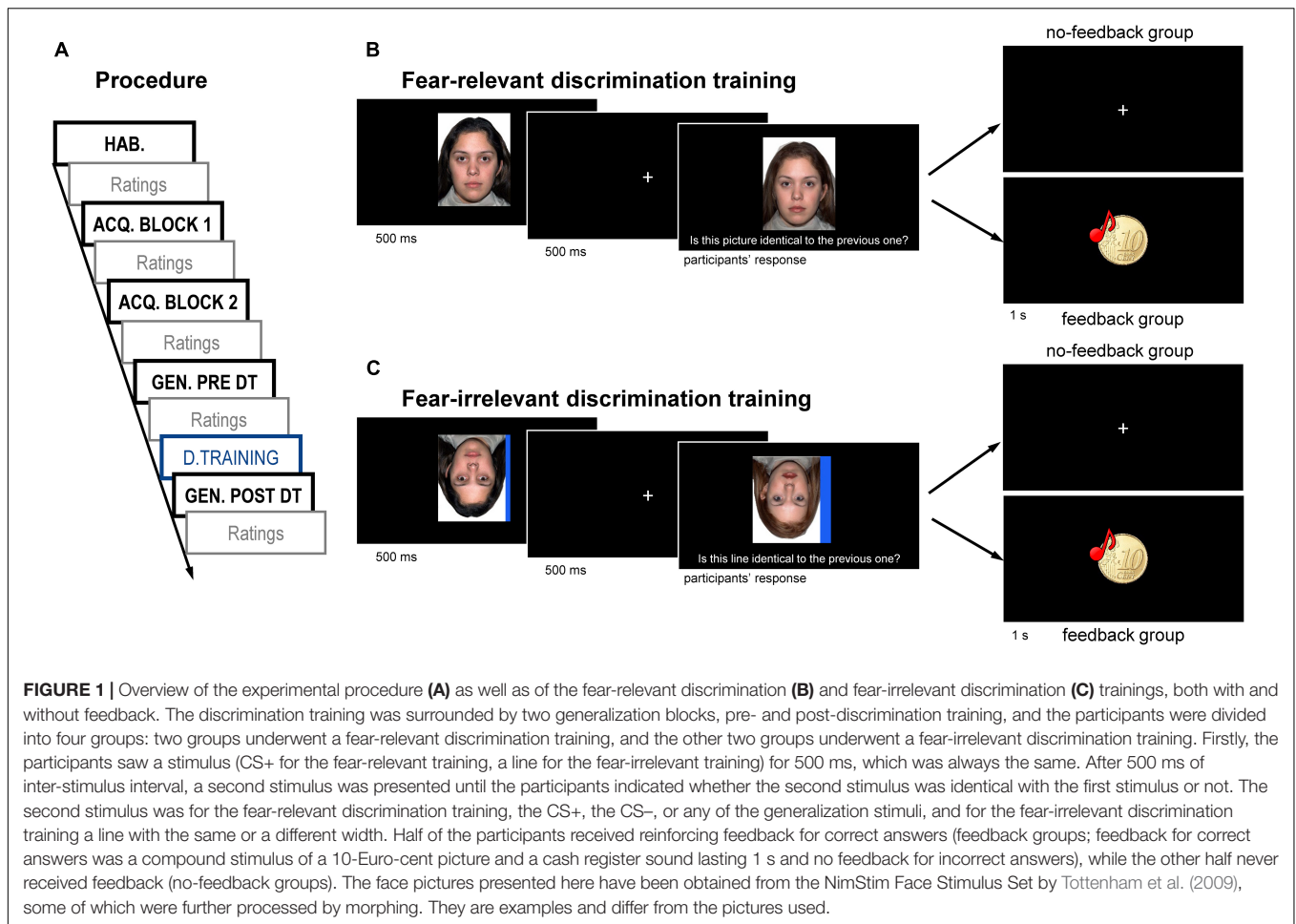
The training protocol was inspired by the discrimination task described previously (Holt et al., 2014). In all trainings, the faces and morphs were equally presented.

For the fear-relevant discrimination training (**Figure 1B**), every trial began with a 500-ms CS+ presentation. After a 500-ms inter-stimulus interval, a second stimulus was presented, which was either CS+, one GS, or CS−, below which the participants could read the question “Is this picture identical with the previous one?” The participants had to respond by pressing the S (yes) or L (no) button on the German keyboard with no time limitation. The face and question lasted on the screen until a response was given.

One group of participants (relevant\_DT\_FB) received reinforcing feedback for correct answers, consisting of a 10-Euro cent picture and a cash register sound (52.4 dB) lasting 1 s. No feedback was provided for incorrect answers. Another group (relevant\_DT\_noFB) received no feedback.

During the fear-irrelevant discrimination training (**Figure 1C**), all faces were turned upside down, and a blue line (35–79 pixels in width) was presented on the right or left side. Importantly, the width of the first stimulus line was always the thinnest line and was always combined with the upside-down CS+ face. All other aspects were as described for the fear-relevant discrimination training, and the participants had to indicate whether the width of the lines was identical or not. One group (irrelevant\_DT\_FB) received feedback again, but not the other group (irrelevant\_DT\_noFB).

In all trainings, the ITI lasted 1–3 s randomly. The participants altogether performed 50 training trials with a short break after 25 trials. For 20 out of the 50 trials, the second stimulus was the CS+ or the thinnest line, i.e., the pictures were identical. For the remaining 30 trials, the two pictures were not identical; each of the remaining five faces or lines were presented five times. While the thin line of the first stimulus was always combined with the upside down CS+, the line-face combinations of the



second stimulus varied. Nevertheless, as during the fear-relevant training, in the fear-irrelevant training every non-CS+ face was presented five times each.

Unfortunately, we cannot report the training performance of the two no-feedback groups due to a programming error. However, the training performances of the two feedback groups are reported in **Supplementary Material**.

## Data Recording and Reduction

Throughout the experiment, SCRs were recorded with two 8-mm Ag/AgCl electrodes attached at the thenar and hypothenar eminences of the non-dominant hand. For recording, Brainproducts V-Amp and BrainVision Recorder software (version 1.21, Brainproducts, Gilching, Germany) were used, having a sampling rate of 1,000 Hz and an online notch filter of 50 Hz. Offline analyses were run with BrainVision Analyzer Software (version 2.1, Brainproducts, Gilching, Germany). The electro-dermal signal was first filtered with a high cutoff filter of 1 Hz. In accordance with the guidelines (Boucsein et al., 2012), SCRs were defined as the difference in  $\mu\text{S}$  between response onset (900–4,000 ms after stimulus onset) and peak (2,000–6,000 ms after stimulus onset). Reactions smaller than  $0.02 \mu\text{S}$  were set to 0. Next, every reaction of each participant was range-corrected,

i.e., divided by the individual's strongest reaction to a face picture, i.e., CS or GS. Besides this, all SCRs were log-transformed into  $\log_{10}(\text{SCR} + 1)$ . The mean values were then calculated for each stimulus and experimental block. The participants ( $n = 5$ ) with an overall raw mean response smaller than  $0.02 \mu\text{S}$  were considered as non-responders and therefore excluded from the statistical testing of SCR. Accordingly,  $n = 20$  of the relevant\_DT\_FB group,  $n = 16$  of the relevant\_DT\_noFB group,  $n = 19$  of the irrelevant\_DT\_FB group, and  $n = 20$  of the irrelevant\_DT\_noFB group were included in the statistical analysis of SCR.

## Statistical Analysis

Statistical analyses were carried out in the R software environment (version 3.6.1) using the packages "afex" (version 0.26-0; Singmann et al., 2015) and "emmeans" (version 1.4.5; Lenth and Lenth, 2018).

Fear acquisition effects were analyzed with ANOVAs having stimulus (CS+, CS-) and block as within-subject factors; the factor block had three levels (habituation, acquisition 1, and acquisition 2) for arousal, valence, and SCR data and two levels (acquisition 1 and acquisition 2) for US expectancy data.

The discrimination training effects were investigated as follows: First, we calculated a generalization index (GI) for each



generalization block defined as the sum of all GS responses divided by the CS+ response:  $GI = [(GS1 + GS2 + GS3 + GS4)/CS+]$  (for details, see Lenaert et al., 2016). In other words, the GI represents the reaction to all GSs relative to CS+. As our fear-relevant discrimination training focuses on the discrimination of all GSs from CS+, the GI fits perfectly to reveal training-related changes in discrimination. To prevent division by zero, all US expectancy values were increased by 10 (i.e., the smallest step on the US expectancy scale), and all values of SCR were increased by  $\log(0.02 + 1)$  (defined as smallest SCR reaction  $>0$ , see above). Then, ANCOVAs on GI post-training were calculated, with fear relevance (relevant\_DT, irrelevant\_DT) and feedback (with, without) as between-subjects factors and GI pre-training as covariate. The ANCOVA is a commonly used method for comparing pre–post change across groups because of its good statistical power and only a slight bias for floor effects (Jennings and Cribbie, 2016).

The alpha level was set at .05, and Greenhouse–Geisser correction was applied for violation of the sphericity assumption. Where necessary, Bonferroni-corrected simple contrasts were calculated as *post hoc* tests. For effect sizes, partial eta-square values are reported.

## RESULTS

### Acquisition of Conditioned Fear

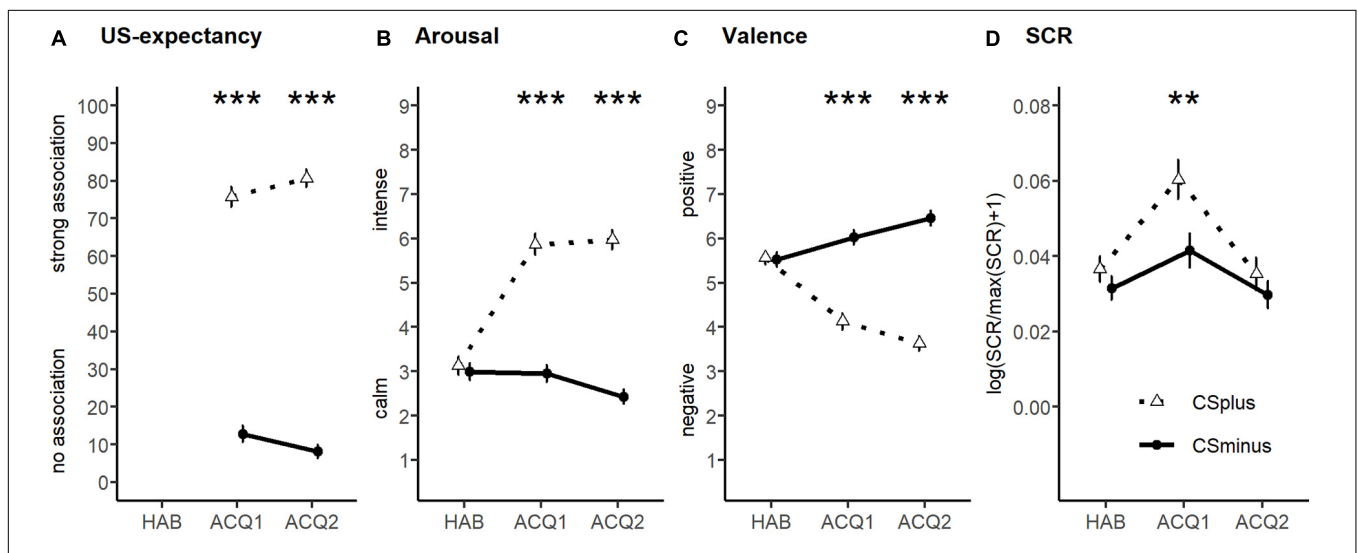
Successful fear acquisition is indicated by significant stimulus  $\times$  block interactions for all dependent variables, i.e., US expectancy [ $F_{(1, 79)} = 19.41, p < 0.001, \eta_p^2 = 0.20$ ; **Figure 2A**], arousal [ $F_{(1.89, 149.12)} = 86.34, p < 0.001, \eta_p^2 = 0.52$ ; **Figure 2B**], and valence ratings [ $F_{(1.86, 146.81)} = 82.83, p < 0.001, \eta_p^2 = 0.51$ ; **Figure 2C**] and for SCR [ $F_{(1.31, 96.87)} = 6.25,$

$p = -0.008, \eta_p^2 = 0.08$ ; **Figure 2D**]. *Post hoc* contrasts confirmed that the ratings were higher for CS+ vs. CS– regarding US expectancy [acquisition1:  $F_{(1, 79)} = 343.24, p < 0.001, \eta_p^2 = 0.81$ ; acquisition2:  $F_{(1, 79)} = 673.82, p < 0.001, \eta_p^2 = 0.90$ ; Bonferroni-corrected  $\alpha < 0.025$ ], arousal [acquisition1:  $F_{(1,79)} = 98.12, p < 0.001, \eta_p^2 = 0.55$ ; acquisition2:  $F_{(1, 79)} = 165.31, p < 0.001, \eta_p^2 = 0.68$ ; Bonferroni corrected  $\alpha < 0.017$ ], and valence [acquisition1:  $F_{(1, 79)} = 59.85, p < 0.001, \eta_p^2 = 0.431$ ; acquisition2:  $F_{(1, 79)} = 129.52, p < 0.001, \eta_p^2 = 0.621$ ; Bonferroni corrected  $\alpha < 0.017$ ], while no CS+ vs. CS– differences were evident for the habituation phase [arousal:  $F_{(1, 79)} = 0.62, p = 0.434, \eta_p^2 = 0.01$ ; valence:  $F_{(1,79)} = 0.05, p = 0.829, \eta_p^2 < 0.01$ ]. *Post hoc* contrasts for SCR also revealed no CS+ vs. CS– differences for habituation phase [ $F_{(1, 74)} = 2.24, p = 0.139, \eta_p^2 < 0.03$ ]. Regarding the acquisition phase, physiological arousal was higher to CS+ vs. CS– for acquisition 1 [ $F_{(1, 74)} = 11.81, p < 0.001, \eta_p^2 = 0.14$ ], but not for acquisition 2 [ $F_{(1, 74)} = 1.50, p = 0.225, \eta_p^2 = 0.02$ ], possibly due to habituation effects.

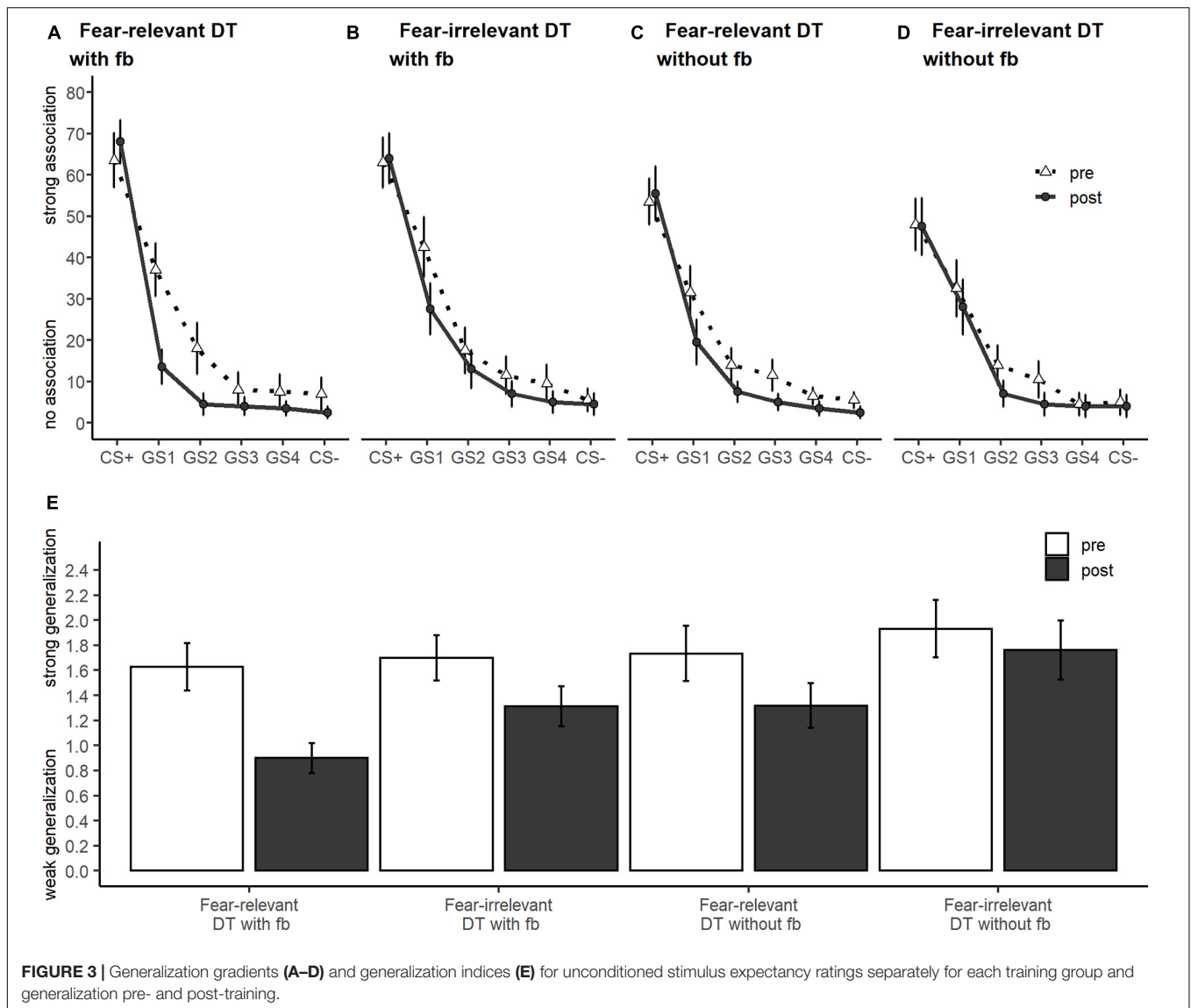
These ANOVAs also returned significant main effects of stimulus for all dependent variables [US expectancy:  $F_{(1, 79)} = 538.75, p < 0.001, \eta_p^2 = 0.87$ ; arousal:  $F_{(1, 79)} = 124.05, p < 0.001, \eta_p^2 = 0.61$ ; valence:  $F_{(1,79)} = 73.29, p < 0.001, \eta_p^2 = 0.48$ ; SCR:  $F_{(1, 72)} = 7.12, p = 0.009, \eta_p^2 = 0.09$ ] and main effects of block for arousal [ $F_{(1.54, 121.81)} = 45.51, p < 0.001, \eta_p^2 = 0.377$ ], valence [ $F_{(1.63, 129.00)} = 14.08, p < 0.001, \eta_p^2 = 0.15$ ], and SCR [ $F_{(1, 72)} = 24.44, p < 0.001, \eta_p^2 = 0.06$ ], but not for US expectancy [ $F_{(1, 79)} = 0.01, p = 0.907, \eta_p^2 < 0.01$ ].

### Discrimination Training Effects

For US expectancy ratings (the corresponding generalization gradients are depicted in **Figures 3A–D**), the ANCOVA on



**FIGURE 2 |** Habituation and acquisition ratings of unconditioned stimulus (US) expectancy (**A**), arousal (**B**), and valence (**C**) as well as skin conductance responses (**D**). Means (with SEs) of ratings and skin conductance responses (SCRs) are depicted for CS+ (white) and CS– (black) for habituation (HAB) for acquisition blocks (ACQ1, ACQ2). Acquisition of conditioned fear was successful as the CS+ vs. the CS– was rated as more likely predicting the US, more arousing, and more negatively valenced, and this elicited larger SCRs. Significance symbols indicate *post hoc* simple contrasts; \*\*\* $p < 0.001$ , \*\* $p < 0.01$ .



GIs revealed significant main effects of training [ $F_{(1,75)} = 6.56$ ,  $p = 0.012$ ,  $\eta_p^2 = 0.08$ ; **Figure 4A**] and feedback [ $F_{(1,75)} = 7.27$ ,  $p = 0.009$ ,  $\eta_p^2 = 0.09$ ; **Figure 4B**], but not their interaction [ $F_{(1,75)} = 0.04$ ,  $p = 0.850$ ,  $\eta_p^2 < 0.01$ ; **Figure 3E**], indicating that fear generalization was reduced more effectively by fear-relevant or rewarding feedback conditions.

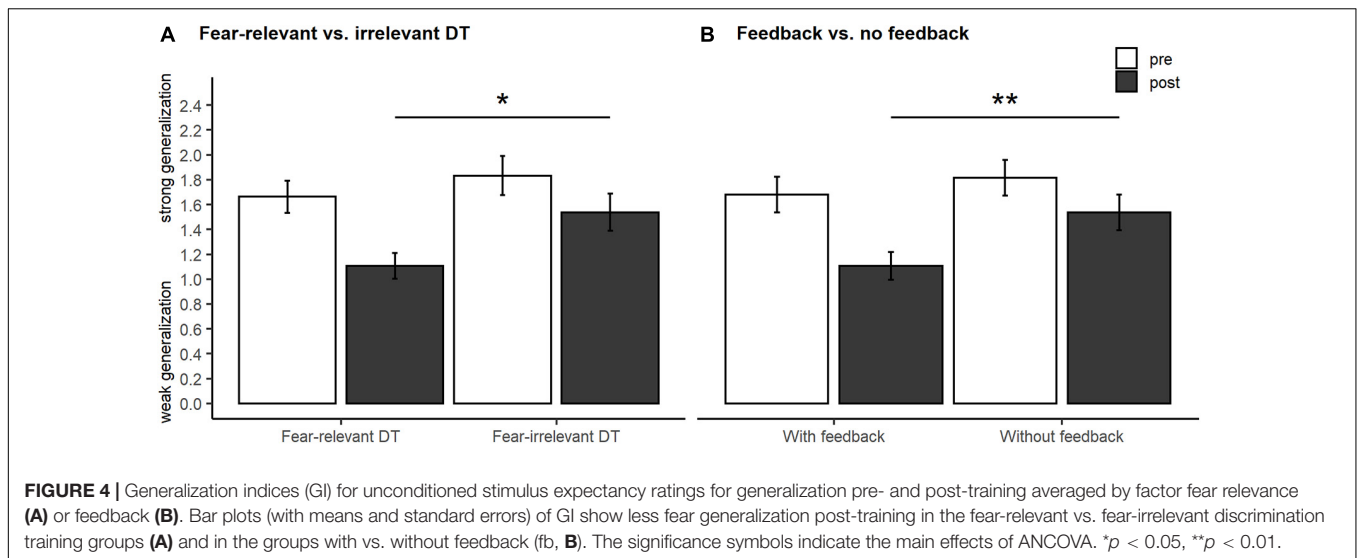
For arousal and valence ratings (for means and SDs, see **Table 2**; the corresponding generalization gradients are depicted in **Figures 5A,B**), the ANCOVAs on GIs revealed no effects involving the between-factors (all  $p > 0.106$ ).

The main effect of covariate was significant for all ratings [US expectancy:  $F_{(1,75)} = 72.13$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.49$ ; arousal:  $F_{(1,75)} = 9.71$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.11$ ; valence:  $F_{(1,75)} = 14.92$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.17$ ], suggesting positive associations between generalizations assessed pre- and post-training [US expectancy:  $r(79) = 0.69$ ,  $p < 0.001$ ; arousal:  $r(79) = 0.35$ ,  $p = 0.001$ ; valence:  $r(79) = 0.72$ ,  $p < 0.001$ ].

The analysis of SCR with ANCOVA on GIs (see **Table 2**; the corresponding generalization gradients are depicted in **Figure 5C**) revealed a marginal significant interaction of fear relevance and feedback [ $F_{(1,70)} = 3.7$ ,  $p = 0.059$ ,  $\eta_p^2 = 0.05$ ] only (all other  $p > 0.153$ ). Bonferroni-corrected ( $\alpha < 0.012$ ) *post hoc* simple contrasts did not confirm a significant impact of the combination of fear relevance and feedback (all  $p > 0.043$ ).

## DISCUSSION

This study was designed to test our hypothesis that the generalization of conditioned fear to perceptually similar stimuli can be reduced significantly better by a discrimination training with fear-relevant vs. fear-irrelevant stimuli. Additionally, we hypothesized that discrimination training effects profit by reinforcing feedback with greatest effects due to fear-relevant training with feedback. The experimental design followed the



logic of a treatment, which normally starts in patients with ascertained fear generalization. Thus, the discrimination training occurred after fear acquisition and in between two generalization tests. In consequence, this is the first study allowing one to verify training effects by comparing the generalization indices pre- vs. post-training.

Analysis of acquisition data validated that successful fear conditioning as the threat cue (CS+) vs. the safety cue (CS-) triggered greater arousal, more negative valence, greater US expectancy, and larger SCRs. Furthermore, analyses of the CS+ vs. CS- responses during the generalization tests pre- and post-discrimination training (see **Supplementary Material**) indicate that these conditioned fear responses remained stable throughout the experiment. Overall, these results corroborate previous studies using similar paradigms (Lissek et al., 2008; Schiele et al., 2016) and indicate the successful acquisition of fear toward the CS+.

Analyses of discrimination training effects on fear generalization suggest, on the one hand, stronger effects for the fear-relevant vs. the fear-irrelevant discrimination training

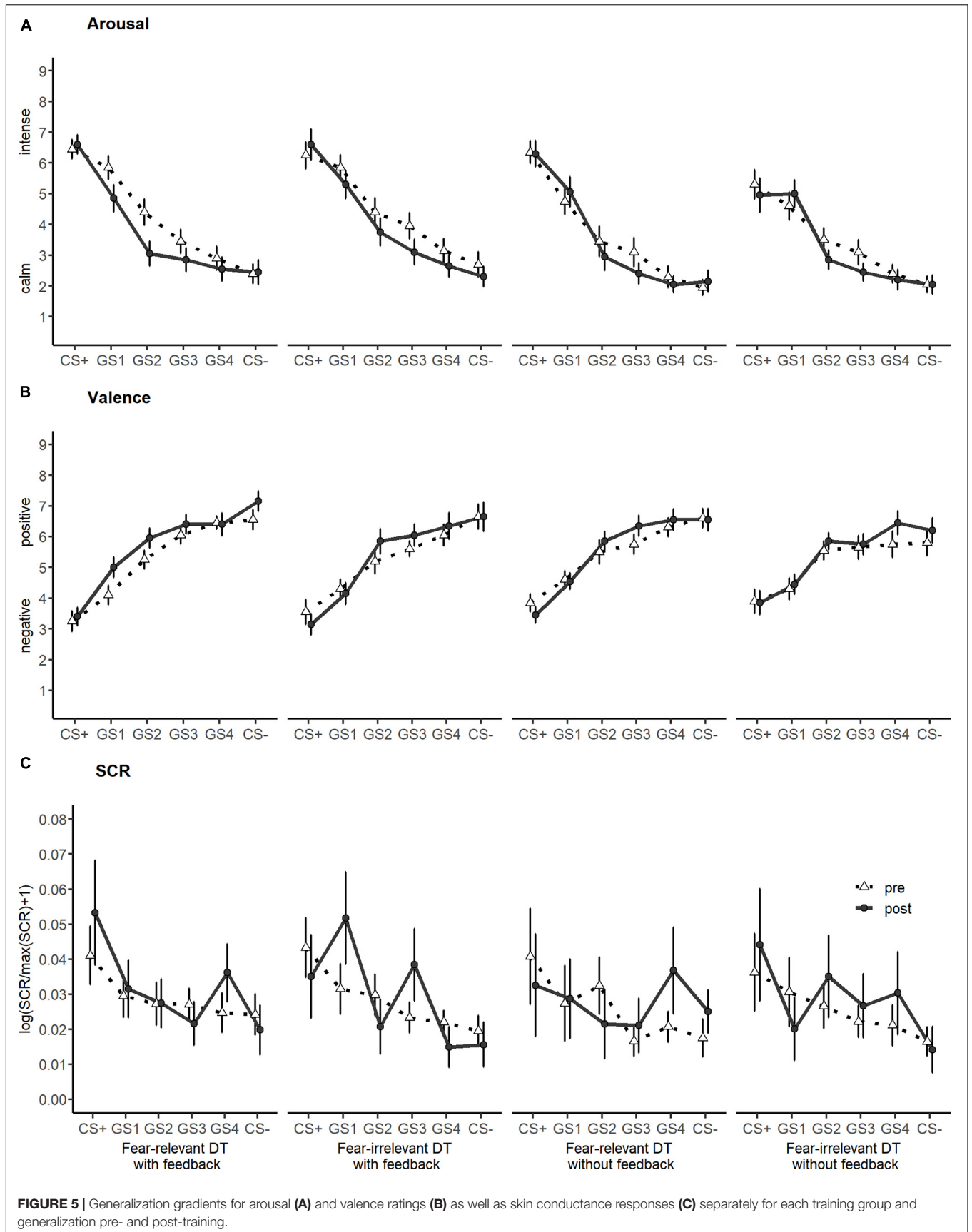
and, on the other hand, stronger effects for discrimination training with reinforcing feedback vs. without feedback, but with no interaction of effects. Thus, we conclude that a reduction of fear generalization is especially effective when the discrimination training is carried out with fear-relevant stimuli vs. fear-irrelevant stimuli or with reinforcing feedback vs. no feedback. However, we found no support for our hypothesis that a combination of training, i.e., training with fear-relevant stimuli *and* feedback, is especially effective in reducing fear generalization. Please note, however, that these findings are restricted to US expectancy ratings.

These results raise the following questions: Firstly, what contributes to the observed superior effect of a discrimination training with fear-relevant vs. fear-irrelevant stimuli on reducing fear generalization? Secondly, why is feedback during discrimination training effective in reducing fear generalization independently from the training stimuli? Thirdly, why are the observed discrimination training effects restricted to the US expectancy ratings?

Regarding the first question, the effect might be related to working memory (WM) processes, which are addressed during discrimination trainings as the participants have to keep in mind the first stimulus in order to compare it with the subsequent one (Lenaert et al., 2016). Considering that WM filters irrelevant from relevant information (Eysenck et al., 2007; Derakshan and Eysenck, 2009; Baddeley, 2012), it is very likely that a fear-relevant discrimination training improves the discrimination between relevant and irrelevant CS+ information, especially strong, and subsequently reduces the generalization effects. Through its relevance in attentional control (Eysenck et al., 2007; Derakshan and Eysenck, 2009), WM might facilitate perceptual learning (Amitay et al., 2014), possibly also thanks to inductive processes like the derivation of rules from WM (e.g., paying attention to visual features of stimuli; Livesey and McLaren, 2009). Rule learning processes might be further promoted by the fixed sequence of stimulus presentations during training (always beginning with CS+ or Line1), which could lead to an

**TABLE 2 |** Generalization indices for arousal and valence ratings as well as skin conductance responses, separately for each training group and generalization pre- and post-training.

	Relevant_DT _FB	Irrelevant_DT _FB	Relevant_DT _noFB	Irrelevant_DT _noFB
<b>Arousal</b>				
Pre (SD)	2.60 (0.70)	2.95 (1.09)	2.15 (0.83)	2.69 (0.76)
Post (SD)	2.11 (1.04)	2.51 (1.45)	2.05 (0.82)	3.26 (2.55)
<b>Valence</b>				
Pre (SD)	2.81 (0.87)	3.41 (2.64)	2.98 (0.74)	3.23 (0.94)
Post (SD)	2.58 (1.01)	2.71 (1.04)	2.62 (0.79)	3.11 (1.11)
<b>SCR</b>				
Pre (SD)	3.78 (2.48)	4.15 (3.05)	4.23 (3.01)	5.34 (4.28)
Post (SD)	5.80 (5.71)	9.83 (10.52)	7.63 (6.90)	5.41 (5.23)





overestimation of perceptual learning (García-Pérez and Alcalá-Quintana, 2020). Secondly, the results go in line with previous reports that a discriminative training with fear-relevant stimuli is able to reduce fear generalization (Lommen et al., 2017), although this study trained the participants before fear acquisition in analogy to prevention, while we trained the participants after fear acquisition in analogy to therapy. Moreover, it goes in line with more general findings regarding the transfer of discrimination learning to a test phase, which was found to be more effective after training relevant compared to irrelevant stimuli (Furmanski and Engel, 2000). In our study, the fear-relevant and fear-irrelevant training groups both have improved discrimination performance during training (for more details, see **Supplementary Material**). Consequently, the more efficient fear reduction of the fear-relevant training group might be related to a more effective transfer to the post-training generalization test. However, we know from previous studies that discrimination training, also with fear-irrelevant stimuli, reduces subsequent generalization effects in general (Ginat-Frolich et al., 2017, 2019). This can explain why the effects of fear-relevance are present but rather small.

Regarding the second question, we argue that feedback affects very general processes unrelated to the processing and discrimination of the presented stimuli. Previous studies revealed that feedback improves performance in general (Sasaki et al., 2009; Weil et al., 2010) as well as mood (Westermann et al., 1996). Moreover, reward facilitates attentional processes toward the reward-associated cue (Chelazzi et al., 2013) and increases motivation (Fishbach et al., 2010). Therefore, we assume that the implemented reinforcing feedback improved our participants' mood, attention, and motivation in general, which consequently improved their discrimination during the following generalization test. To verify this hypothesis, future studies might consider adding a control condition without discrimination training but a motivation condition, e.g., a task with comparable cognitive load and reinforcing feedback for correct responses.

Thirdly, we have to consider why the effects of the implemented discrimination trainings were restricted to US expectancy ratings. One possible explanation lies in the typology of the ratings. As previously suggested (Lonsdorf et al., 2017), US expectancy reflects cognitive learning processes, while valence and arousal ratings rather reflect affective learning processes. Therefore, new learning experiences can be seen more easily and earlier in US expectancy ratings. In line with this, the beneficial effects of our training were visible for US expectancy ratings. In contrast, extinction after evaluative conditioning, reflected in changes in valence, is hard to achieve (Vansteenwegen et al., 2006; Hofmann et al., 2010). This might also be true for discrimination learning, which could explain why we did not find training effects for the affective ratings. Moreover, the dissociation between affective and cognitive ratings suggests that, apart from perception, other (higher-order) processes (e.g., WM and inference rules) might have also determined fear generalization and its reduction in our paradigm. We suggest that future studies should explore whether and at what level an extensive discrimination training can improve discrimination as reflected in affective responses and, in consequence, reduce generalization.

In order to reveal discrimination training effects on the fear generalization of affective responses, future studies might consider assessing online ratings of valence and arousal during tests. We assessed the ratings at the end of every generalization block only and therefore could not determine discrimination training effects by comparing ratings from the end of generalization pre-training with the beginning of generalization post-training. Considering that online ratings might draw attention to US contingencies (Lonsdorf et al., 2017), we decided to assess intermittent ratings as several previous studies did (Haddad et al., 2012; Meulders et al., 2013; Holt et al., 2014; Roesmann et al., 2020).

Finally, we want to address the limitations of the current study. First, one limitation concerns the potential role of extinction and safety learning as additional mechanisms contributing to the higher efficacy of fear-relevant discrimination training. During the fear-relevant training, the subjects are exposed to the CS+ and the GSs without the USs presence, likely triggering extinction and safety learning processes. We can assume that our measured effects do not include extinction processes for CS+ because USs were presented again during generalization testing, thereby very likely undoing the CS+ extinction learning (as indicated by stable conditioned fear response to CS+ both pre- and post-training; see **Supplementary Material**). However, we cannot rule out that lower responses to certain GSs (and in consequence less generalization as reflected in the generalization index) stem from safety learning rather than some effects of perceptual discrimination training. Therefore, future studies should consider examining an additional group that did not undergo discrimination training but was exposed to the same amount of stimuli to control for potential extinction and safety learning effects. Second, no effects of fear relevance or feedback could be confirmed for physiological arousal, i.e., SCRs, which limits the effectiveness of our “therapeutic” approach of discrimination training to the verbal level. However, previous studies also did not find training effects for physiological measures (Ginat-Frolich et al., 2017) or did not include physiological measures (Lommen et al., 2017). In our study, contrary to all ratings, SCRs post-training were not significantly influenced by the pre-training level, as shown by the absent main effect of covariate. This means that all trainings strongly influenced the electrodermal activity in a way that dissolved the pre-training generalization pattern. Conceivably, the main reason lies in the nature of a training. While the participants passively observed the stimuli during the generalization blocks, during the discrimination training they actively fulfilled a task. The active task might have overlaid specific training effects. Besides this, the discriminative physiological arousal found during the first acquisition block disappeared in the second acquisition block, which is presumably due to habituation and could explain the absence of training effects for SCR. Third, the sample sizes are rather small. It should be considered, however, that the study presented here is a proof-of-principle study to establish a therapeutical approach of fear reduction by discrimination trainings. Fourth, we examined healthy participants who show moderate fear generalization as previous studies have also reported (Lissek et al., 2008). Considering the over-generalization

in highly anxious individuals and anxiety patients (Lissek et al., 2010, 2014; Laufer et al., 2016; Sep et al., 2019), future studies should investigate discrimination training in more clinic-relevant samples. Fifth, we have to acknowledge that, due to a programming error, we have no data on the discrimination training performance for the no-feedback groups. Therefore, we are unable to test whether and how feedback improved discrimination performance during training, and consequently, we cannot relate feedback effects on performance to later generalization effects.

In conclusion, this proof-of-principle study demonstrated the successful reduction of existing fear generalization by a fear-relevant discrimination training in healthy individuals at least for cognitive fear parameters, i.e., US expectancy. Moreover, we revealed that reinforcing feedback during discrimination training reduces the generalization of US expectancy, presumably *via* motivational mechanisms. Importantly, our study is the first to mimic a therapeutic approach as the discrimination training was performed after fear acquisition and a first demonstration of fear generalization. Accordingly, this is the first indication that pre-existing fear generalization can be reduced by means of a fear-relevant discrimination training or a reinforcing discrimination feedback, which in the future might be successfully used in patients with anxiety disorders.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the Medical Board of the University of Würzburg, Würzburg, Germany. The

## REFERENCES

- Amitay, S., Zhang, Y.-X., Jones, P. R., and Moore, D. R. (2014). Perceptual learning: top to bottom. *Vision Res.* 99, 69–77. doi: 10.1016/j.visres.2013.11.006
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Ann. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422
- Baumann, C., Schiele, M. A., Herrmann, M. J., Lonsdorf, T. B., Zwanzger, P., Domschke, K., et al. (2017). Effects of an anxiety-specific psychometric factor on fear conditioning and fear generalization. *Z. Psychol.* 225, 200–213. doi: 10.1027/2151-2604/a000304
- Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., et al. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology* 49, 1017–1034. doi: 10.1111/j.1469-8986.2012.01384.x
- Bradley, M. M., and Lang, P. J. (1999). *International Affective Digitized Sounds (Iads): Stimuli, Intruction Manual and Affective Ratings*. Gainesville, FL: University of Florida.
- Chelazzi, L., Perlato, A., Santandrea, E., and Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Res.* 85, 58–72. doi: 10.1016/j.visres.2012.12.005

patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

KH, MA, KD, MR, JD, and PP contributed to conception and design of the study. KS was involved in the conception and design of the discrimination training. MS contributed to the design of the fear generalization paradigm. KH organized the database and performed the statistical analysis. KH, MA, and PP wrote the first draft of the manuscript and critically revised each section. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

The work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project no. 44541416—TRR 58 “Fear, Anxiety, Anxiety Disorders” within project Z02 to PP, JD, MR, and KD. This publication was supported by the Open Access Publication Fund of the University of Würzburg.

## ACKNOWLEDGMENTS

We thank Paula Engelke, Johanna Holzäpfel, and Laura Rösler for helping us with collecting the data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.665711/full#supplementary-material>

- Derakshan, N., and Eysenck, M. W. (2009). Anxiety, processing efficiency, and cognitive performance. *Eur. Psychol.* 14, 168–176. doi: 10.1027/1016-9040.14.2.168
- Dunsmoor, J. E., and Paz, R. (2015). Fear generalization and anxiety: behavioral and neural mechanisms. *Biol. Psychiatry* 78, 336–343. doi: 10.1016/j.biopsych.2015.04.010
- Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., and Hermans, D. (2015). Fear generalization in humans: systematic review and implications for anxiety disorder research. *Behav. Ther.* 46, 561–582. doi: 10.1016/j.beth.2014.10.001
- Eysenck, M. W., Derakshan, N., Santos, R., and Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion* 7, 336–353. doi: 10.1037/1528-3542.7.2.336
- Fishbach, A., Eyal, T., and Finkelstein, S. R. (2010). How positive and negative feedback motivate goal pursuit. *Soc. Pers. Psychol. Compass* 4, 517–530. doi: 10.1111/j.1751-9004.2010.00285.x
- Furmanski, C. S., and Engel, S. A. (2000). Perceptual learning in object recognition: object specificity and size invariance. *Vision Res.* 40, 473–484. doi: 10.1016/S0042-6989(99)00134-0
- García-Pérez, M. A., and Alcalá-Quintana, R. (2020). Order effects in two-alternative forced-choice tasks invalidate adaptive threshold estimates. *Behav. Res. Methods* 52, 2168–2187. doi: 10.3758/s13428-020-01384-6

- Ginat-Frolich, R., Gendler, T., Marzan, D., Tsuk, Y., and Shechner, T. (2019). Reducing fear overgeneralization in children using a novel perceptual discrimination task. *Behav. Res. Ther.* 116, 131–139. doi: 10.1016/j.brat.2019.03.008
- Ginat-Frolich, R., Klein, Z., Katz, O., and Shechner, T. (2017). A novel perceptual discrimination training task: reducing fear overgeneralization in the context of fear learning. *Behav. Res. Ther.* 93, 29–37. doi: 10.1016/j.brat.2017.03.010
- Haddad, A., Pritchett, D., Lissek, S., and Lau, J. (2012). Trait anxiety and fear responses to safety cues: stimulus generalization or sensitization? *J. Psychopathol. Behav. Assess.* 34, 323–331. doi: 10.1007/s10862-012-9284-7
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., and Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychol. Bull.* 136, 390–421. doi: 10.1037/a0018916
- Holt, D. J., Boeke, E. A., Wolthuisen, R. P. F., Nasr, S., Milad, M. R., and Tootell, R. B. H. (2014). A parametric study of fear generalization to faces and non-face objects: relationship to discrimination thresholds. *Front. Hum. Neurosci.* 8:624. doi: 10.3389/fnhum.2014.00624
- Jennings, M., and Cribbie, R. A. (2016). Comparing pre-post change across groups: guidelines for choosing between difference scores, ANCOVA, and residual change scores. *J. Data Sci.* 14, 205–230. doi: 10.6339/jds.201604\_14(2).0002
- Lau, J. Y. F., Lissek, S., Nelson, E. E., Lee, Y., Roberson-Nay, R., Poeth, K., et al. (2008). Fear conditioning in adolescents with anxiety disorders: results from a novel experimental paradigm. *J. Am. Acad. Child Adolesc. Psychiatry* 47, 94–102. doi: 10.1097/chi.0b01e31815a5f01
- Laufer, O., Israeli, D., and Paz, R. (2016). Behavioral and neural mechanisms of overgeneralization in anxiety. *Curr. Biol.* 26, 713–722. doi: 10.1016/j.cub.2016.01.023
- LeDoux, J. E., and Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: a two-system framework. *Am. J. Psychiatry* 173, 1083–1093. doi: 10.1176/appi.ajp.2016.16030353
- Lenaert, B., van de Ven, V., Kaas, A. L., and Vlaeyen, J. W. S. (2016). Generalization on the basis of prior experience is predicted by individual differences in working memory. *Behav. Ther.* 47, 130–140. doi: 10.1016/j.beth.2015.10.001
- Lenth, R., and Lenth, M. R. (2018). Package 'lsmeans'. *Am. Stat.* 34, 216–221.
- Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., et al. (2008). Generalization of conditioned fear-conditioned startle in humans: experimental validation and clinical relevance. *Behav. Res. Ther.* 46, 678–687. doi: 10.1016/j.brat.2008.02.005
- Lissek, S., Kaczurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., and Grillon, C. (2014). Generalized anxiety disorder is associated with overgeneralization of classically conditioned fear. *Biol. Psychiatry* 75, 909–915. doi: 10.1016/j.biopsych.2013.07.025
- Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., et al. (2010). Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. *Am. J. Psychiatry* 167, 47–55. doi: 10.1176/appi.ajp.2009.09030410
- Livesey, E. J., and McLaren, I. P. (2009). Discrimination and generalization along a simple dimension: peak shift and rule-governed responding. *J. Exp. Psychol. Anim. Behav. Process.* 35, 554–565. doi: 10.1037/a0015524
- Lommen, M. J. J., Duta, M., Vanbrabant, K., de Jong, R., Juechems, K., and Ehlers, A. (2017). Training discrimination diminishes maladaptive avoidance of innocuous stimuli in a fear conditioning paradigm. *PLoS One* 12:e0184485. doi: 10.1371/journal.pone.0184485
- Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., et al. (2017). Don't fear 'fear conditioning': methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neurosci. Biobehav. Rev.* 77, 247–285. doi: 10.1016/j.neubiorev.2017.02.026
- Meulders, A., Vandebroek, N., Vervliet, B., and Vlaeyen, J. (2013). Generalization gradients in cued and contextual pain-related fear: an experimental study in healthy participants. *Front. Hum. Neurosci.* 7:345. doi: 10.3389/fnhum.2013.00345
- Raymond, J. G., Steele, J. D., and Seriès, P. (2017). Modeling trait anxiety: from computational processes to personality. *Front. Psychiatry* 8:1. doi: 10.3389/fpsy.2017.00001
- Roesmann, K., Wiens, N., Winker, C., Rehbein, M. A., Wessing, I., and Junghoefer, M. (2020). Fear generalization of implicit conditioned facial features – behavioral and magnetoencephalographic correlates. *Neuroimage* 205:116302. doi: 10.1016/j.neuroimage.2019.116302
- Sasaki, Y., Nanez, J. E., and Watanabe, T. (2009). Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* 11, 53–60. doi: 10.1038/nrn2737
- Schiele, M. A., Reinhard, J., Reif, A., Domschke, K., Romanos, M., Deckert, J., et al. (2016). Developmental aspects of fear: comparing the acquisition and generalization of conditioned fear in children and adults. *Dev. Psychobiol.* 58, 471–481. doi: 10.1002/dev.21393
- Sep, M. S. C., Steenmeijer, A., and Kennis, M. (2019). The relation between anxious personality traits and fear generalization in healthy subjects: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 107, 320–328. doi: 10.1016/j.neubiorev.2019.09.029
- Singmann, H., Bolker, B., Westfall, J., and Aust, F. (2015). *afex: Analysis of factorial experiments. R package version 0.13–145*.
- Stegmann, Y., Schiele, M. A., Schumann, D., Lonsdorf, T. B., Zwanzger, P., Romanos, M., et al. (2019). Individual differences in human fear generalization—pattern identification and implications for anxiety disorders. *Transl. Psychiatry* 9:307. doi: 10.1038/s41398-019-0646-8
- Struyf, D., Hermans, D., and Vervliet, B. (2018). Maximizing the generalization of fear extinction: exposures to a peak generalization stimulus. *Behav. Res. Ther.* 111, 1–8. doi: 10.1016/j.brat.2018.09.005
- Torrents-Rodas, D., Fullana, M. A., Bonillo, A., Caseras, X., Andión, O., and Torrubia, R. (2013). No effect of trait anxiety on differential fear conditioning or fear generalization. *Biol. Psychol.* 92, 185–190. doi: 10.1016/j.biopsycho.2012.10.006
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006
- Vansteenwegen, D., Francken, G., Vervliet, B., De Clercq, A., and Eelen, P. (2006). Resistance to extinction in evaluative conditioning. *J. Exp. Psychol. Anim. Behav. Process.* 32, 71–79. doi: 10.1037/0097-7403.32.1.71
- Weil, R. S., Furl, N., Ruff, C. C., Symmonds, G., Flandin, G., Dolan, R. J., et al. (2010). Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. *J. Neurophysiol.* 104, 1746–1757. doi: 10.1152/jn.00870.2009
- Westermann, R., Spies, K., Stahl, G., and Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. *Eur. J. Soc. Psychol.* 26, 557–580. doi: 10.1002/(sici)1099-0992(199607)26:4<557::aid-ejps769>3.0.co;2-4
- Zaman, J., Ceulemans, E., Hermans, D., and Beckers, T. (2019). Direct and indirect effects of perception on generalization gradients. *Behav. Res. Ther.* 114, 44–50. doi: 10.1016/j.brat.2019.01.006

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Herzog, Andreatta, Schneider, Schiele, Domschke, Romanos, Deckert and Pauli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.