

# Learning to fear depends on emotion and gaze interaction: The role of self-relevance in fear learning



Yoann Stussi<sup>a,b,\*</sup>, Tobias Brosch<sup>a,b</sup>, David Sander<sup>a,b</sup>

<sup>a</sup> Swiss Center for Affective Sciences, University of Geneva, Campus Biotech, Chemin des Mines 9, CH-1202 Geneva, Switzerland

<sup>b</sup> Laboratory for the Study of Emotion Elicitation and Expression (E3Lab), Department of Psychology, University of Geneva, Boulevard du Pont d'Arve 40, 1205 Geneva, Switzerland

## ARTICLE INFO

### Article history:

Received 12 December 2014

Received in revised form 17 June 2015

Accepted 18 June 2015

Available online 23 June 2015

### Keywords:

Emotional learning

Fear conditioning

Relevance detection

Emotion-gaze interaction

Skin conductance response

## ABSTRACT

Emotional learning is an adaptive function, however its psychological determinants are unclear. Here, we propose a new theoretical framework based on appraisal theories of emotion, which holds that emotional learning is modulated by a process of relevance detection. Testing the model, we predicted faster, larger acquisition and greater resistance to extinction of the conditioned response (CR) to self-relevant stimuli relative to stimuli with less relevance. We manipulated self-relevance through emotion and gaze direction of synthetic dynamic facial expressions during differential aversive conditioning. Results provided mixed evidence for our hypotheses. Critically, we revealed faster acquisition of the CR to angry faces with direct compared with averted gaze and greater resistance to extinction to fearful faces with averted relative to direct gaze. We conclude that the relevance detection hypothesis offers an appropriate theoretical framework allowing to (re)interpret existing evidence, incorporate our results, and propose a new research perspective in the study of emotional learning.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Emotional learning refers to the process by which a stimulus acquires an emotional significance. It represents a crucial adaptive function enabling an organism to respond appropriately to environmental stimuli by learning to identify their aversive or appetitive characteristics. Although neural mechanisms of emotional learning have been widely studied (see Phelps, 2006), psychological determinants underpinning the attribution of emotional value to stimuli, situations, events or behaviors remain unclear.

According to Seligman's (1970, 1971) preparedness theory, organisms are prepared to develop fear reactions to specific threatening stimuli based on biological predispositions shaped by the evolution of the species. The concept of preparedness has received strong empirical support demonstrating faster (e.g., Ho & Lipp, 2014; Öhman, Eriksson, & Olofsson, 1975) and larger (e.g., Fredrikson, Öhman, & Hugdahl, 1976; Öhman, Fredrikson, & Hugdahl, 1978) acquisition of a fear response to evolutionary fear-relevant stimuli than to fear-irrelevant stimuli. More importantly, critical evidence supporting the preparedness theory is the

resistance to extinction of learned fear to fear-relevant stimuli from phylogenetic origin, such as snakes (e.g., Öhman et al., 1975; Öhman, Fredrikson, Hugdahl, & Rimmö, 1976) and angry faces (e.g., Öhman & Dimberg, 1978), whereas extinction occurs rapidly for fear-irrelevant stimuli, such as flowers and happy faces (for a review, see Öhman & Mineka, 2001). Other studies showed that the learned fear response to phylogenetically fear-relevant stimuli was resistant to extinction even without awareness of these stimuli (e.g., Esteves, Parra, Dimberg, & Öhman, 1994; Öhman & Soares, 1993; for a review, see Öhman & Mineka, 2001). In a related line of research, Öhman and colleagues (Öhman, Flykt, & Esteves, 2001; Öhman, Lundqvist, & Esteves, 2001) showed that evolutionary relevant threatening stimuli automatically captured attention. Integrating these empirical findings with preparedness theory, Öhman and Mineka (2001) proposed the existence of an evolved fear module implemented in the human brain to explain the initiation and learning of fear, positing in particular that threatening stimuli encountered by the species during their evolutionary past benefit from enhanced learning compared with threatening stimuli from ontogenetic origin or non-threatening stimuli.

However, several empirical results have contradicted the hypothesized superiority of phylogenetically fear-relevant stimuli relative to ontogenetically fear-relevant stimuli. For instance, Hugdahl and Johnsen (1989) showed that resistance to extinction for gun pictures pointed toward the participants (i.e., a cultural

\* Corresponding author at: Campus Biotech, CISA – University of Geneva, Case Postale 60, CH-1211 Geneva 20, Switzerland.

E-mail address: [yoann.stussi@unige.ch](mailto:yoann.stussi@unige.ch) (Y. Stussi).

threat) associated with a loud noise was not statistically different from resistance to extinction for snake pictures directed toward the participants (i.e., a biological threat) associated with electrical shocks. Flykt, Esteves, and Öhman (2007) reported resistance to extinction for masked presentation of snake and gun pictures only when they were directed toward the participants, but not when they were directed away from the participants. These findings are consistent with the existence of a similar conditioning effect for biological and cultural threats. Similar results were also reported regarding attention processes. For instance, Brosch and Sharma (2005) showed that threatening stimuli from phylogenetic origin were not detected faster or more easily than threatening stimuli from ontogenetic origin in visual search. Jointly, these empirical data suggest, critically, that the key factor influencing fear learning seems to be fear-relevance rather than evolutionary history.

In line with this suggestion, appraisal theories of emotion may offer a different view, proposing that fear-relevant stimuli benefit from enhanced learning not because they are associated to fear through evolution, but because they are highly relevant for the individual (see Sander, Grafman, & Zalla, 2003). From this perspective, fear-relevant stimuli from phylogenetic origin are likely to be detected and automatically appraised as highly relevant to the organism's survival, which could account for the preferential allocation of attentional resources and enhanced fear learning to these stimuli. However, the process of relevance detection is not thought to differ for fear-relevant stimuli from phylogenetic and ontogenetic origin, respectively, or even as compared with relevant positive stimuli. Indeed, appraisal theories predict that highly relevant stimuli are prioritized in attentional processing, better learned and memorized regardless of their valence and evolutionary status per se (see Sander, Grandjean, & Scherer, 2005). Several studies have shown that attention (Brosch, Sander, Pourtois, & Scherer, 2008; Brosch, Sander, & Scherer, 2007; Pool, Brosch, Delplanque, & Sander, 2014) and memory (Montagrin, Brosch, & Sander, 2013) processes are specifically modulated on the basis of the affective relevance of the stimulus events. In particular, Brosch et al. (2008) demonstrated that spatial attention orienting processes were modulated equally strongly by angry adult faces and by baby faces, both at the behavioral and the neural level, indicating that relevant threatening and relevant positive stimuli benefited from a similar prioritization.

Taken together, these observations support no special status of fear-relevant stimuli from phylogenetic origin as compared with fear-relevant stimuli from ontogenetic origin and relevant positive stimuli. More importantly, they suggest that the preferential learning of "evolutionarily prepared" threat stimuli, such as snakes and angry faces, could be due to a more general mechanism of relevance detection rather than biological preparedness. To the best of our knowledge, no research has, however, directly investigated the role of relevance detection in emotional learning.

The purpose of this research was thus to make a first contribution to the study of relevance detection as a determinant of emotional learning. In this perspective, we investigated the impact of self-relevance on fear learning by manipulating the interaction between emotion and gaze in facial expressions (see Cristinzio, N'Diaye, Seeck, Vuilleumier, & Sander, 2010; N'Diaye, Sander, & Vuilleumier, 2009; Sander, Grandjean, Kaiser, Wehrle, & Scherer, 2007). According to appraisal theories, the processing of gaze direction modulates the detection and appraisal of self-relevance of a facial expression (Sander et al., 2003, 2007). Consistent with this perspective, previous work has shown the importance of the direction of facial and gaze display both in conditioning (Dimberg & Öhman, 1983) and attention (e.g., Juth, Lundqvist, Karlsson, & Öhman, 2005) to threat. For instance, Dimberg and Öhman (1983) reported resistance to extinction of conditioned fear to directed (i.e., head and eyes directed toward the participants)

angry faces but not to averted angry faces. The study by Juth et al. (2005) showed that directed faces were consistently detected more quickly and accurately than averted faces, both for angry and fearful expressions. In discrepancy with the results of Juth et al. (2005), appraisal theories posit, however, that a facial expression of anger is more self-relevant with direct gaze relative to averted gaze because it signals danger of being attacked, whereas a facial expression of fear is more self-relevant with averted gaze, since it signals a danger in the proximal environment (Sander et al., 2003). Building on these predictions and previous evidence of superior conditioning to fear-relevant stimuli as compared with fear-irrelevant stimuli (see Öhman & Mineka, 2001), we propose that – if a relevance detection mechanism is involved in fear learning – highly self-relevant stimuli should hence lead to the acquisition of a conditioned fear response that is more rapidly and largely acquired, and that is more resistant to extinction than stimuli with less relevance.

To test this hypothesis, we presented synthetic dynamic facial expressions to participants by manipulating self-relevance through emotion and gaze interaction in a differential aversive conditioning paradigm. We hypothesized that (a) angry faces with direct gaze compared with averted gaze and (b) fearful faces with averted gaze relative to direct gaze would lead to (1) a faster acquisition of the conditioned fear response, (2) the acquisition of a larger conditioned response, and (3) an enhanced resistance to extinction of the conditioned response.

## 2. Method

### 2.1. Participants

Ninety-three undergraduate students from the University of Geneva participated in the experiment for course credit and provided informed consent. Twelve participants were excluded from conditioning analysis due to technical problems, seven for displaying virtually no skin conductance responses, and 13 because of failure to acquire a conditioned response to at least one of the two conditioned stimuli positively predictive of the unconditioned stimulus. The exclusion criteria applied are widely accepted in the human conditioning literature (e.g., Olsson, Ebert, Banaji, & Phelps, 2005). The final sample included 61 students (12 men) aged between 18 and 42 years ( $M = 21.59$ ,  $SD = 3.52$ ).

### 2.2. Stimuli and apparatus

Four synthetic dynamic facial expressions<sup>1</sup> (two of anger and two of fear) either with direct gaze or averted gaze served as conditioned stimuli (CS). Dynamic expressions were created from four male identities of the Radboud Faces Database (Langner et al., 2010) using FACSGen (see Krumhuber, Tamarit, Roesch, & Scherer, 2012; Mumenthaler & Sander, 2015; Roesch et al., 2011). The same four face identities were presented to all participants. Each face identity served both as positively conditioned stimulus (CS+) and negatively conditioned stimulus (CS-), counterbalanced across participants. The stimuli were presented using MATLAB (The MathWorks, Inc., Natick, Massachusetts) with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and MATLAB Compiler Runtime. The unconditioned stimulus (US) was an electric stimulation (200-ms duration, 50 pulses/s) delivered to the right wrist through a Grass

<sup>1</sup> We used synthetic dynamic stimuli in order to increase ecological validity – emotional facial expressions and gaze shifts being dynamic in nature – while being able to perform highly controlled manipulations of (a) the time course of eye movements and unfolding of dynamic facial expressions, and (b) action units implied in the expressions and their intensity.

SD9 stimulator (Grass Medical Instruments, West Warwick, Rhodes Island) charged by a stabilized current.

The conditioned fear response (CR) was assessed through skin conductance response (SCR) measured with two 6-mm contact diameter Ag-AgCl electrodes filled with 0.5% NaCl electrolyte gel and attached to the medial phalanges of the second and third digit of the left hand. Data were continuously recorded at 1000 Hz through a BIOPAC MP150 system (Santa Barbara, California). An offline analysis of the SCR was carried out with AcqKnowledge 4.1 software (BIOPAC Systems Inc., Goleta, California).

### 2.3. Design and procedure

Emotion (Anger vs. Fear) was manipulated within-subjects and gaze direction (Averted vs. Direct) between-subjects ( $n=28$  for averted gaze condition,  $n=33$  for direct gaze condition) in order to facilitate the discrimination between the CSs. During the task, each participant viewed two different face identities showing anger expression, and two identities showing fear expression, either with averted gaze or direct gaze.

Before conditioning, participants set individually the stimulation amplitude to a level reported to be “uncomfortable, but not painful” (Olsson et al., 2005). The stimulation amplitude varied between 20 and 50 V. The conditioning procedure comprised three contiguous phases. The initial habituation phase consisted of four unreinforced presentations of each one of the four CSs. During the acquisition phase (six presentations of each CS), one stimulus (CS+) from each emotion was systematically paired with stimulation, while the other stimulus (CS−) from each emotion was never reinforced. The acquisition phase always started with a CS+ trial. During the extinction phase (six presentations of each CS), no stimulation was delivered.

Each trial began with a fixation cross for a duration that varied between 250 and 500 ms. A CS was then presented with a static neutral expression during 500 ms. A gaze shift was next initiated by a 100-ms movement, either from direct to averted (leftward or rightward), or from averted (leftward or rightward) to direct, preceding a 100-ms delay where the face remained static (see N'Diaye et al., 2009). Gaze shifts were controlled to occur the same amount of times from the center to the left and from the center to the right (averted gaze condition), or from the left to the center and from the right to the center (direct gaze condition). After the delay, facial expression evolved to an expression of anger or fear, the dynamic sequence lasting 400 ms (see N'Diaye et al., 2009), and was then presented statically during 4900 ms.<sup>2</sup> Each CS was presented for 6 s in a pseudorandomized order with an interstimulus interval ranging from 12 s to 15 s. During acquisition, the US was delivered 5.8 s after CS+ onset.

At the end of the experiment, the experimenter assessed participants' awareness of reinforcement contingencies with a recall test, in which participants were asked to describe the relation between the CSs and the US. Finally, participants were debriefed.

### 2.4. Response definition

SCR was measured for each trial as the peak-to-peak amplitude difference in skin conductance to the largest deflection starting in the 0.5–4.5 s temporal window following the emotional expression apex (1.6–5.6 s after CS onset) to capture the response to the

<sup>2</sup> The dynamics of eye movements and facial expressions was created by superimposing sequential static images generated by FACSGen corresponding to each stage of movement with MATLAB (see Mumenthaler & Sander, 2012, 2015) at a rate of 60 Hz. Six images were used to animate the gaze shift to the left, center, or right. Twenty-four images were used for the dynamics of emotional facial expressions.

interaction display of emotion and gaze. The minimal response criterion was  $0.02 \mu\text{S}$ . Responses below this criterion were scored as '0' and remained in the analyses. SCR data were low-pass filtered (Blackman  $-92 \text{ dB}$ , 1 Hz). SCRs were detected automatically with AcqKnowledge 4.1 software and checked for artifacts manually. The raw SCR scores were square-root-transformed to reduce positive skew and scaled according to each participant's mean square-root-transformed unconditioned response. The habituation means comprised the first four presentations of each CS. To examine the CR acquisition speed, the acquisition means were separated into an early (the first three presentations of each CS ensuing the first association of the CS+ with the US [i.e., trials 6–8; see Fig. 1]) and a late (the subsequent three presentations of each CS [i.e., trials 9–11; see Fig. 1]) phase (see Olsson, Carmona, Bolger, Downey, & Ochsner, 2013). The extinction means included the last five presentations of each CS following the first omission of the US (i.e., trials 12–16; see Fig. 1). The CR was operationalized as the mean scaled differential SCR (CS+ minus CS−) from the same emotion (see Olsson et al., 2005).

### 2.5. Statistical analyses

Separate mixed-design analyses of variance (ANOVAs) were conducted to analyze habituation, acquisition, and extinction data. To test specifically our a priori hypotheses, we performed planned contrasts analyses comparing the CR to (a) angry faces with direct vs. averted gaze, and (b) fearful faces with averted vs. direct gaze in (1) early acquisition, (2) late acquisition, and (3) extinction. Descriptive and residual analyses revealed the presence of outliers in the direct gaze condition for fear expression in extinction data (value smaller than the lower quartile minus three times the interquartile range, or value larger than the upper quartile plus three times the interquartile range; see Tukey, 1977). These outliers ( $n=7$ ) were therefore removed from the analyses of the extinction phase.

## 3. Results

Fig. 1 displays the means of scaled SCR magnitudes to angry and fearful faces throughout the habituation, acquisition, and extinction phases as a function of gaze direction, CS type, and trial. Results regarding the CR are depicted in Fig. 2.

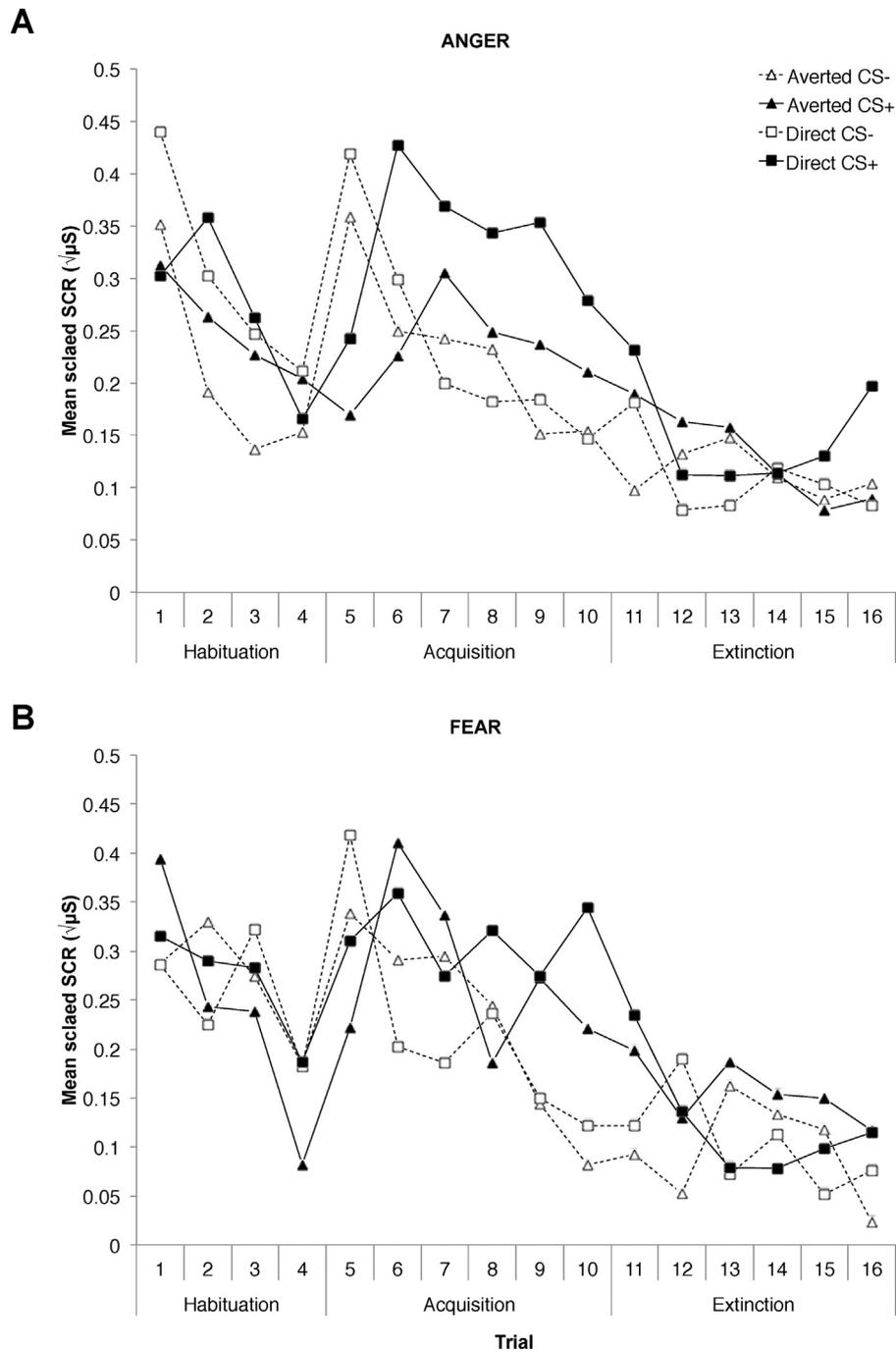
### 3.1. Habituation

The 2 (Emotion: Anger vs. Fear)  $\times$  2 (Gaze Direction: Averted vs. Direct) mixed-design ANOVA revealed a marginally significant interaction between emotion and gaze,  $F(1, 59)=3.12$ ,  $p=.083$ , partial  $\eta^2=.050$ . Post-hoc Tukey's HSD comparisons showed no significant differences between CSs (all  $ps>.27$ ).

### 3.2. Acquisition

The 2 (Emotion: Anger vs. Fear)  $\times$  2 (Gaze Direction: Averted vs. Direct)  $\times$  2 (Time: Early vs. Late) mixed-design ANOVA showed a main effect of gaze direction, indicating larger CR to direct-gaze faces than to averted-gaze faces,  $F(1, 59)=5.04$ ,  $p=.029$ , partial  $\eta^2=.079$ . Interaction effects between emotion and gaze direction, and between emotion, gaze direction, and time did not reach significance (both  $F_s < 1$ ). No other effects were significant within this phase (all  $ps > .14$ ).

With regard to the a priori hypothesis of faster CR acquisition to highly self-relevant faces compared with less self-relevant faces, the CR to angry faces with direct gaze ( $M=0.15$ ,  $SD=0.22$ ) was acquired faster than to angry faces with averted gaze ( $M=0.03$ ,  $SD=0.18$ ) in early acquisition trials,  $t(59)=2.24$ ,  $p=.029$ , 95% CI



**Fig. 1.** Mean scaled SCRs to (A) facial expressions of anger and (B) facial expressions of fear during habituation, acquisition, and extinction phases as a function of gaze direction (averted, direct), CS type (CS+, CS-), and trial.

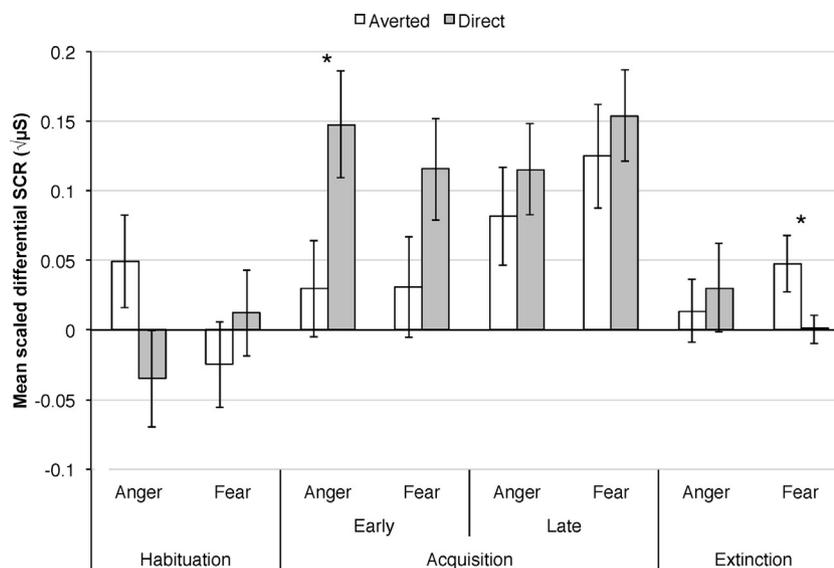
[0.01, 0.22],  $d=0.575$ . In contrast, the CR to fearful faces with averted gaze ( $M=0.03$ ,  $SD=0.19$ ) was not acquired faster than to fearful faces with direct gaze ( $M=0.12$ ,  $SD=0.21$ ),  $t(59)=-1.64$ ,  $p=.107$ , 95% CI [-0.19, 0.02]. Further analyses using one-sample  $t$ -tests showed no reliable differential conditioning to averted-gaze faces in early acquisition both for anger expressions,  $t(27)=0.86$ ,  $p=.397$ , 95% CI [-0.04, 0.10], and fear expressions,  $t(27)=0.85$ ,  $p=.404$ , 95% CI [-0.04, 0.10].

Concerning the hypothesis of larger CR acquisition to highly self-relevant faces comparatively to less self-relevant faces, the CR to angry faces with direct gaze ( $M=0.12$ ,  $SD=0.19$ ) was not larger than to angry faces with averted gaze ( $M=0.08$ ,  $SD=0.19$ )

in late acquisition trials,  $t(59)=0.70$ ,  $p=.488$ , 95% CI [-0.06, 0.13]. Similarly, there was no significant difference between fearful faces with averted gaze ( $M=0.12$ ,  $SD=0.20$ ) relative to direct gaze ( $M=0.15$ ,  $SD=0.19$ ),  $t(59)=-0.58$ ,  $p=.563$ , 95% CI [-0.13, 0.07].

### 3.3. Extinction

The 2 (Emotion: Anger vs. Fear)  $\times$  2 (Gaze Direction: Averted vs. Direct) mixed-design ANOVA revealed no interaction effect between emotion and gaze direction of facial expressions,  $F(1,$



**Fig. 2.** Mean conditioned response (scaled differential SCR) as a function of emotion (anger, fear) and gaze direction (averted, direct) during the habituation, (early and late) acquisition, and extinction phases. Error bars indicate standard errors. Asterisks indicate significant differences between conditions (\* $p < .05$ ).

52) = 2.07,  $p = .156$ , and no main effect of emotion, nor gaze direction (both  $F_s < 1$ ).

The analysis testing our a priori hypothesis that the CR to facial expressions of anger with direct gaze would be more resistant to extinction than to expressions of anger with averted gaze showed that angry faces with direct gaze ( $M = 0.03$ ,  $SD = 0.16$ ) did not lead to more resistance to extinction than angry faces with averted gaze ( $M = 0.01$ ,  $SD = 0.12$ ),  $t(52) = 0.43$ ,  $p = .666$ , 95% CI [-0.06, 0.09]. Additional analyses using one-sample  $t$ -tests revealed that the CR to angry faces exhibited no reliable resistance to extinction, neither with averted gaze,  $t(27) = 0.60$ ,  $p = .555$ , 95% CI [-0.03, 0.06], nor direct gaze,  $t(25) = 0.95$ ,  $p = .349$ , 95% CI [-0.04, 0.10]. Conversely and in accordance with our a priori hypothesis, the CR to fearful faces with averted gaze ( $M = 0.05$ ,  $SD = 0.11$ ) was more resistant to extinction than to fearful faces with direct gaze ( $M = 0.00$ ,  $SD = 0.05$ ),  $t(52) = 2.01$ ,  $p = .049$ , 95% CI [0.0001, 0.09],  $d = 0.661$ .<sup>3</sup>

#### 4. Discussion

In this study, we tested the hypothesis that relevance detection represents a general determinant of emotional learning, as suggested by appraisal theories of emotion (e.g., Sander et al., 2003). More specifically, we investigated the influence of self-relevance on fear learning by manipulating the emotion  $\times$  gaze interaction in facial expressions during differential aversive conditioning. We hypothesized faster, larger acquisition and enhanced resistance to extinction of the conditioned fear response to highly self-relevant stimuli relative to stimuli with less relevance.

Taken together, our results provide mixed evidence regarding our hypotheses. Although the overall data pattern in acquisition did not fully support the predicted influence of self-relevance on conditioned fear acquisition, faster fear acquisition to direct-gaze angry faces was congruent with our predictions and suggests that facial

<sup>3</sup> By comparison, the planned contrast comparing the CR to fearful faces with averted vs. direct gaze including outliers did not yield significance,  $t(59) = 1.31$ ,  $p = .195$ , 95% CI [-0.02, 0.11]. A more robust, nonparametric Mann-Whitney  $U$  test including outliers confirmed, however, a more resistant-to-extinction CR to fearful faces with averted gaze than to fearful faces with direct gaze,  $U = 300$ ,  $Z = 2.34$ ,  $p = .019$ ,  $r = .351$ . The results of the overall ANOVA and the planned contrast comparing the CR to angry faces with direct vs. averted gaze did not change with the inclusion of outliers.

expressions of anger are more rapidly associated with an aversive event when they are higher in self-relevance. However, the conditioned response to fearful faces was not more readily acquired to averted gaze than to direct gaze, which is at odds with our predictions.

In contrast to our second hypothesis predicting larger CR acquisition to self-relevant stimuli than to less self-relevant stimuli, we found similar responding to self-relevant and less self-relevant stimuli toward the end of acquisition. In the human conditioning literature, only few studies found overall superior acquisition to fear-relevant compared with fear-irrelevant stimuli (e.g., Fredrikson et al., 1976; Öhman et al., 1978; Olsson et al., 2005, Experiment 2). This could be explained by arguing that previously non-relevant (or fear-irrelevant) stimuli became relevant in signaling a threat through association with an aversive event (Sander et al., 2003) and thus showed no differential responding in respect to highly relevant (or fear-relevant) stimuli due to a ceiling effect. From this perspective, it may be argued that the investigation of superior fear acquisition to highly self-relevant as compared with less relevant stimuli does not seem to represent an optimal test to assess the hypothesis that emotional learning is modulated by a process of relevance detection.

On the other hand and although the overall pattern observed in extinction data was not supportive of the predicted impact of self-relevance on resistance to extinction, the greater persistence of the conditioned fear response to fearful faces with averted relative to direct gaze represents the most critical finding of our study. To the best of our knowledge, this is the first evidence establishing that facial expressions of fear with averted gaze previously associated with an aversive event can produce a more persistent conditioned fear response than facial expressions of fear with direct gaze. According to Öhman and Mineka (2001), the enhanced resistance to extinction of learned fear to fear-relevant stimuli from phylogenetic origin is a critical support of the fear module and preparedness theories. In addition, appraisal theories predict that facial expressions of fear with averted gaze are more self-relevant than facial expressions of fear with direct gaze (Cristinzio et al., 2010; N'Diaye et al., 2009; Sander et al., 2003, 2007). The more persistent learning to fearful faces with averted gaze relative to fearful faces with direct gaze underlies the influence of self-relevance on fear learning. In this view, it seems that this result represents the best evidence to date of the role of relevance detection in fear

learning. The lack of differential resistance to extinction of the conditioned fear to angry faces with direct gaze as compared with averted gaze is however more problematic and fails to replicate previous findings (Dimberg & Öhman, 1983). This absence of effect – as well as the failure to find statistical support for the interaction between emotion and gaze direction in the extinction phase – could reflect on the failure to replicate the enhanced resistance to extinction of the conditioned response to direct-gaze angry faces consistently reported in the human conditioning literature (e.g., Dimberg & Öhman, 1983, 1996; Öhman & Dimberg, 1978; Öhman & Mineka, 2001; Olsson & Phelps, 2004; Rowles, Lipp, & Mallan, 2012). The lack of resistance to extinction to angry faces with direct gaze turns out to be rather difficult to explain since a clear differential conditioning was observed during the acquisition phase. It may be due to the use of a mixed design instead of a between-subjects design, which implied the presentation of two different CSs+ that were no longer associated to the US during the extinction phase. This may have facilitated extinction of the conditioned fear in comparison with between-subjects designs using a single CS+ (see Bramwell, Mallan, & Lipp, 2014; but see also Olsson et al., 2005). It remains however unclear to what extent the presentation of two CSs+ unpaired with the US during extinction should have preferentially facilitated extinction of the conditioned response to angry faces with direct gaze as compared with, for instance, fearful faces with averted gaze. Another potential explanation could refer to the use of synthetic faces instead of real human faces. It could be argued that synthetic faces expressing anger may have a lesser inherent threat value than human angry faces. However, it is again not clear why the use of synthetic faces should have prevented the resistance to extinction of the conditioned response to direct-gaze angry faces, but not to averted-gaze fearful faces.

An interpretation of our results in terms of biological preparedness could still be advanced by asserting that direct-gaze angry faces and averted-gaze fearful faces represent a case of stimuli that are “more prepared” than averted-gaze angry faces and direct-gaze fearful faces. However, we do not think that this account provides a better explanation of our findings than the relevance detection framework. If the biological preparedness perspective seems to hold for angry faces (see Dimberg & Öhman, 1983, 1996), the case of fearful faces is less clear. It could be indeed argued that the fear module and preparedness theories consider fear expressions as biologically prepared stimuli regardless of gaze direction because both averted- and direct-gaze fearful faces convey information about potential threat and should predict enhanced persistence of learned fear to fearful faces both with averted and direct gaze. In accord with this view, previous findings showing superior conditioning to fearful faces with direct gaze have been interpreted in terms of biological preparedness (see Dimberg & Öhman, 1996; Lanzetta & Orr, 1986). In sum, the fear module and preparedness theories fail to be clear in predicting enhanced resistance to extinction to averted-gaze fearful faces as compared with direct-gaze fearful faces and offer at best post-hoc explanations of this finding, while appraisal theories provide a clear a priori prediction. In this regard, we consider that the relevance detection framework positing enhanced learning of highly relevant stimuli is the most appropriate theoretical framework to incorporate our results.

Nonetheless, it is important to highlight that the data collected in this study only partially support the predictions derived from the relevance detection hypothesis of emotional learning. Furthermore, we acknowledge limitations in this study. At first, the control of gaze-shifts direction within-subjects could have specifically hindered the acquisition and the extinction of the conditioned response to averted-gaze stimuli relative to direct-gaze stimuli. Participants in the direct gaze condition had to learn that the CS+ angry and fearful faces with gaze directed toward them were

followed by an electric stimulation, whereas participants in the averted gaze condition had to learn that a stimulation followed the CS+ angry and fearful faces both when their gaze was averted leftward and rightward, which could have conducted to a general slower learning in the averted gaze condition both in acquisition – as indicated by the lack of reliable differential conditioning to averted-gaze faces during early acquisition – and extinction. This may represent potential grounds for the failure to find strong statistical support for the interaction between emotion and gaze direction in acquisition and extinction. However, we found only faster acquisition to highly self-relevant stimuli respective to direct gaze (i.e., angry faces with direct gaze), but not to stimuli with less self-relevance (i.e., fearful faces with direct gaze). Similarly, greater resistance to extinction was revealed only to highly self-relevant stimuli with regard to averted gaze (i.e., fearful faces with averted gaze), but not when stimuli were less self-relevant (i.e., angry faces with averted gaze). Hence, significant differences were found only to faces with high self-relevance respective to gaze direction in the early acquisition and extinction phases, thus providing further support for our hypotheses. Another limitation could be the manipulation of gaze direction as a between-subjects factor. As mentioned earlier, we preferred to manipulate emotion rather than gaze direction within-subjects in order to facilitate the discrimination between the CSs. However, manipulating gaze direction within-subjects could have allowed a more pure test of differences between averted and direct gaze conditions.

Notwithstanding these limitations and the fact that we did not find all the predicted effects, our study provides initial evidence that self-relevance may influence fear learning and that relevance detection may represent a general mechanism determining emotional learning. Our results suggest that the theoretical framework derived from appraisal theories could provide a credible alternative to the fear module and preparedness theories. Importantly, the notion of relevance detection captures the dimension of evolutionary significance but is not limited to it by referring also to the other concerns of the individual (Sander, 2013). In this fashion, the relevance detection framework answers the recent call for a theoretical model accommodating the influence of biological and cultural factors on fear learning (see Mallan, Lipp, & Cochrane, 2013). A conceptual approach based on relevance detection goes even beyond this call by proposing – in contrast to the fear module and preparedness theories – a general mechanism of emotional learning that is not limited to fear learning and predicting that stimuli detected and appraised as highly relevant to the organism’s needs, goals, values or well-being benefit from enhanced learning independently of their intrinsic valence and evolutionary status per se. Thus, relevance detection represents a more flexible approach than biological preparedness and emerges, in our opinion, as a very promising new theoretical framework to give a better insight into the understanding of basic mechanisms underlying emotional learning in humans. However, it is important to note that the present experiment represents only a first step in the study of the role of relevance detection in emotional learning and further research is needed to better outline this role, primarily by replicating the results that do support the specific predictions of the relevance detection framework, while replicating at the same time the basic findings of human conditioning literature. In this perspective, future studies should investigate, in particular, the learning of relevant positive stimuli in conditioning.

## Acknowledgements

This research was supported by the National Center of Competence in Research (NCCR) Affective Sciences, financed by the Swiss National Science Foundation (51NF40-104897), and hosted by the

University of Geneva. We thank Sylvain Delplanque and Christian Mumenthaler for their valuable help in this project.

## References

- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. <http://dx.doi.org/10.1163/156856897x00357>
- Bramwell, S., Mallan, K. M., & Lipp, O. V. (2014). Are two threats worse than one? The effects of face race and emotional expression on fear conditioning. *Psychophysiology*, 51, 152–158. <http://dx.doi.org/10.1111/psyp.12155>
- Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: Rapid spatial orienting toward positive emotional stimuli. *Psychological Science*, 19, 362–370. <http://dx.doi.org/10.1111/j.1467-9280.2008.02094.x>
- Brosch, T., Sander, D., & Scherer, K. R. (2007). That baby caught my eye... Attention capture by infant faces. *Emotion*, 7, 685–689. <http://dx.doi.org/10.1037/1528-3542.7.3.685>
- Brosch, T., & Sharma, D. (2005). The role of fear-relevant stimuli in visual search: A comparison of phylogenetic and ontogenetic stimuli. *Emotion*, 5, 360–364. <http://dx.doi.org/10.1037/1528-3542.5.3.360>
- Cristinzio, C., N'Diaye, K., Seeck, M., Vuilleumier, P., & Sander, D. (2010). Integration of gaze direction and facial expression in patients with unilateral amygdala damage. *Brain*, 133, 248–261. <http://dx.doi.org/10.1093/brain/awp255>
- Dimberg, U., & Öhman, A. (1983). The effects of directional facial cues on electrodermal conditioning to facial stimuli. *Psychophysiology*, 20, 160–167. <http://dx.doi.org/10.1111/j.1469-8986.1983.tb03282.x>
- Dimberg, U., & Öhman, A. (1996). Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion*, 20, 149–182. <http://dx.doi.org/10.1007/BF02253869>
- Esteves, F., Parra, C., Dimberg, U., & Öhman, A. (1994). Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear-relevant facial stimuli. *Psychophysiology*, 31, 375–385. <http://dx.doi.org/10.1111/j.1469-8986.1994.tb02446.x>
- Flykt, A., Esteves, F., & Öhman, A. (2007). Skin conductance responses to masked conditioned stimuli: Phylogenetic/ontogenetic factors versus direction of threat? *Biological Psychology*, 74, 328–336. <http://dx.doi.org/10.1016/j.biopsycho.2006.08.004>
- Fredrikson, M., Hugdahl, K., & Öhman, A. (1976). Electrodermal conditioning to potentially phobic stimuli in male and female subjects. *Biological Psychology*, 4, 305–314. [http://dx.doi.org/10.1016/0301-0511\(76\)90021-1](http://dx.doi.org/10.1016/0301-0511(76)90021-1)
- Ho, Y., & Lipp, O. V. (2014). Faster acquisition of conditioned fear to fear-relevant than to nonfear-relevant conditional stimuli. *Psychophysiology*, 51, 810–813. <http://dx.doi.org/10.1111/psyp.12223>
- Hugdahl, K., & Johnsen, B. H. (1989). Preparedness and electrodermal fear-conditioning: Ontogenetic vs. phylogenetic explanations. *Behaviour Research Therapy*, 27, 269–278. [http://dx.doi.org/10.1016/0005-7967\(89\)90046-6](http://dx.doi.org/10.1016/0005-7967(89)90046-6)
- Juth, P., Lundqvist, D., Karlsson, A., & Öhman, A. (2005). Looking for foes and friends: Perceptual and emotional factors when finding a face in the crowd. *Emotion*, 5, 379–395. <http://dx.doi.org/10.1037/1528-3542.5.4.379>
- Krumhuber, E. G., Tamarit, L., Roesch, E. B., & Scherer, K. R. (2012). FACSGen 2.0 animation software: Generating three-dimensional FACS-valid facial expressions for emotion research. *Emotion*, 12, 351–363. <http://dx.doi.org/10.1037/a0026632>
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, 24, 1377–1388. <http://dx.doi.org/10.1080/0269930903485076>
- Lanzetta, J. T., & Orr, S. P. (1986). Excitatory strength of expressive faces: Effects of happy and fear expressions and context on the extinction of a conditioned fear response. *Journal of Personality and Social Psychology*, 50, 190–194. <http://dx.doi.org/10.1037/0022-3514.50.1.190>
- Mallan, K. M., Lipp, O. V., & Cochrane, B. (2013). Slithering snakes, angry men and out-group members: What and whom are we evolved to fear? *Cognition & Emotion*, 27, 1168–1180. <http://dx.doi.org/10.1080/02699931.2013.778195>
- Montagrin, A., Brosch, T., & Sander, D. (2013). Goal conduciveness as a key determinant of memory facilitation. *Emotion*, 13, 622–628. <http://dx.doi.org/10.1037/a0033066>
- Mumenthaler, C., & Sander, D. (2012). Social appraisal influences emotion recognition. *Journal of Personality and Social Psychology*, 102, 1118–1135. <http://dx.doi.org/10.1037/a0026885>
- Mumenthaler, C., & Sander, D. (2015). Automatic integration of social information in emotion recognition. *Journal of Experimental Psychology: General*, 144, 392–399. <http://dx.doi.org/10.1037/xge0000059>
- N'Diaye, K., Sander, D., & Vuilleumier, P. (2009). Self-relevance processing in the human amygdala: Gaze direction, facial expression, and emotion intensity. *Emotion*, 9, 798–806. <http://dx.doi.org/10.1037/a0017845>
- Öhman, A., & Dimberg, U. (1978). Facial expressions as conditioned stimuli for electrodermal responses: A case of preparedness? *Journal of Personality and Social Psychology*, 36, 1251–1258. <http://dx.doi.org/10.1037/0022-3514.36.11.1251>
- Öhman, A., Eriksson, A., & Olofsson, C. (1975). One-trial learning and superior resistance to extinction of autonomic responses conditioned to potentially phobic stimuli. *Journal of Comparative and Physiological Psychology*, 88, 619–627. <http://dx.doi.org/10.1037/h0078388>
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130, 466–478. <http://dx.doi.org/10.1037/0096-3445.130.3.466>
- Öhman, A., Fredrikson, M., & Hugdahl, K. (1978). Orienting and defensive responding in the electrodermal system: Palmar-dorsal differences and recovery rate during conditioning to potentially phobic stimuli. *Psychophysiology*, 15, 93–101. <http://dx.doi.org/10.1111/j.1469-8986.1978.tb01342.x>
- Öhman, A., Fredrikson, M., Hugdahl, K., & Rimmö, P.-A. (1976). The premise of equipotentiality in human classical conditioning: Conditioned electrodermal responses to potentially phobic stimuli. *Journal of Experimental Psychology: General*, 105, 313–337. <http://dx.doi.org/10.1037/0096-3445.105.4.313>
- Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, 80, 381–396. <http://dx.doi.org/10.1037/0022-3514.80.3.381>
- Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483–522. <http://dx.doi.org/10.1037/0033-295X.108.3.483>
- Öhman, A., & Soares, J. F. (1993). On the automatic nature of phobic stimuli: Conditioned electrodermal responses to masked fear-relevant stimuli. *Journal of Abnormal Psychology*, 102, 121–123. <http://dx.doi.org/10.1037/0021-843X.102.1.121>
- Olsson, A., Carmona, S., Downey, G., Bolger, N., & Ochsner, K. N. (2013). Learning biases underlying individual differences in sensitivity to social rejection. *Emotion*, 13, 616–621. <http://dx.doi.org/10.1037/a0033150>
- Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science*, 309, 785–787. <http://dx.doi.org/10.1126/science.1113551>
- Olsson, A., & Phelps, E. A. (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychological Science*, 15, 822–828. <http://dx.doi.org/10.1111/j.0956-7976.2004.00762.x>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. <http://dx.doi.org/10.1163/156856897x00366>
- Phelps, E. A. (2006). Emotion and cognition: Insights from studies on the human amygdala. *Annual Review of Psychology*, 57, 27–53. <http://dx.doi.org/10.1146/annurev.psych.56.091103.070234>
- Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2014). Where is the chocolate? Rapid spatial orienting toward stimuli associated with primary rewards. *Cognition*, 130, 348–359. <http://dx.doi.org/10.1016/j.cognition.2013.12.002>
- Roesch, E. B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., & Scherer, K. R. (2011). FACSGen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units. *Journal of Nonverbal Behavior*, 35, 1–16. <http://dx.doi.org/10.1007/s10919-010-0095-9>
- Rowles, M. E., Lipp, O. V., & Mallan, K. M. (2012). On the resistance to extinction of fear conditioned to angry faces. *Psychophysiology*, 49, 375–380. <http://dx.doi.org/10.1111/j.1469-8986.2011.01308.x>
- Sander, D. (2013). *Models of emotion: The affective neuroscience approach*. In J. L. Armony, & P. Vuilleumier (Eds.), *The Cambridge Handbook of human affective neuroscience* (pp. 5–53). Cambridge, UK: Cambridge University Press.
- Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences*, 14, 303–316. <http://dx.doi.org/10.1515/REVNEURO.2003.14.4.303>
- Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology*, 19, 470–480. <http://dx.doi.org/10.1080/09541440600757426>
- Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks*, 18, 317–352. <http://dx.doi.org/10.1016/j.neunet.2005.03.001>
- Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review*, 77, 406–418. <http://dx.doi.org/10.1037/h0029790>
- Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy*, 2, 307–320. [http://dx.doi.org/10.1016/S0005-7894\(71\)80064-3](http://dx.doi.org/10.1016/S0005-7894(71)80064-3)
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.