Supplement to:

Individual patterns of visual exploration predict the extent of fear generalization in humans

Mario Reutter and Matthias Gamer

Department of Psychology I, Julius-Maximilians-University Würzburg, Germany

Materials, data, and results can be accessed via https://osf.io/4gz7f/

Corresponding author:

Mario Reutter

Department of Psychology I

Julius-Maximilians-University

Marcusstr. 9-11

97070 Würzburg, Germany.

Phone: +49 931 31-84757

e-mail: mario.reutter@uni-wuerzburg.de

Electrodermal Activity

Method

Skin conductance was measured using a constant voltage system (0.5 V) with two 22/10 mm Ag/AgCl electrodes (Berger Medizintechnik GmbH, Gleisdorf, Österreich) filled with 0.5 % NaCl electrolyte gel (PAR Medizintechnik GmbH & Co. KG, Berlin, Germany) that were placed on the thenar and hypothenar eminences of the left hand. Data were collected via the same apparatus as cardiac activity with a sampling rate of 500 Hz.

Further processing was implemented in R 3.5.1. We adhered to the procedure by Ahrens et al. (2016) to obtain comparable results. In order to identify electrodermal non-responders, subjects' responses to the unconditioned stimuli were evaluated as the difference between the maximum and minimum electrodermal activity within the first 5 seconds after administration of electric shocks. Differences less than 0.1 μS were scored as zero. Participants' electrodermal responses were only analyzed further if their average reaction to the painful stimulus exceeded 0.1 μS and more than 50% of responses to the electric stimulation were non-zero. Due to these criteria, electrodermal data of 10 subjects were excluded. Subsequently, skin conductance responses (SCRs) to the photographs were scored as the difference between the maximum and minimum activity within the first 8 seconds after stimulus onset (i.e., the minimum time before presentation of the next stimulus). Differences less than 0.1 μS were scored as zero and trials containing or following a painful stimulus were excluded in order to avoid overlapping of unconditioned responses.

Since substantial overlap between individual skin-conductance responses has to be expected (see Discussion), we also reanalyzed the data using Ledalab 3.4.9 (Benedek & Kaernbach, 2010). This toolbox decomposes the continuous electrodermal activity into tonic and phasic components. Focusing on the phasic electrodermal activity, the average driver was determined within a response window of 0.5 to 3 seconds after stimulus onset (minimum amplitude of 0.1 μS). Again, subjects were only analyzed further if more than 50% of responses

to the electric shock were non-zero, leading to exclusion of 8 individuals for this analysis. Individually determined SCR amplitudes and phasic components of electrodermal responding as calculated by Ledalab were log-transformed to reduce the skew of the amplitude distribution.

Results

SCR amplitudes in the generalization phase were analyzed using a 6×2 within-subjects ANOVA with factors *threat level* and *diagnostic region*. Responses were on average different from zero as indicated by a significant intercept ($F(1, 33) = 78.18$, $p < .001$, $\eta_p^2 = .70$) but they were not predicted by either factor or their interaction (Fs ≤ 1.51, ps $\geq .209$).

Since no differential fear response was observed in skin conductance reactivity, it is not surprising that electrodermal fear generalization profiles were also not predicted by the cumulative dwell time into diagnostic ROIs ($\beta = .11$, $t(62) = 0.63$, $p = .532$) or by the latency of the first fixation ($\beta = -.04$, $t(62) = -0.25$, $p = .806$). Other effects also did not reach significance ($|\beta|$s $\leq .24$, $|t|$s ≤ 1.57, ps $\geq .121$).

Employing the same ANOVA on the deconvoluted data did not change the results. Aside from a significant intercept ($F(1, 34) = 61.22$, $p < .001$, $\eta_p^2 = .64$), no other statistically significant effects were observed (Fs ≤ 1.15, ps $\geq .335$).

Discussion

The currently observed pattern of autonomic responses differs from the results of Ahrens and colleagues (2016). In their study, healthy participants showed a fear generalization gradient in electrodermal activity but virtually no change in heart rate toward any stimulus. Our data, however, indicates differential cardiac fear responses in healthy subjects (see main text; for an overview on fear bradycardia see Roelofs, 2017) while we did not observe this for skin conductance. One reason for this is that we did not optimize the experimental design for the measurement and analysis of skin conductance responses. Since the primary outcome measures

respond quickly to visual stimulation, we chose to keep the inter-trial-interval rather short. This, however, resulted in large overlap of skin conductance responses to the face stimuli, the ratings, and the US. Hence, our data do not permit a valid isolation of electrodermal responses to the onset of the photographs. This interpretation is supported by the pattern of pupil responses in the current study (see main text), that reflected differences in threat value after the acquisition phase and mirrored the generalization gradient that was also observed for heart rate deceleration. Since these pupil responses are driven by sympathetic nervous system activity (Bradley, Miccoli, Escrig, & Lang, 2008) – similar to the skin conductance responses measured in other studies – it seems sensible to assume that the current stimulus timing was not optimal for recordings of electrodermal activity.

Further Methods

Before starting the generalization paradigm, participants completed several preparatory tasks to ensure that they could perceptually distinguish the employed stimuli, to calibrate the intensity of the electrotactile stimulation and to practice the timing of behavioral responses during the task. These tasks are described in detail in the following.

Discrimination Task. Since the stimulus material was perceptually highly similar, the experiment started with a discrimination task that was run with two pairs of stimuli that were to be employed as CS+ and CS− in the subsequent fear generalization paradigm. At the beginning of the task, the four selected pictures were shown in succession for 6 seconds each with their respective response button, separated by a fixation cross for 1 second. After having viewed each stimulus and the according response button once, participants were asked to indicate at the end of consecutive trials via button press which picture they have just seen. After the response, feedback was provided for 2 seconds. Every picture was presented five times, totaling to ten repetitions per pair and 20 trials in total. If a subject achieved less than 80% correct answers for any of the two pairs, the task was repeated and if necessary, the selected pairs were switched until the discrimination criterion was satisfied.

Pain Calibration. After the setup of the electrodes (see Data Recording and Processing), the pain calibration procedure was conducted. Subjects were exposed to two alternatingly ascending and descending trains of electrotactile stimuli. The first impulse started at an intensity of 0.25 mA and was adjusted in steps of 0.25 mA. In succession to each stimulation, participants indicated their sensation on an eleven-point scale ranging from *0 = no sensation* over *4 = minimally painful* to a theoretical maximum of *10 = worst pain imaginable*. The intensity was increased during ascending trains until a 4 or higher was indicated and amperage was decreased again until stimulation was not perceived as painful anymore. For each of the four trains, the minimum intensity that still obtained a painful sensation was identified, the results were averaged across trains and augmented by 50%. If this

intensity did not lead to a moderately painful sensation denoted by a six or higher on the scale, amperage was slightly adjusted once more with consultation of the subjects. After the experiment, an electrotactile stimulus of the same intensity was again applied and participants were asked to report their sensation using the same eleven-point scale as before. The average US intensity that was applied during the experiment amounted to 1.26 ± 0.90 mA and resulted in moderate pain ratings right after calibration (6.4 ± 0.5) and until after the experiment (5.3 ± 0.9). The decline in US painfulness across the experiment was significant ($t(43) = 7.95$, $p < .001$, $d = 1.20$) as well as an increase in variance ($F(43, 43) = 0.33$, $p < .001$).

Timing Training. Having calibrated the aversive electrotactile stimulation, a training phase was conducted with geometric shapes in order to get subjects accustomed to the experimental task that included a trial-by-trial rating of shock expectancies. We used three geometric shapes each with a height of 400 pixels: a green circle, a blue square, and a red equilateral triangle rotated by 180°. In the training phase, we established a contingency between the shapes and the pain stimulation of 0%, 50%, and 100% respectively. The trial structure was identical to the Fear Generalization Task (cp. Figure 1) except that a brief feedback of 1 second with the participant's rating was displayed after the painful stimulation was applied or omitted. If subjects pressed too early or too late, the answer was marked as invalid and an according feedback was provided. Every shape was presented four times in a randomized order. If participants had more than 2 invalid responses out of 12 trials in total, the procedure was explained again and repeated.

Figure S1. Exemplary depiction of a trial with original rating prompt in German ["How likely does a painful stimulus follow?"]. The response scale ranges from *1 = no stimulus* across *3 = unsure* to *5 = stimulus certain*. The corresponding keys on the German keyboard layout are space, L, Ö, Ä, and right shift respectively. The figure was adapted using photographs from the Oslo Face Database (https://sirileknes.com/oslo-face-database/). Source: Chelnokova et al., 2014.

Physiological Responses during Habituation & Acquisition

Pupillary Responses

While pupil responses did not differ between faces during habituation ($t(42) = 0.12$, $p = .904$, $d = 0.02$, 95% CI [−0.28, 0.32]), a sizeable effect emerged during acquisition with pupil dilation being significantly greater for CS+ compared to CS− ($t(42) = 5.29$, $p < .001$, $d = 0.81$, 95% CI [0.46, 1.15]). This effect decreased in magnitude during the generalization phase but remained statistically significant ($t(42) = 3.13$, $p = .002$, $d = 0.48$, 95% CI [0.16, 0.79]).

Heart Rate

During habituation, there was no significant difference between phasic heart rate changes in response to faces that later became a CS+ or CS− ($t(42) = -1.47$, $p = .148$, $d = -0.22$, 95% CI [−0.53, 0.08]). Averaging across the trials of the acquisition phase, there was still no difference between CSs ($t(42) = -1.44$, $p = .157$, $d = -0.22$, 95% CI [−0.52, 0.08]). Descriptively, the effect during acquisition even pointed into the opposite direction than expected, i.e., greater deceleration for CS− ($p = .921$, one-tailed). Only during the generalization phase, a significant difference between CS+ and CS− was detected ($t(42) = 2.78$, $p = .008$, $d = 0.42$, 95% CI [0.11, 0.73]).

Discussion

Pupillary responses and heart rate changes showed different temporal dynamics across the experiment. While modulations in pupil size emerged quickly during acquisition and already declined again in the generalization phase, threat-specific cardiac deceleration built up more slowly such that it could only be reliably detected in the last stage of the experiment. It is unclear, however, whether the weaking of differential pupillary responses was due to habituation effects or to the drop in US-contingency from 75% (acquisition) to 50% (generalization). Additionally, future research should elucidate if threat-contingent heart rate decelerations are not only slower to acquire but also more resistant to extinction.

Extended models (generalization phase)

Shock Expectancy Ratings

In order to further analyze the ratings of the generalization phase, a 6×2×2×2×2 mixed-effects ANOVA was conducted with the two within-subject factors *threat level* (CS−, GS1 through GS4, CS+) and *diagnostic region* (eyes vs. mouth/nose) as well as three additional between-subjects factors *pairs* (male eyes & female mouth/nose vs. female eyes & male mouth/nose), *cs-male* (face 1 vs. 2 of the male pair was used as CS+) *cs-female* (face 1vs. 2 of the female pair was used as CS+). Along with the already discussed main effect of *threat level* ($F(5, 200) = 123.89$, GG-ε = .55, $p < .001$, $\eta_p^2 = .78$, 95% CI [.70, .83]) and the marginally significant interaction of *threat level* × *diagnostic region* ($F(5, 200) = 2.36$, GG-ε = .52, $p = .080$, $\eta_p^2 = .06$, 95% CI [.00, .16]), a significant interaction of *threat level* × *cs-female* emerged ($F(5, 200) = 4.19$, GG-ε = .55, $p = .010$, $\eta_p^2 = .11$, 95% CI [.01, .22]) that was further classified by a four-way interaction *threat level* × *cs-female* × *pairs* × *diagnostic region* ($F(5, 200) = 3.71$, GG-ε = .52, $p = .016$, $\eta_p^2 = .10$, 95% CI [.00, .20]). As can be seen in Figure S2, CS-discrimination was increased and fear generalization reduced, when the second female stimulus (of either female pair) was used as CS+. This suggests a mismatch between the perceived inherent threat level of the stimuli and their assigned role in the fear acquisition paradigm. Checking the ratings of the original stimuli, it becomes apparent that there may be differences in perceived dominance between the female pairs (differences scores for eyes male: 0.24, mouth/nose male: 0.03, eyes female: 0.62, mouth/nose female: 1.36). However, assignment of stimuli to the CS+ and CS− was counterbalanced. Thus, 50% of participants experienced the more dominant stimulus as safety signal, balancing out the aforementioned effect.

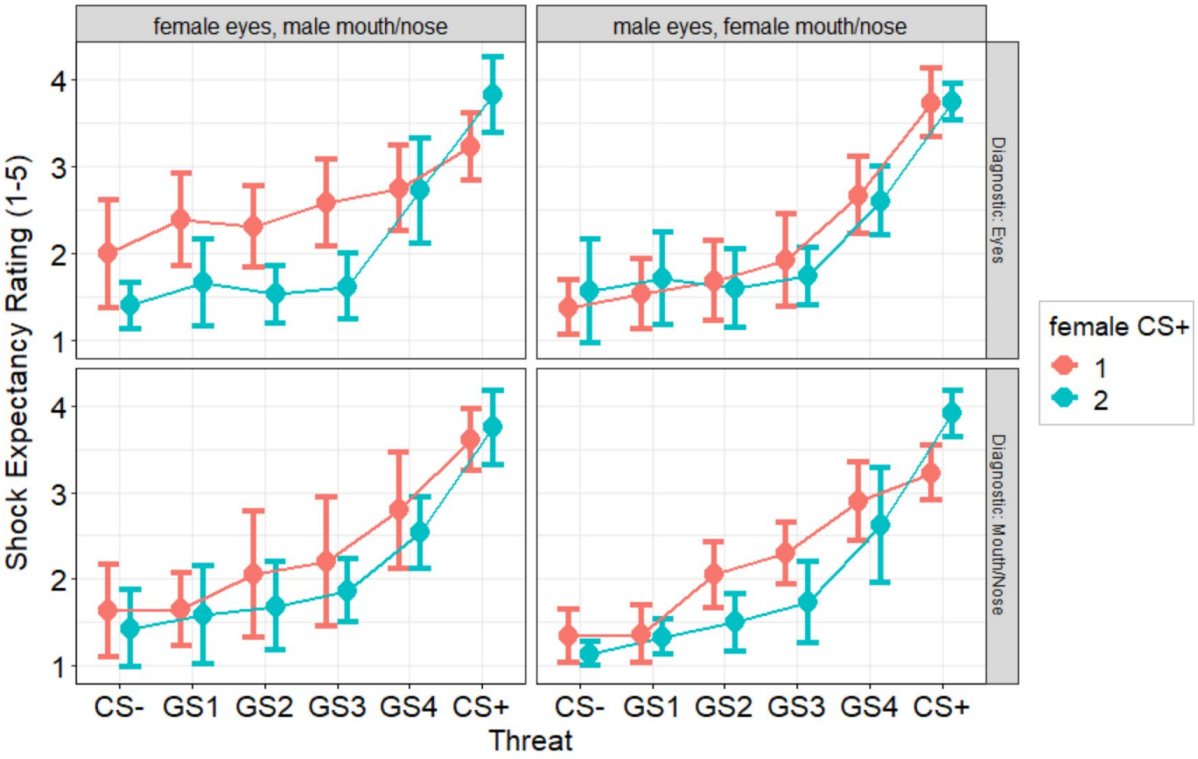

Figure S2. Illustration of the effect of stimulus assignment. Differences in the fear gradient emerged for female pairs with diagnostic eyes (top left) and diagnostic mouth and nose (bottom right). Error bars indicate 95% confidence intervals of between-subjects estimates.

Eye-Tracking Data

To explore potential temporal dynamics within the fixation behavior, we employed a 2×2×6×8 within-subjects ANOVA with factors *ROI* (eyes vs. mouth/nose), *diagnosticity* (diagnostic vs. non-diagnostic), *threat level* (CS−, GS1 through GS4, CS+) and *trial time* (0 to 4 sec in bins of 0.5 sec). The effects that are not depicted in the main text are illustrated here.

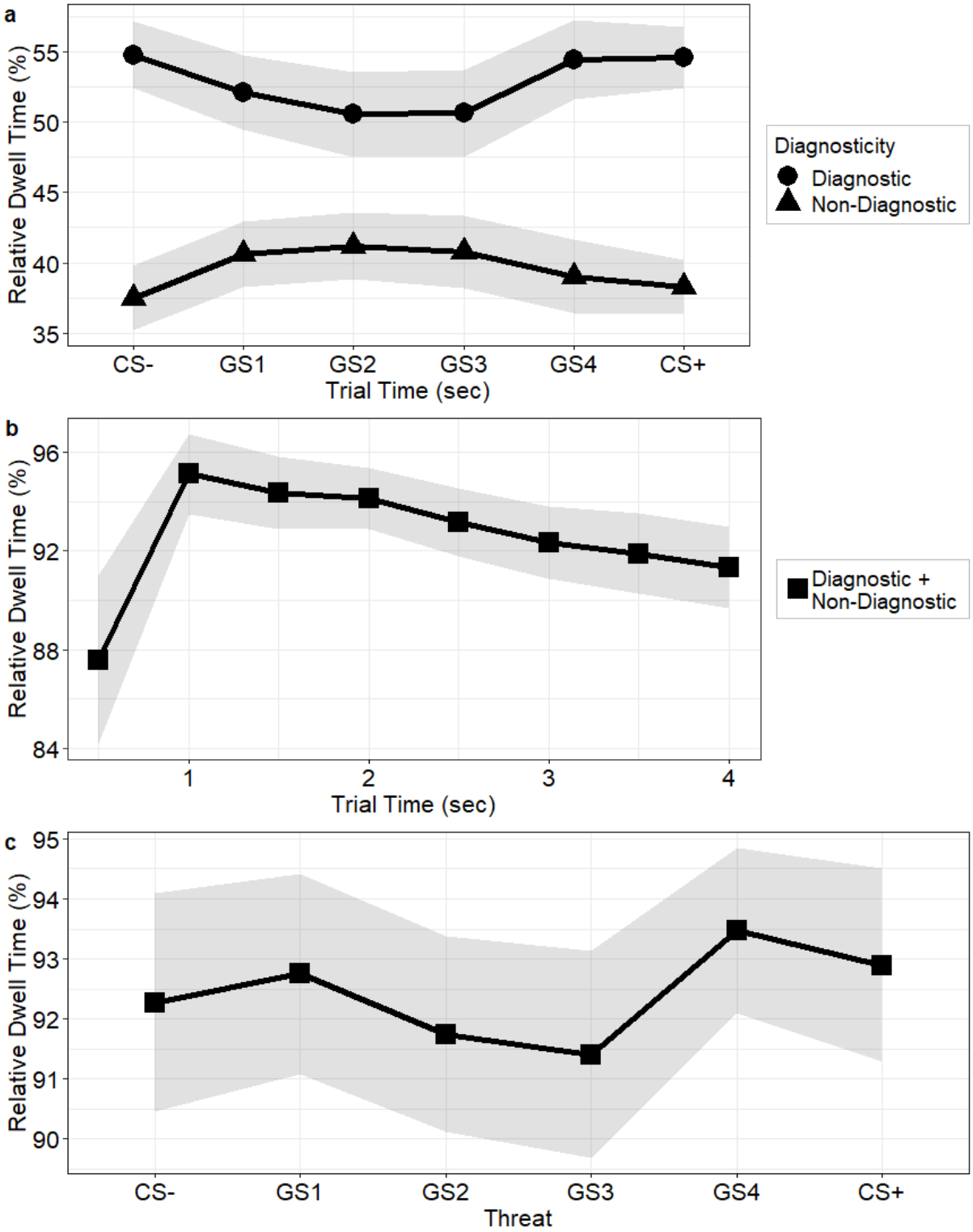

Figure S3. Visualization of ANOVA effects: a) *diagnosticity* × *threat level*, b) main effect of *trial time*, and c) main effect of *threat*. Confidence bands indicate 95% confidence intervals of between-subjects estimates.

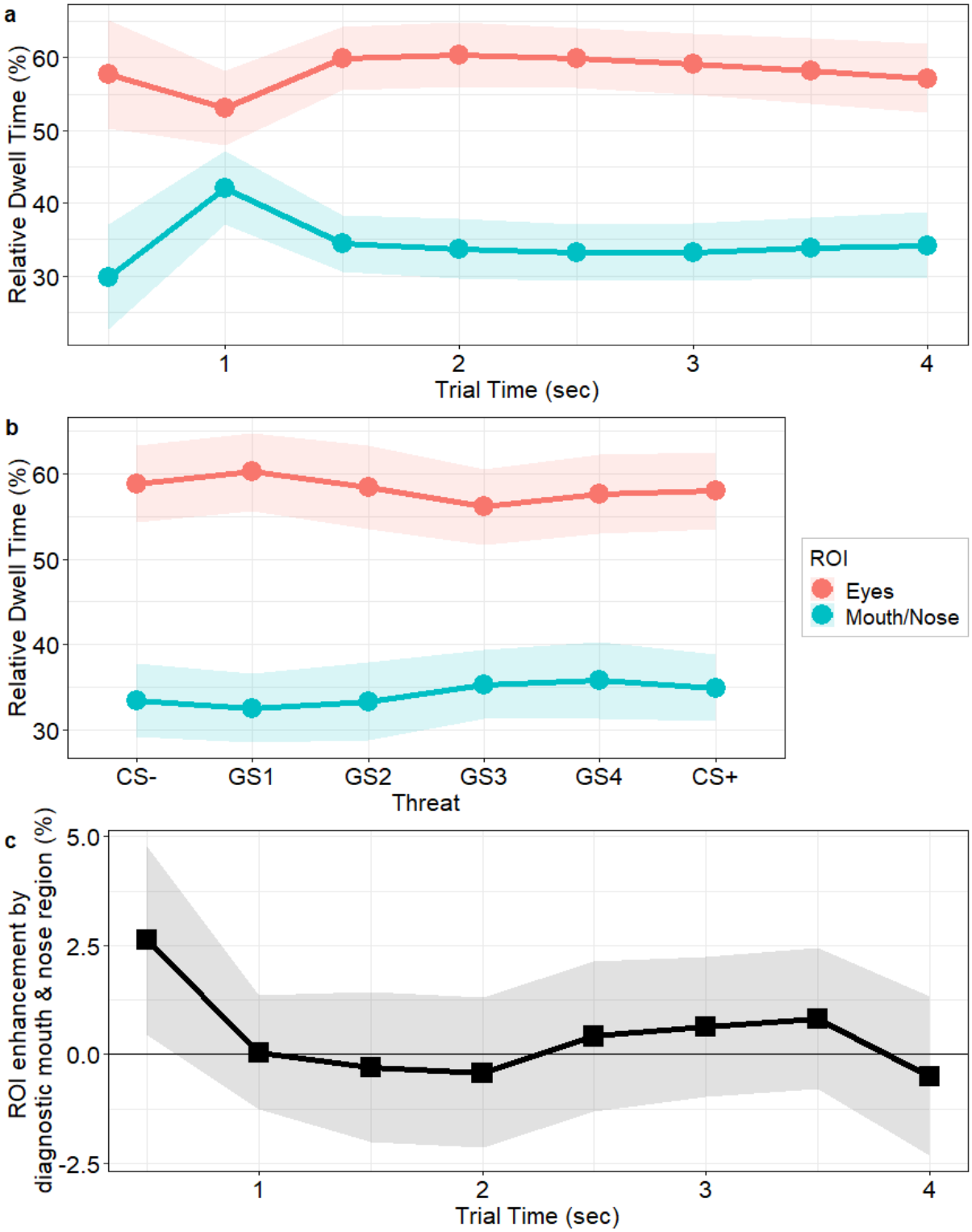

Figure S4. Visualization of ANOVA effects: a) *ROI* × *trial time*, b) *ROI* × *threat level*, and c) the marginal *diagnosticity* × *ROI* × *trial time* simplified as the temporal progression of the difference score between the relative dwell times on any ROI for stimuli with diagnostic mouth & nose area minus diagnostic eyes (cp. *diagnosticity* × *ROI* in confirmatory analysis). Confidence bands indicate 95% confidence intervals of between-subjects estimates.

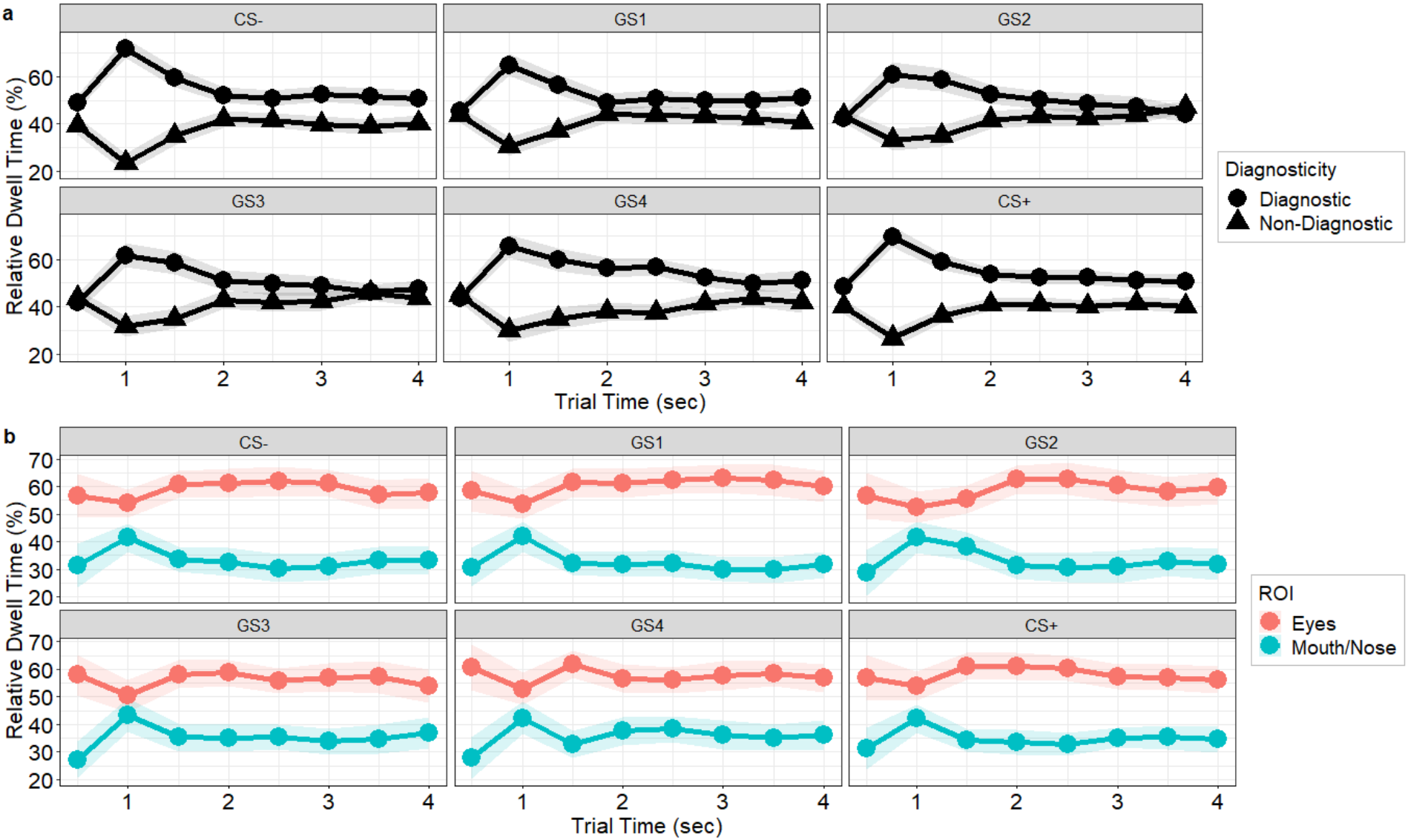

Figure S5. Visualization of ANOVA effects: a) *diagnosticity* × *trial time* × *threat level* and b) *ROI* × *trial time* × *threat level*. Confidence bands indicate 95% confidence intervals of between-subjects estimates.

Influence of Visual Exploration on Fear Generalization

The confirmatory linear mixed model analysis indicated that fear generalization as reflected in shock expectancy ratings was predicted by the cumulative dwell time into diagnostic regions of interest (β = .30, 95% CI [.07, .50], $t(50.27) = 2.57$, $p = .007$ one-tailed; cf. Figure 6). Even though diagnostic dwell was strongly modulated across trial time (at least for diagnostic mouth/nose; cf. Figure 5b), there was no evidence for temporal dynamics of this relationship ($|\beta|s \leq .15$, $|t|s \leq 1.32$, $ps \geq .186$). Descriptively, beta estimates ranged from .171 (at 3 seconds) to.344 (at 1 second). The parameter estimates across trial time are depicted in Figure S6.

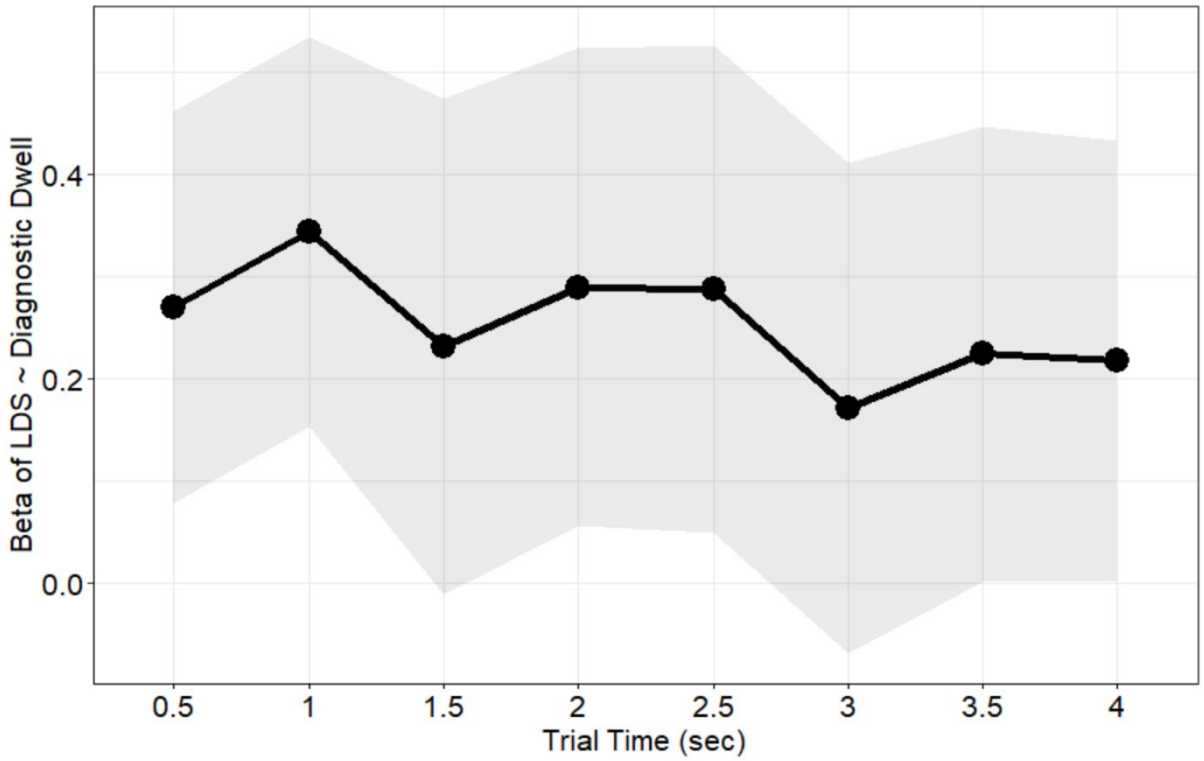

Figure S6. Estimated beta weights of the association between dwell time into diagnostic regions and fear generalization (as indexed by linear deviation scores of shock expectancy ratings) across trial time. The temporal modulation of this relationship was not significant. Confidence bands indicate 95% confidence intervals of parameter estimates.

The explorative linear mixed model analysis on the shock expectancy gradient using the square root of latencies of first fixations into the diagnostic region yielded highly significant predictive value (β = −.35, 95% CI [−.53, −.15], $t(59.11) = -3.47$, $p < .001$ one-tailed). Without this normalization, the effect was almost identical (β = −.34, 95% CI [−.51, −.14],

$t(62.83) = -3.29$, $p < .001$ one-tailed). Other predictors did not reach statistical significance ($|\beta|$s $\leq .07$, $|t|$s ≤ 0.65, ps $\geq .519$). The respective scatter plot can be seen in Figure S7.

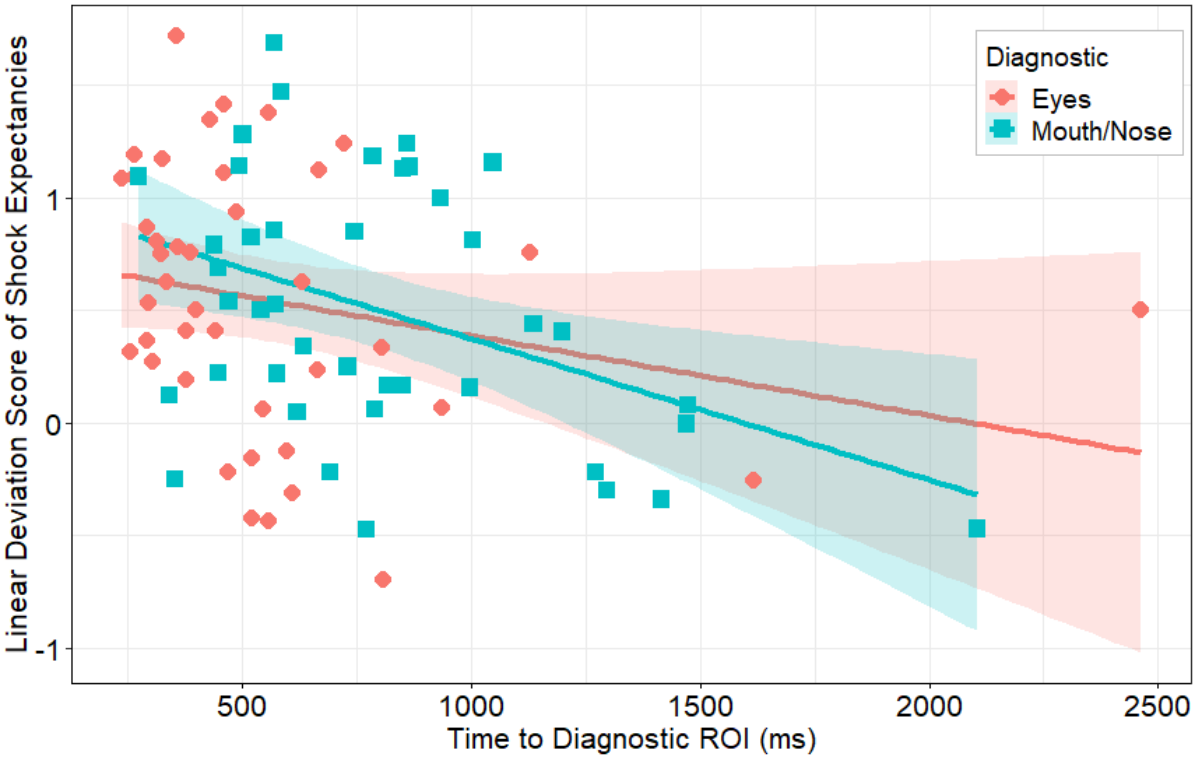

Figure S7. Scatter plot for linear deviation scores of shock expectancy ratings during generalization as a function of the latency until the first fixation on the diagnostic ROI. Confidence bands indicate 95% confidence intervals of regression line estimates.

References

Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T., & Wieser, M. J. (2016). Fear conditioning and stimulus generalization in patients with social anxiety disorder. *Journal of Anxiety Disorders*, *44*, 36–46. https://doi.org/10.1016/j.janxdis.2016.10.003

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. https://doi.org/10.1111/j.1469-8986.2008.00654.x

Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Løseth, G., Maurud, H., . . . Leknes, S. (2014). Rewards of beauty: The opioid system mediates social motivation in humans. *Molecular Psychiatry*, *19*(7), 746–747. https://doi.org/10.1038/mp.2014.1

Roelofs, K. (2017). Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1718). https://doi.org/10.1098/rstb.2016.0206