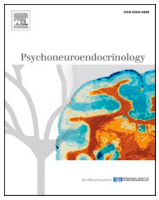
Psychoneuroendocrinology 187 (2026) 107793

Contents lists available at ScienceDirect

Psychoneuroendocrinology

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



The effects of pre-extinction stress vs. physical exercise on contextual retrieval and generalization

Lianne N. Wolsink *, Oliver T. Wolf, Christian J. Merz, Valerie L. Jentsch

Department of Cognitive Psychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, Universitätsstraße 150, Bochum 44780, Germany

ARTICLE INFO

Keywords:
Aerobic exercise
Context
Extinction learning
Fear renewal
Stress hormones

ABSTRACT

Exposure to stress hormones before extinction learning leads to strong and less context-dependent extinction memories. However, laboratory stress induction protocols have disadvantages, such as increases in negative affect. In this study, we investigated physical exercise as another stress hormone associated, but healthy and more positively experienced modulator of fear extinction memories. We compared the effects of a vigorous-intensity exercise intervention (treadmill running) and psychosocial stress before extinction training on later contextual retrieval and generalization. To this end, 120 (60 women, 60 men) healthy participants underwent fear acquisition training in context A on day 1, the exercise, stress, or control intervention followed by extinction training in context B on day 2, and a retrieval and reinstatement test in context A, B, and C on day 3. The stress and exercise intervention both significantly increased heart rate, salivary cortisol and alpha amylase levels compared to the control intervention. For both skin conductance responses and pupil dilation, fear acquisition was successful and fear responding decreased over extinction trials. The groups did not differ in fear responding during retrieval 24 h later, but psychosocial stress seemed to reduce fear renewal in a novel context. Moreover, exploratory analyses revealed that stress reduced the context-dependency of the extinction memory after reinstatement in men, whereas exercise reduced overall responding in the extinction context. These findings shed light on potential boundary conditions of stress and exercise effects on contextual retrieval of extinction memories, including the generalization to novel contexts with high therapeutic value.

1. Introduction

Fear extinction learning represents the key underlying mechanism of exposure therapy for anxiety disorders (Craske et al., 2018). During extinction learning, a new memory is formed that inhibits the original fear memory (Bouton, 2004). To prevent a fear response upon later confrontation with a reminder (i.e., conditioned stimulus (CS)), the extinction memory needs to be stronger than the fear memory. In addition to memory strength, extinction memories need to generalize to other contexts to prevent the return of fear in the original (ABA) or a novel context (ABC; i.e., fear renewal). However, whereas fear memories often generalize, extinction memories are usually context-specific (Bouton, 2004; Vervliet et al., 2013). Several lines of research therefore focus on improving extinction memories (Fitzgerald et al., 2014).

One approach to improve the strength and generalization of extinction memories is exposure to stress hormones (de Quervain et al., 2019; Merz and Wolf, 2022). Acute stress before extinction training improved extinction memory retrieval 24 h later in men only (Bentz et al., 2013) and prevented renewal of fear (in men; Meir Drexler et al., 2018) and neutral memories (in men and women; Meir Drexler et al., 2017). Based on these findings, the Stress Timing affects Relapse (STaR) model (Meir Drexler et al., 2019) states that exposure to acute stress or administration of cortisol before extinction training creates a strong and less context-dependent extinction memory, such that a change in context does not result in return of fear.

Typical procedures to increase stress hormone levels have some disadvantages, however. Laboratory stress induction protocols, such as the Trier Social Stress Test (TSST; Kirschbaum et al., 1993), usually induce negative emotions (Allen et al., 2014), thereby questioning the emotional and ethical tolerability for use in clinical practice. Moreover, pharmacologically elevated cortisol combined with exposure therapy can improve treatment outcomes (de Quervain et al., 2017), but anxiety patients tend to prefer psychological over pharmacological approaches (McHugh et al., 2013). In search of a more positively appraised

* Corresponding author.
E-mail address: lianne.wolsink@ruhr-uni-bochum.de (L.N. Wolsink).

https://doi.org/10.1016/j.psyneuen.2026.107793
Received 22 September 2025; Received in revised form 3 February 2026; Accepted 10 February 2026
Available online 12 February 2026

behavioral method to enhance extinction memories, physical exercise could be suitable. Similar to stress induction via the TSST, a bout of physical exercise can activate the two major stress response systems: the sympathetic nervous system (SNS) and the hypothalamus-pituitary-adrenocortical (HPA) axis (Ponce et al., 2019), leading to the release of (nor)adrenaline and glucocorticoids (GCs; mainly cortisol in humans; Joëls and Baram, 2009), respectively. Yet, exercise is healthy and often associated with increases in positive and reductions in negative emotions (Basso and Suzuki, 2017).

Acute physical exercise can also modulate (emotional) memory consolidation (e.g., Jentsch and Wolf, 2020; Keyan and Bryant, 2019a, 2019b; Loprinzi et al., 2019) and improve extinction memory recall in rodents (e.g., Keyan and Bryant, 2019a; Tanner et al., 2018; but see Jacquart et al., 2017). Initial findings on extinction retrieval in humans are promising (Crombie et al., 2021a, 2023; Jentsch et al., 2023; Keyan and Bryant, 2019b), but often difficult to compare due to variability in study design and sample (Azar et al., 2024). Potential moderating factors are exercise characteristics such as duration or intensity, and the outcome measure of fear. For instance, a bout of moderate-intensity exercise reduced threat expectancy during retrieval in healthy women (Crombie et al., 2023) and women with posttraumatic stress disorder (PTSD; Crombie et al., 2021a) but had no effect on physiological measures. In contrast, another study found that high-intensity exercise reduced fear potentiated startle responses during retrieval in men and women (Keyan and Bryant, 2019b). Furthermore, exercise timing relative to extinction training could play a crucial role. Whereas most human studies applied exercise after extinction training (Crombie et al., 2021a, 2023; Jentsch et al., 2023; Keyan and Bryant, 2019b), the effects of pre-extinction (or in clinical settings: pre-exposure) exercise have only been studied for exposure therapy, leading to mixed results. Moderate-intensity exercise before exposure reduced PTSD symptoms in a small pilot study involving mainly women (Powers et al., 2015), while vigorous-intensity exercise before exposure did not affect symptom reduction in men and women with fear of heights (Jacquart et al., 2017). In another study, low-to-moderate-intensity exercise before exposure was less effective than after exposure in men and women with PTSD (Voorendonk et al., 2021).

Taken together, physical exercise seems a promising intervention for improving extinction memories, but heterogeneity in study design and results also show that moderating factors play a role. Moreover, the effect of physical exercise on generalization of the extinction memory to novel contexts (i.e., ABC renewal), which has high therapeutic value, has not been studied before. Even though post-extinction exercise was generally found to be more beneficial for item memory, pre-extinction exercise may have larger effects on context memory, which would be in line with the STaR model. Additionally, the underlying mechanisms are not well understood. Besides the contributing role of brain-derived neurotrophic factor (Crombie et al., 2021b; Keyan and Bryant, 2019b), other mechanisms could also be involved (Basso and Suzuki, 2017), such as stress hormones. In the current study, we tested whether a single bout of vigorous-intensity physical exercise compared to psychosocial stress before extinction training could enhance extinction memory retrieval and its generalization to a novel context 24 h later. We used a contextual fear conditioning paradigm: 120 healthy men and women underwent fear acquisition training in context A on day 1, the exercise, stress, or a control intervention followed by extinction training in context B on day 2, and a retrieval and reinstatement test in context A, B, and a novel context C on day 3. As the STaR model (Meir Drexler et al., 2019) has only been tested using physiological stressors or cortisol administration, we used the TSST to test whether the STaR model also applies to psychosocial stress alongside the effects of physical exercise. Moreover, as sex hormones can influence fear and extinction memory processes (Merz, Kinner, et al., 2018), we tested both men and women to examine sex differences in an exploratory manner.

In line with the STaR model, we predicted that pre-extinction stress and vigorous-intensity exercise both facilitate extinction memory consolidation in a context-independent way. More specifically, we expected reduced fear retrieval in context A, enhanced extinction retrieval in context B, that also generalizes to context C in the stress and exercise group as compared to the control group as indicated by reduced differential (CS+ minus CS-) skin conductance responses (SCRs) and pupillary responses in all three contexts. We further expected successful extinction retrieval to be related to increased stress-/exercise-induced heart rate, alpha amylase, and cortisol responses.

2. Methods

We preregistered this study on OSF (https://doi.org/10.17605/OSF.IO/45R7Y).

2.1. Participants

An a priori power analysis revealed a required sample size of 40 participants for each group (see supplement S1.1). We tested 120 healthy participants (60 women; age: M=23.60, SD=4.01 years, body mass index: M=23.03, SD=2.33 kg/m^2; S2.1), mainly recruited at Ruhr University Bochum (Germany). All women were naturally cycling and not tested during menses, pregnancy, or breast-feeding. Further participant inclusion criteria are reported in the supplement (S1.2). Participants provided written informed consent before participation and received €50 or course credits for participation. All procedures were approved by the ethics committee of the Faculty of Psychology at Ruhr University Bochum (reg. nr. 656).

Table 1 shows the sample characteristics per group per sex. The groups did not significantly differ in age, BMI, or self-reported physical activity in daily life as measured by the LTEQ (Godin, 2011). Men and women significantly differed in age and BMI, but this did not interact with the factor group. Statistics are reported in the supplement (S2.1).

2.2. Task and procedure

Participants were tested on three consecutive days, with the different experimental phases (Fig. 1) starting around (+/- 2 h) the same time between 12:00 pm and 6:00 pm.

The fear conditioning paradigm was presented in MATLAB (version 2018a, The MathWorks Inc., 2018) using the Psychophysics toolbox (Brainard and Vision, 1997) and the OTBR Toolbox (Rose et al., 2008). Stimuli and procedure were adopted from Milad et al. (2007) as realized before (e.g., Hermann et al., 2016). One picture of an office room, of a room with a shelf, and of a conference room served as contexts A, B, and C. In all contexts, the same desk lamp lit up in either yellow or blue serving as the CS (assignment of colors to CS+ and CS- counterbalanced across participants). All pictures and the fixation cross presented during the inter-trial interval (ITI) were of equal size (1440 ×1080 pixel) and luminance matched to a brightness of 125 cd/m^2 to ensure proper pupillometric assessment. An electrical stimulation served as unconditioned stimulus (UCS), individually adjusted as being “unpleasant but not painful”. The UCS was applied using a constant voltage stimulator (STM200; BIOPAC Systems, Goleta, CA) with two electrodes (surface size: 1 cm^2) attached to the middle of the left calf.

Before each experimental phase started, participants were instructed to carefully watch the stimulus presentation to observe any possible regularity between lamplight colors and electrical stimulation. They were informed that if they discover such a regularity (e.g., yellow is associated with an electrical stimulation), it will not change during the experimental phases (thus, e.g., blue will never be associated with a stimulation). However, they were not explicitly informed about the actual contingencies or any change in contingencies (e.g., that no stimulations are given during the extinction phase) over the course of the experiment.

All trials started with the presentation of the context without a CS (turned-off lamp) for 3 s, followed by 6 s of CS presentation (lamp

Table 1
Sample characteristics per group and per sex.

	Stress			Exercise			Control		
	Overall ($N = 40$)	Men ($N = 20$)	Women ($N = 20$)	Overall ($N = 40$)	Men ($N = 20$)	Women ($N = 20$)	Overall ($N = 40$)	Men ($N = 20$)	Women ($N = 20$)
Age	23.55 ± 4.39	25.50 ± 4.94	21.60 ± 2.66	23.68 ± 4.20	24.40 ± 5.02	22.95 ± 3.14	23.58 ± 3.48	23.70 ± 3.25	23.45 ± 3.78
BMI	23.06 ± 2.10	23.90 ± 2.02	22.26 ± 1.89	23.26 ± 2.42	24.05 ± 2.64	22.47 ± 1.94	22.77 ± 2.50	23.50 ± 2.47	22.00 ± 2.34
PA	699.25 ± 725.25	608.25 ± 777.57	790.25 ± 676.47	449.00 ± 420.15	458.25 ± 339.08	439.75 ± 497.17	603.75 ± 588.206	633.50 ± 661.49	574.00 ± 520.34

Note. Values are reported as M±SD; BMI: Body mass index in kg/m^2; PA: Composite score of self-reported physical activity in daily life: frequency during the last 4 weeks (in days) × duration (in min).

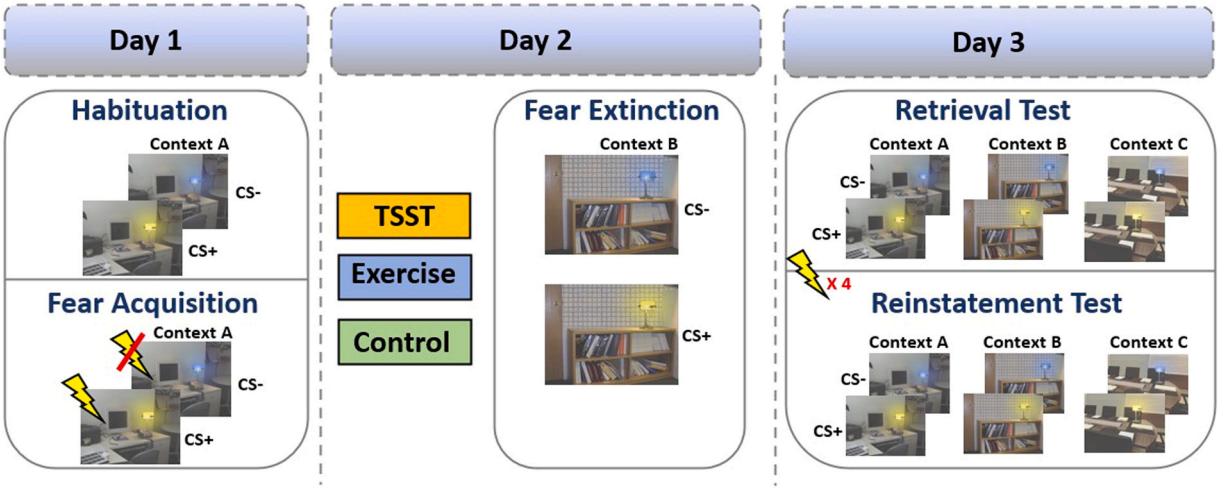


Fig. 1. Overview of experimental procedure, Note. TSST: Trier Social Stress Test.

within the context picture lighting up in yellow or blue). The UCS (100 ms) was delivered immediately after the offset of the CS for reinforced CS+ trials during fear acquisition training only. An intertrial-interval depicting a black cross hair on a grey background was shown from CS offset until the start of the next context presentation for 11 s.

Day 1 started with habituation, including the presentation of two CS+ and CS- each without any UCS in context A (Fig. 1). During fear acquisition training, both CS were presented eight times each in context A. Whereas the CS+ was paired with the UCS in five out of eight trials (62.5 % partial reinforcement rate), the CS- was never coupled with the UCS. After fear acquisition training, we checked contingency awareness using the criteria in the supplement (S1.3). Day 2 started with the stress, exercise, or control intervention. Thereafter, participants performed extinction training, during which both CS were shown without any UCS for eight times each in context B. On day 3, during the retrieval test, the CS+ and CS- were presented four times each in context A, B, and the new context C without any electrical stimulation. Then, four un-signaled UCS separated by 5 s were given in a neutral context (i.e., black background for 20 s) as reinstatement of the acquired fear on day 1. During the subsequent reinstatement test, both CS+ and CS- were presented four times in context A, B, and C, again without electrical stimulation. The retrieval and reinstatement test were arranged in four blocks each including one presentation of the CS+ and the CS- in context A, B, and C. Each combination of CS and context could occur in every block, reflecting a counterbalanced design.

Pseudo-randomized stimulus orders were used for all experimental phases with no more than two presentations of the same CS in a row and a counterbalanced beginning of each experimental phase with either CS+ or CS- (for a detailed description: S1.4). Stimulus presentation and allocation to the CS were individualized and matched between the three groups.

2.3. Skin conductance responses (SCRs)

SCRs were measured with Ag/AgCl electrodes filled with an isotonic (0.05NaCl) electrolyte medium attached to the hypothenar of the participant's left hand. We used a commercial SCR coupler and amplifying system (MP150 +GSR100C, BIOPAC Systems, Inc.; software: AcqKnowledge 5.0) and a sampling rate of 100 Hz. Trough-to-peak maximum amplitudes (in µS) were calculated in multiple time windows using the EDA-Analysis App (Otto et al., 2023): 1–4 s after context onset as contextual response, 1–6.5 s after CS onset as conditioned response, 0.5–4 s after UCS onset as unconditioned response. The time window for the conditioned response includes the typical time between stimulus and SCR peak (Dawson et al., 2017) and the time window generally used in fear conditioning research (Lonsdorf et al., 2017).

2.4. Pupillary responses

Pupillary responses were recorded with a high-accuracy head-stabilized EyeLink Portable Duo eye tracker (SR Research Ltd., Ottawa, Canada) mounted on a tripod, placed below the screen, and connected to a Host PC (ThinkPad T470 W10DG, Lenovo Notebook). A high-speed USB camera including a near-infrared illuminator for dark pupil detection measured retinal and corneal reflections to obtain participants' pupil diameter (in arbitrary units (AU)) of both eyes with a sampling rate of 250 Hz. A 12-point calibration was carried out to ensure correct tracking of the participant's pupils. To minimize head movements, participants placed their chin and forehead on a headrest. The testing room was moderately lit without daylight luminance to control for variance in light influences.

Preprocessing of the raw pupil size time series was performed using

MATLAB (Version 2024b, The MathWorks Inc., 2024) based on routines reported in Jentsch et al. (2023); S1.5). For each participant and each trial, baseline pupil size was defined as the average pupil diameter recorded during the last 300 ms of context presentation and subtracted from the pupil size during CS presentation to account for random fluctuations in pupil size over time (Mathôt et al., 2018). Mean pupil size for CS+ and CS- trials were determined within a time window from 0 to 6 s after CS onset in bins of 2 s (0–2 s, 2–4 s, and 4–6 s) with a primary focus on the last time bin (i.e., 4–6 s, time bin directly preceding the UCS; Finke et al., 2021) as conditioned response.

2.5. The stress, exercise and control interventions

Participants were randomly assigned to one of three groups: stress ($N = 40$; 20 women), exercise ($N = 40$; 20 women) or control ($N = 40$; 20 women). The interventions took place in a different room than the room in which the fear conditioning procedure was performed. In the stress intervention, participants completed a modified version of the TSST with a 20 min duration (S1.6).

In the vigorous-intensity exercise intervention, participants completed a 20 min running task on a treadmill with a 3 % slope. The intervention consisted of a 4 min warm-up during which speed was stepwise increased until an individually set target heart rate (HR) was reached, a 15 min vigorous-intensity training at target HR, and a 1 min cool-down. Similar to previous studies from our lab (Jentsch and Wolf, 2020; Jentsch et al., 2023), the target HR was defined as 85 % of the participant's HR reserve (S1.7). Speed was set for each participant individually and was adjusted throughout the intervention to ensure that participants maintained their target heart rate (i.e., reduced when HR increased above the target HR, or increased again when HR dropped). At four time points during the exercise intervention, participants' exertion levels were monitored using the Borg rating of perceived exertion (Borg, 1998; from no exertion (6) to maximum effort (20); ratings after 5 min: $M = 12.50$, $SD = 3.00$; after 10 min: $M = 15.21$, $SD = 2.01$; after 15 min: $M = 16.19$, $SD = 2.03$; after 20 min: $M = 15.68$, $SD = 2.46$).

In the control intervention, participants were requested to sit for 20 min while watching two emotionally neutral and not arousing documentaries (on a 9-point Likert scale ranging from calming to arousing: $M=2.18$, $SD=1.54$; and from negative to positive: $M=6.96$, $SD=1.59$).

2.6. Cardiovascular and neuroendocrine measures

2.6.1. Heart rate

HR responses were recorded on day 2 using a Polar watch device (Polar V800, Polar® Electro, Finland) at the beginning of the testing day (6 min), during (20 min) and after (6 min) the respective intervention. HR data was processed in Kubios HRV Standard (Version 3.5.0; Tarvainen et al., 2014). Recordings were shortened to 5 min (baseline and post) and to 15 min (intervention), detrended (smoothn prior: $\lambda=500$), and abnormal or biologically implausible beats were automatically detected and corrected using interpolation. Mean HR in beats per minute (bpm) was calculated for each time interval.

2.6.2. Salivary cortisol and alpha amylase

A total of nine saliva samples were collected using Salivette sampling devices (Sarstedt, Nümbrecht, Germany) to assess salivary alpha amylase (sAA) and free cortisol concentrations. On day 1 and 3, samples were taken at the beginning and end of the testing day. On day 2, they were taken at the beginning of the testing day (baseline), −1 min before the stress/exercise/control intervention and + 1, + 10 (before extinction training), and + 20 min (after extinction training) after its offset. Analysis procedures are reported in the supplement (S1.8). Concurrently with each saliva sample, self-reported affect was assessed by the Differential Affective Scale (DAS; S1.9 and S2.4).

2.7. Data exclusions

Participants with missing data were excluded only from the respective analysis. The final sample for cortisol and sAA analysis on day 2 was $N = 118$ ($N = 120$ on day 1 and 3), for self-reported affect $N = 118$, and for HR $N = 114$. For SCRs and pupillometry, the final sample contained data from $N = 114$ for fear acquisition, $N = 114$ and $N = 110$ for extinction, $N = 116$ and $N = 93$ for retrieval, and $N = 113$ and $N = 91$ for reinstatement test, respectively (exclusion criteria: S1.5 and S1.10).

2.8. Statistical analyses

Statistical analyses were performed in R (version 4.4.2, R Team, 2024) using RStudio (version 2024.09.1 +394, R Studio Team, 2024). The significance level was set to $\alpha = .05$. In case of non-normally distributed data, data was log-transformed, which was the case for salivary cortisol, sAA, and skin conductance data. For ANOVA, we report Greenhouse-Geisser corrected p-values if assumptions of sphericity were violated. Significant ANOVA results were followed by planned Holm-adjusted post-hoc tests. For salivary cortisol, sAA and affect ratings, we ran separate mixed ANOVA with the within-subjects factor time (day 1: pre- and post-acquisition; day 2: baseline, pre-intervention, + 1, + 10 and + 20 min after stress/exercise/control offset; day3: pre- and post-retrieval/reinstatement test) and the between-subjects factors group (stress vs. exercise vs. control) and sex (women vs. men). To confirm a rise in HR, we ran a mixed ANOVA with the within-subjects factor time (baseline, during stress/exercise/control, post) and the between-subjects factors group and sex.

We conducted mixed ANOVA for SCRs and pupil responses, separately, with the within-subjects factors CS (CS+ vs. CS-), time (per stimulus: habituation: two trials; fear acquisition and extinction training: eight trials each; retrieval test and reinstatement test: four trials each) and context (context A vs. B vs. C during the retrieval and reinstatement test only) and the between-subjects factors group and sex. ANOVA were conducted separately for each experimental phase with the relevant factors included. Additional mixed ANOVA with the within-subjects factors CS and context and the between-subjects factors group and sex were conducted on the first trial of the retrieval test and the first trial of the reinstatement test only.

To test whether changes in cortisol, sAA, and HR were related to extinction memory retrieval, we conducted Spearman correlations for the entire sample and for the groups separately. To do so, we calculated the delta between peak and baseline cortisol and sAA concentrations. The peak was individually defined as the highest value of time point + 1 min, + 10 min, or + 20 min. The delta between peak and baseline was also calculated for HR. Extinction memory retrieval was calculated as CS+ /CS- differences in SCRs and pupil responses, separately, during the retrieval and reinstatement test (separately for context A, B, and C; for all trials and the first trial only). We used Holm's method to correct for multiple testing.

3. Results

The results section was written using the R package Papaja (Aust and Barth, 2024) and reproducible using the R markdown file and data (https://doi.org/10.17605/OSF.IO/792HB). A sample description is reported in the supplement (S2.1).

3.1. Cardiovascular and neuroendocrine measures

3.1.1. Heart rate

The stress and exercise intervention both significantly increased HR compared to the control intervention (Fig. 2a; group×time interaction: $F(3.41, 183.95) = 841.19$, $p = <.001$, partial $\eta^2 = 0.94$; main effect group: $F(2, 108) = 185.93$, $p = <.001$, partial $\eta^2 = 0.78$; main effect time: $F(1.70, 183.95) = 1365.15$, $p = <.001$, partial $\eta^2 = 0.93$). While HR did

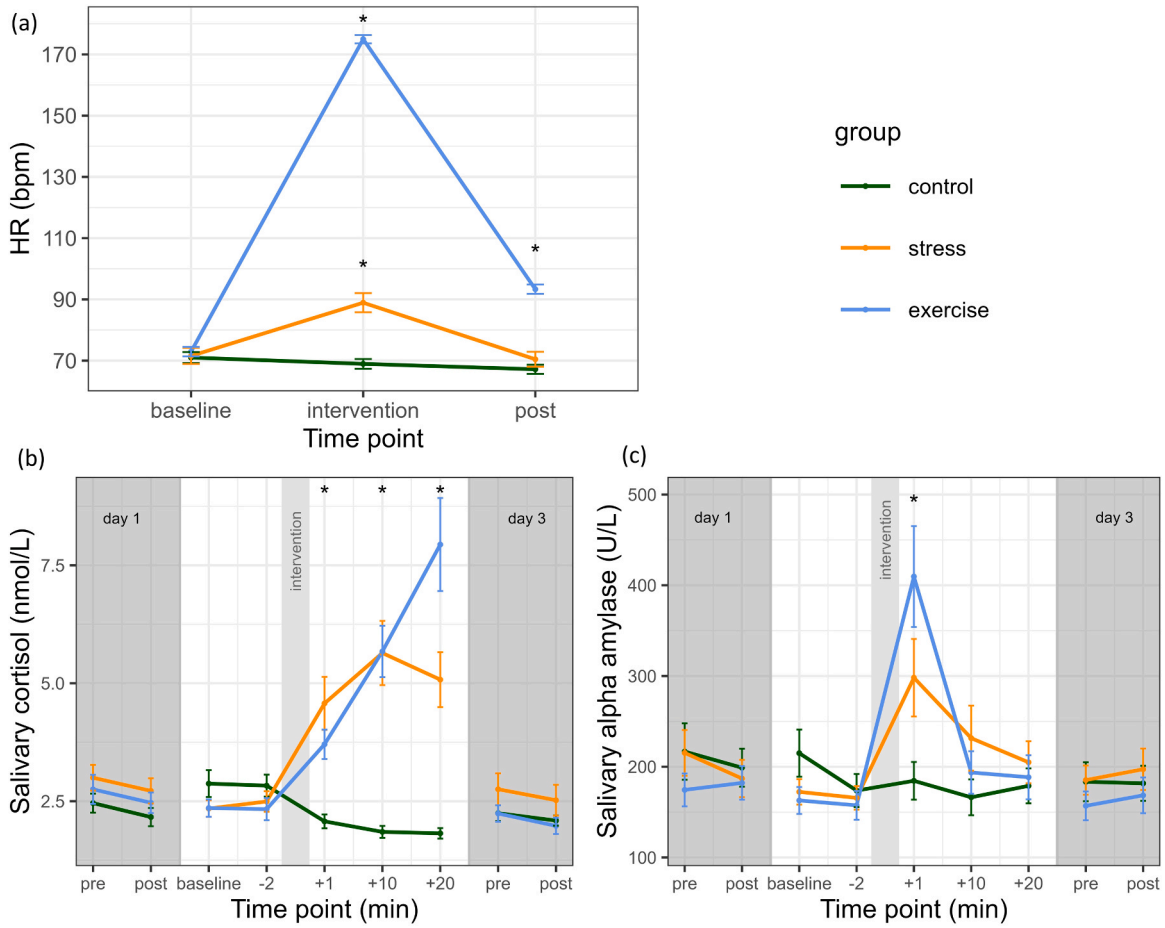


Fig. 2.. *Cardiovascular and neuroendocrine responses to the stress, exercise, and control intervention, Note.* (a) Mean (±SEM) heart rate in bpm, (b) salivary cortisol concentrations and (c) salivary alpha amylase levels per group per time point. The stress and exercise intervention both significantly increased heart rate, salivary cortisol and salivary alpha amylase levels compared to the control intervention. Heart rate was also higher in the exercise compared to the stress group during the intervention and post-measurement, and cortisol levels were higher in the exercise compared to the stress group at time point + 20 min. *$p < .05$.

not differ at baseline ($F(2, 111) = 0.24$, $p = .789$), the exercise and stress group both showed significantly higher HR than the control group during ($F(2, 111) = 672.40$, $p = <.001$, partial $\eta^2 = 0.92$; exercise>control: $t(72.35) = -50.51$, $p = <.001$, $d = 11.59$; stress>control $t(55.02) = -5.70$, $p = <.001$, $d = 1.31$) and after the intervention ($F(2, 111) = 58.23$, $p = <.001$, partial $\eta^2 = 0.51$; exercise>control: ($t(74.00) = -12.27$, $p = <.001$, $d = 2.81$). Exercise elicited a higher HR than the stress group ($t(50.69) = -25.24$, $p = <.001$, $d = 5.79$), that remained during the post measurement ($t(61.83) = -7.99$, $p = <.001$, $d = 1.83$). Women (M=88.90 bpm, SD=34.39) had an overall higher HR than men (M=84.43 bpm, SD=35.06; main effect sex: $F(1, 108) = 6.36$, $p = .013$, partial $\eta^2 = 0.06$), especially in the stress group (sex×group interaction: $F(2, 108) = 5.94$, $p = .004$, partial $\eta^2 = 0.10$; $t(111.45) = -4.50$, $p = <.001$, $d = 0.84$, but not in the exercise ($t(109.80) = -0.32$, $p = .750$, $d = 0.06$) or control group ($t(107.90) = 1.05$, $p = .297$, $d = -0.20$; S2.2 and Figure S1).

3.1.2. Salivary cortisol and alpha amylase

The stress and the exercise intervention both significantly increased salivary cortisol compared to the control intervention (Fig. 2b; group×time interaction: $F(5.75, 321.86) = 17.32$, $p = <.001$, partial $\eta^2 = 0.24$; main effect group: $F(2, 112) = 6.86$, $p = .002$, partial $\eta^2 = 0.11$; main effect time: $F(2.87, 321.86) = 28.07$, $p = <.001$, partial $\eta^2 = 0.20$). Cortisol levels were significantly higher + 1 min ($F(2, 117) = 11.80$, $p = <.001$, partial $\eta^2 = 0.17$) and + 10 min ($F(2, 117) = 25.98$, $p = <.001$, partial $\eta^2 = 0.31$) after the stress ($t(56.07) = -4.42$, $p = <.001$, $d = 0.99$ and $t(51.54) = -6.14$, $p = <.001$, $d = 1.37$) and exercise intervention ($t(68.08) = -4.71$, $p = <.001$, $d = 1.05$ and $t(56.10) = -7.78$, $p = <.001$, $d = 1.74$) compared to the control intervention, but did not differ between stress and exercise ($t(70.75) = 0.68$, $p = .499$ and $t(75.91) = -0.51$, $p = .610$). At time point + 20 min, cortisol levels were still highest in the exercise group, followed by the stress (exercise>stress: $t(75.73) = -2.21$, $p = .030$, $d = 0.49$; exercise>control: $t(48.07) = -8.02$, $p = <.001$, $d = 1.79$), and control group (stress>control: $t(51.69) = -6.18$, $p = <.001$, $d = 1.38$; $F(2, 117) = 30.63$, $p = <.001$, partial $\eta^2 = 0.34$). Cortisol levels at the other time points did not significantly differ between groups (all p's > .05).

For the exploratory sex analyses, there was a significant group×sex interaction ($F(2, 112) = 3.45$, $p = .035$, partial $\eta^2 = 0.06$) and a group×time×sex interaction ($F(5.75, 321.86) = 3.41$, $p = .003$, partial $\eta^2 = 0.06$). In the stress group, men had significantly higher cortisol levels compared to women at time point + 1 min ($t(37.60) = 2.12$, $p = .040$, $d = 0.67$). At time point + 10 min, cortisol levels differed between men and women both in the stress group (men > women; $t(36.52) = 2.30$, $p = .028$, $d = 0.73$) and in the exercise group (women > men; $t(38.00) = -2.30$, $p = .027$, $d = -0.73$). In the exercise group, women also had significantly higher cortisol levels at time point + 20 min ($t(37.95) = -2.12$, $p = .041$, $d = -0.67$). Moreover, in the stress group, men had significantly higher cortisol levels at the end of the first testing day ($t(37.38) = 2.37$, $p = .023$, $d = 0.75$) and at the beginning of the third testing day ($t(36.56) = 2.46$, $p = .019$, $d = 0.78$). All other comparisons were not significant (all p's > .05). The main effect of sex and the group×time interaction were also not significant (both p's > .05; S2.3 and Figure S2).

The stress and exercise intervention both significantly increased sAA levels compared to the control intervention (Fig. 2c; group×time interaction: $F(9.66, 541.24) = 6.10$, $p = <.001$, partial $\eta^2 = 0.10$; main effect group: $F(2, 112) = 0.31$, $p = .731$; main effect time: $F(4.83, 541.24) = 15.88$, $p = <.001$, partial $\eta^2 = 0.12$). sAA was significantly increased immediately after the stress ($t(77.97) = -2.58$, $p = .024$, $d = 0.58$) and exercise ($t(75.42) = -3.46$, $p = .003$, $d = 0.77$) but not after the control intervention ($F(2, 117) = 6.70$, $p = .018$, partial $\eta^2 = 0.10$). sAA did not significantly differ between groups at other time points (all p's > .05). There were no main or interaction effects for sex (all p's > .05; Figure S3).

3.2. Skin conductance responses

3.2.1. Fear acquisition

Fear acquisition as measured by SCRs was successful (Fig. 3a; main effect trial: $F(5.85, 632.31) = 9.85$, $p = <.001$, partial $\eta^2 = 0.08$; main effect CS: $F(1, 108) = 108.24$, $p = <.001$, partial $\eta^2 = 0.50$; trial×CS interaction: $F(5.98, 645.89) = 10.88$, $p = <.001$, partial $\eta^2 = 0.09$). From trial 2 onwards, SCRs were significantly higher in response to the CS+ compared to the CS- (all p's < .001). There were no significant main or interaction effects with group (all p's > .05). However, there was a significant sex×trial interaction ($F(5.85, 632.31) = 2.40$, $p = .028$, partial $\eta^2 = 0.02$) and a sex×trial×CS interaction ($F(5.98, 645.89) = 2.91$, $p = .008$, partial $\eta^2 = 0.03$). Men had significantly larger SCR to the CS+ on the last acquisition trial compared to women ($t(108.79) = 2.47$, $p = .015$, $d = 0.46$).

3.2.2. Fear extinction

For fear extinction training, the main effect of trial (Fig. 3b; $F(4.94, 533.65) = 41.10$, $p = <.001$, partial $\eta^2 = 0.28$), CS ($F(1, 108) = 91.33$, $p = <.001$, partial $\eta^2 = 0.46$) and trial×CS interaction ($F(6.12, 661.35) = 7.08$, $p = <.001$, partial $\eta^2 = 0.06$) were significant. Although Fig. 3b shows a decrease in differential fear responding over trials, post-hoc comparisons revealed significant differences between CS+ and CS- on all trials (all p's < .001), indicating successful, but incomplete extinction learning. Again, there were no main or interaction effects with group or sex (all p's > .05).

3.2.3. Retrieval and reinstatement test

For the retrieval phase, a main effect of context ($F(1.85, 201.85) = 5.11$, $p = .008$, partial $\eta^2 = 0.04$), trial ($F(2.20, 239.46) = 48.53$, $p = <.001$, partial $\eta^2 = 0.31$), and CS ($F(1, 109) = 41.22$, $p = <.001$, partial $\eta^2 = 0.27$), and a trial×CS interaction ($F(2.82, 307.39) = 11.18$, $p = <.001$, partial $\eta^2 = 0.09$) occurred (Fig. 4a). Independent of CS, SCRs were larger in context A compared to B ($t(925) = 3.28$, $p = .003$, $d = 0.11$) and C ($t(925) = 2.49$, $p = .026$, $d = 0.08$), while they did not significantly differ between context B and C ($t(925) = -0.86$, $p = .391$), indicating increased responding in the original threat context. However, neither the main effect of group ($F(2, 109) = 1.64$, $p = .198$), nor the expected interactions with group were significant (all p's > .05). There were no main or interaction effects with the exploratory factor sex (all p's > .05). When considering only the first trial in order, to exclude the re-extinction process during the retrieval phase, the main effect of context ($F(1.90, 208.45) = 4.03$, $p = .021$, partial $\eta^2 = 0.04$) and CS ($F(1, 110) = 66.84$, $p = <.001$, partial $\eta^2 = 0.38$) were still significant, but there was again no main effect ($F(2, 110) = 3.06$, $p = .051$) or any interaction with group or sex (all p's > .05). Post-hoc comparisons again revealed higher SCRs in context A relative to B ($t(231) = 2.64$, $p = .026$, $d = 0.17$). As some previous studies (Meir Drexler et al., 2017, 2018) tested only men, we re-ran the analyses for men and women separately, for exploratory purposes, but they did not yield any significant main effect or interaction including the factor group (all p's > .05; Supplement Figure S6).

For the reinstatement test, SCRs were larger towards the CS+ as compared to the CS- (main effect CS: $F(1, 106) = 11.19$, $p = .001$, partial $\eta^2 = 0.10$) and there was a significant effect of trial ($F(3, 318) = 4.48$, $p = .004$, partial $\eta^2 = 0.04$) and significant trial×CS ($F(3, 318) = 4.63$, $p = .003$, partial $\eta^2 = 0.04$) and trial×CS×context×sex interactions ($F(5.27, 558.14) = 2.33$, $p = .039$, partial $\eta^2 = 0.02$; S2.5.2; Fig. 4b). For the exploratory sex analyses, men had larger SCRs for CS+ on trial 1 in context A ($t(110.96) = 1.99$, $p = .049$, $d = 0.38$), and for CS- on trial 2 in context B ($t(91.38) = 2.25$, $p = .027$, $d = 0.43$), and for CS- on trial 3 in context C ($t(104.73) = 2.06$, $p = .041$, $d = 0.39$), compared to women. Neither the main effect of group ($F(2, 106) = 1.29$, $p = .280$) nor the expected interactions with group were significant (all p's > .05). The main effect or other interaction effects with the exploratory factor sex were not significant (all p's > .05). When only considering the first trial in order, there was a significant main effect of CS ($F(1, 107) = 19.01$, $p = <.001$, partial $\eta^2 = 0.15$) and a sex×CS×context interaction ($F(2, 214) = 3.23$, $p = .042$, partial $\eta^2 = 0.03$), indicating that men had higher SCR values to CS+ relative to women in context A ($t(110.96) = 1.99$, $p = .049$, $d = 0.15$). Again, the main effect of group ($F(2, 107) = 0.69$, $p = .502$) or any interaction with group were not significant (all p's > .05).

Separate exploratory analyses for men only revealed a group×context interaction ($F(4, 102) = 3.13$, $p = .018$, partial $\eta^2 = 0.11$). Independent of CS, SCRs in context A were lower in the stress compared to the control group ($t(260.12) = 2.46$, $p = .044$, $d = 0.29$; Fig. 5a), indicating a generalized reinstatement of fear in the control but less so in the stress group. In context B, the exercise group showed overall decreased SCRs compared to the control ($t(255.04) = 2.33$, $p = .041$, $d = 0.27$) and stress group ($t(279.54) = 2.86$, $p = .014$, $d = 0.33$; Fig. 5a), suggesting generally less responding after reinstatement in the extinction context in the exercise relative to the other groups. Analyses for women separately did not yield any significant effect or any interaction including the factor group (all p's > .05; Fig. 5b).

3.3. Pupillary responses

3.3.1. Fear acquisition

Fear acquisition as measured by pupil dilation was successful (Fig. 3c; main effect trial: $F(7, 616) = 3.69$, $p = <.001$, partial $\eta^2 = 0.04$); main effect CS: $F(1, 88) = 68.04$, $p = <.001$, partial $\eta^2 = 0.44$; trial×CS interaction: $F(6.16, 541.94) = 4.12$, $p = <.001$, partial $\eta^2 = 0.04$). From trial 2 onwards, pupil size was significantly larger in response to CS+ compared to CS- (all p's < .001). There were no main or interaction effects with group (all p's > .05), except for a group×trial×CS interaction ($F(12.32, 541.94) = 2.22$, $p = .009$, partial $\eta^2 = 0.05$). This interaction was driven by differences between groups in differential responding on trial 5 (no differential responding in the exercise group), trial 6 and 7 (no differential responding in the control group) as well as trial 8 (no differential responding in the stress group). Moreover, men had larger pupil sizes over all trials and CS types ($F(1, 88) = 4.04$, $p = .047$, partial $\eta^2 = 0.04$). There were no interaction effects with sex (all p's > .05).

3.3.2. Fear extinction

For fear extinction, the main effect of trial (Fig. 3d; $F(7, 595) = 3.21$, $p = .002$, partial $\eta^2 = 0.04$), CS ($F(1, 85) = 62.67$, $p = <.001$, partial $\eta^2 = 0.42$), and the trial×CS interaction ($F(7, 595) = 2.58$, $p = .012$, partial $\eta^2 = 0.03$) were significant. Although Fig. 3d shows a decrease in differential fear responding over trials, post hoc comparisons revealed significant differences on all trials (all p's < .05), except for trial 7 ($t(103) = -1.92$, $p = .057$), indicating successful, but incomplete extinction learning. Moreover, men had significantly larger pupil sizes than women over all trials and CS types ($F(1, 85) = 4.66$, $p = .034$, partial $\eta^2 = 0.05$). There were no main or interaction effects with group and no interaction effects with sex (all p's > .05).

3.3.3. Retrieval and reinstatement test

For the retrieval phase, the main effect of context ($F(2, 156) = 6.30$, $p = .002$, partial $\eta^2 = 0.08$), trial ($F(3, 234) = 7.25$, $p = <.001$, partial

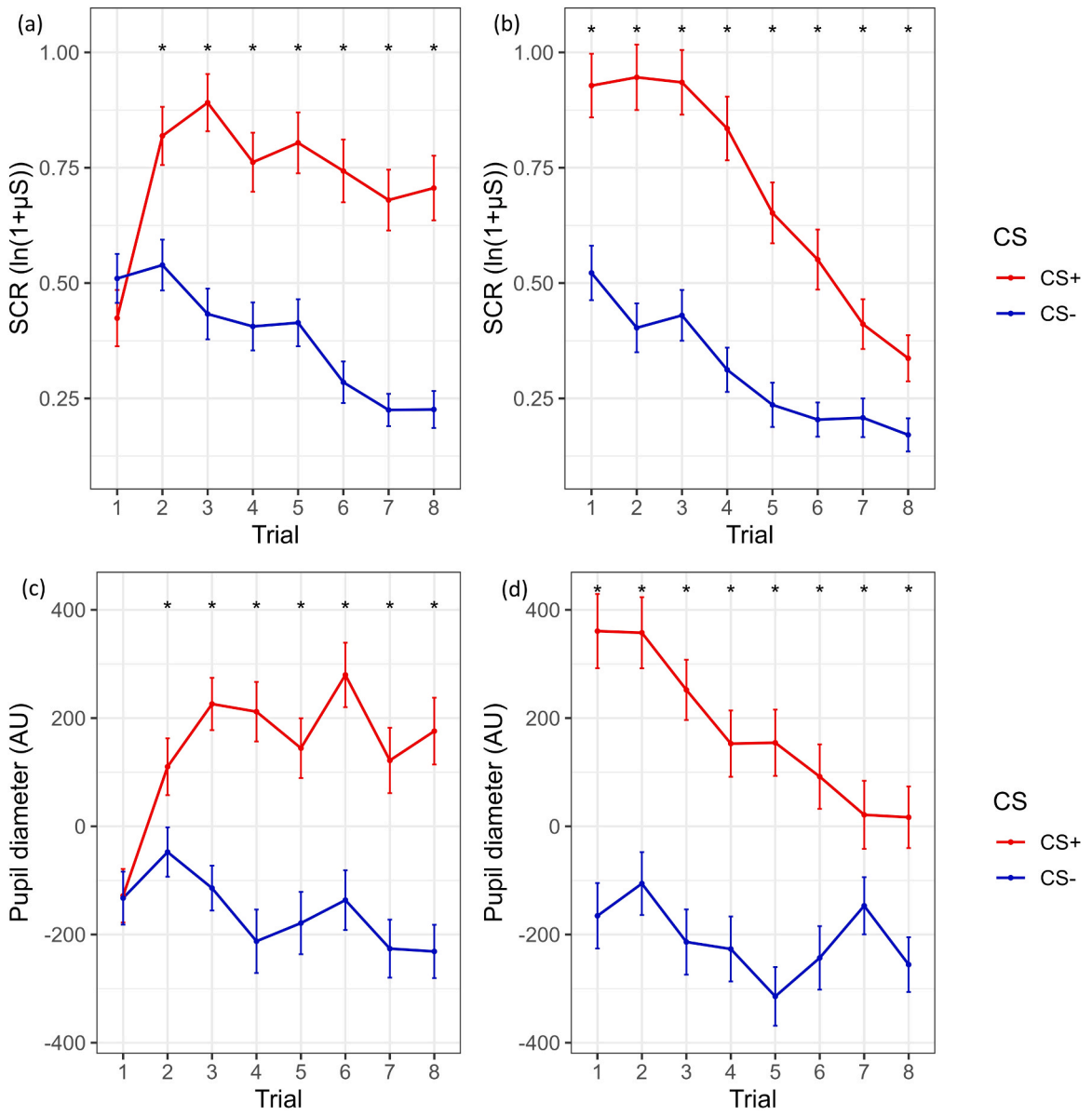


Fig. 3. *SCRs and pupillary responses during fear acquisition and extinction training, Note.* Mean (±SEM) amplitudes of SCRs (a, b) and pupil dilation in arbitrary units (AU; c, d) for CS+ and CS- separately during (a, c) fear acquisition and (b, d) extinction training. Both SCRs and pupil dilation were significantly larger to CS+ compared to CS- during fear acquisition, indicating successful fear learning. Differential responding decreased during extinction, but remained significant on the last trial, indicating successful, but incomplete extinction learning. *$p < .05$.

η^2= 0.08) and CS ($F(1, 78)$= 63.34, p = <.001, partial η^2= 0.45) as well as the trial×CS interaction ($F(3, 234)$= 8.77, p = <.001, partial η^2= 0.10) were significant (Fig. 4c). Consistent with SCR data, pupillary responses were overall significantly larger in context A compared to B ($t(735) = 3.43$, $p = .002$, d= 0.13) and context C ($t(734) = 3.45$, $p = .002$, d= 0.13), while they did not differ between context B and C ($t(733) = -0.03$, $p = .973$), again indicating increased responding in the original threat context. The main effect of group ($F(2,78)$= 0.41, $p = .666$) and the expected interactions with group were not significant (all p's > .05). However, we found a group×trial interaction ($F(6, 234)$= 2.28, $p = .037$, partial η^2= 0.06). The stress group had a larger pupil size on trial 1 compared to trial 3 ($t(190) = 4.42$, p = <.001, d= 0.32) and trial 4 ($t(191) = 3.50$, $p = .003$, d= 0.25), the exercise group had a larger pupil size on trial 1 compared to trial 4 ($t(187) = 3.26$, $p = .008$, d= 0.24), and the control group had a larger pupil size on trial 1 compared to trial 2 ($t(171) = 2.76$, $p = .039$, d= 0.21), irrespective of the CS. No other post-hoc comparisons were significant (all p's > .05). Moreover, the group×sex×context interaction was significant ($F(4, 156)$= 2.53, $p = .043$, partial η^2= 0.06), driven by generally larger pupil sizes in context A vs. B in men in the exercise group ($t(129) = 3.09$, $p = .007$, d= 0.27). There were no main or other interaction effects with the exploratory factor sex (all p's > .05). When only considering the first trial, there was a significant main effect of CS ($F(1, 86)$= 63.91, p = <.001, partial η^2= 0.43), but no main effect of group ($F(2, 86)$= 0.81, $p = .448$) or any interaction with group or sex (all p's > .05).

For the reinstatement test, only the main effect of context was significant ($F(2, 142)$= 4.53, $p = .012$, partial η^2= 0.06; Fig. 4d). Pupillary responses were significantly larger in context A compared to B ($t(716) = 3.27$, $p = .003$, d= 0.12) and C ($t(720) = 2.44$, $p = .029$, d= 0.09), while they did not differ between context B and C ($t(721) = -1.11$, $p = .266$). The main effect of group ($F(2, 71)$= 0.51, $p = .600$) and the expected interactions with group were not significant (all p's > .05). There were no main or interaction effects with the exploratory factor sex

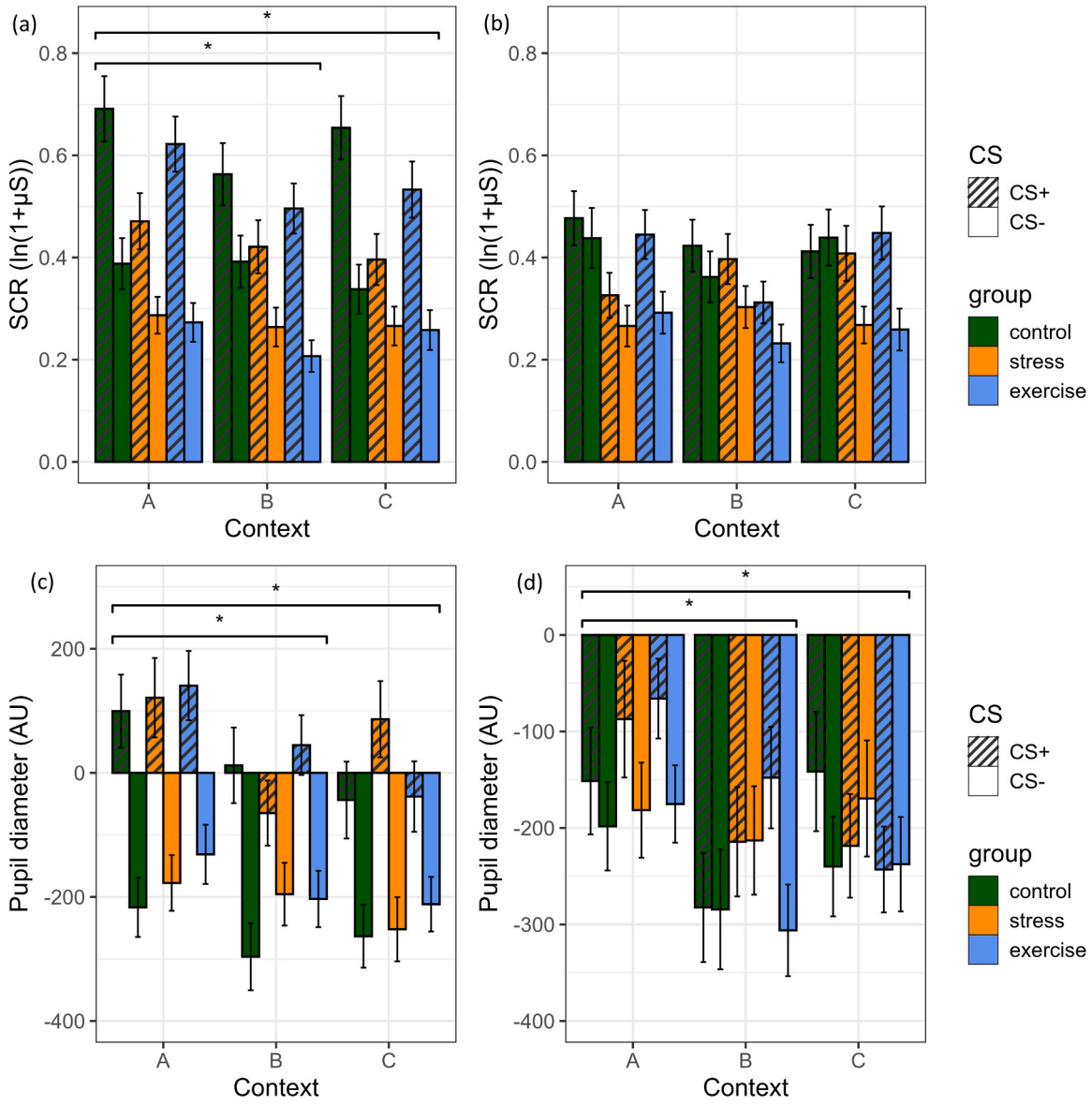


Fig. 4. *SCRs and pupillary responses during retrieval and reinstatement test, Note.* Mean (±SEM) amplitudes of the SCR (a, b) and pupil dilation in arbitrary units (AU; c, d) over all four trials, per group, per context, for CS+ and CS- separately, during (a, c) retrieval and (b, d) reinstatement test. Pupillary responses are baseline corrected using the context presentation as baseline period. Positive values reflect pupil dilation as compared to the baseline period. SCRs and pupil dilation were larger in context A compared to B and C, while they did not significantly differ between context B and C, indicating increased responding in the original threat context. *$p < .05$.

(all p's $> .05$). We repeated this mixed ANOVA for trial 1 only, but this did not yield any significant effects (all p's $> .05$). The preregistered mixed ANOVA including the factor phase (retrieval vs. reinstatement) and the mixed ANOVA on the comparison of the last retrieval trial vs. the first reinstatement trial did not yield any hypothesis-relevant significant results, neither for SCRs nor for pupillary responses (S2.5.2; S2.6.1).

3.4. Correlational analyses: heart rate, cortisol, alpha amylase

We correlated delta HR, cortisol, and sAA concentrations during the intervention on day 2, with differential fear responding during the retrieval and reinstatement test for SCRs and pupil dilation separately, and we applied Holm's method to correct for multiple testing (correlation matrices: S2.7). For SCRs during the reinstatement test, delta cortisol negatively correlated with differential fear responding in the stress group in context B ($N = 37$, ρ=-0.51, $p = .046$), but not in context A ($N = 37$, ρ=0.25, $p > .999$) or C ($N = 37$, ρ=-0.01, $p > .999$), indicating that stronger cortisol responding in response to the stress intervention was related to less CS+ /CS- differentiation in the extinction context after reinstatement on the following day (Figure S7). This correlation was also present in the control group ($N = 38$, ρ=-0.56, $p = .008$), but not in the exercise group ($N = 38$, ρ=0.04, $p > .999$). No other correlations were significant.

3.5. Exploratory analyses

Exploratory analyses on responses to context during day 3, chronic stress (S1.11), self-reported physical activity in daily life, and the interaction between extinction learning performance, cardiovascular and neuroendocrine responses to the intervention, and differential fear responding during retrieval are reported in the supplement (S2.8).

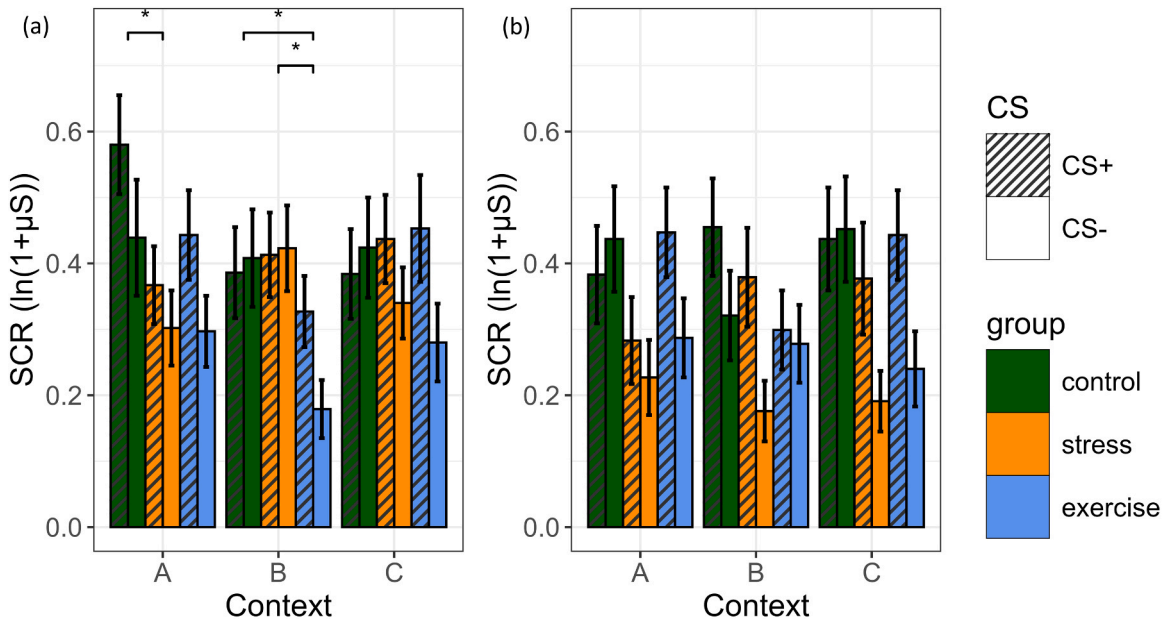


Fig. 5. *SCRs during reinstatement test for men and women separately, Note.* Mean (±SEM) amplitudes of the SCR over all four trials, per group, per context, for CS+ and CS- separately, for (a) men and (b) women. For men, SCRs in context A were lower in the stress compared to the control group. In context B, SCRs were lower in the exercise compared to the control and stress group. *p < .05.

4. Discussion

We compared the effects of pre-extinction vigorous-intensity physical exercise and psychosocial stress on contextual retrieval and generalization 24 h later. Both interventions successfully activated the SNS and the HPA axis compared to the control intervention. Fear acquisition was successful, conditioned responding decreased over extinction trials, and no group differences emerged during these phases. In contrast to our hypotheses, pre-extinction stress and exercise did not significantly improve extinction memory consolidation or generalization across contexts, neither for SCRs nor for pupillary responses, as compared to the control group. Exploratory analyses revealed that when tested after reinstatement, stress reduced the context-dependency of the extinction memory, whereas exercise reduced overall responding in the extinction context in men only. This finding is consistent with the STaR model (Meir Drexler et al., 2019) and could indicate more promising effects after reinstatement compared to retrieval and rather restricted to men than to women. However, as this was a subgroup analysis, it may have been underpowered, and further research clearly needs to test these ideas.[1]

Our main findings do not fully support the prediction of the STaR model (Meir Drexler et al., 2019) that stress before extinction training improves later retrieval and contextual generalization. In contrast to previous studies, we used the TSST as stress induction method, instead of the (Socially Evaluated) Cold Pressor Test (Bentz et al., 2013; Meir Drexler et al., 2017, 2018). The TSST has a stronger psychosocial component, has a longer duration, and more reliably induces cortisol increases (Skoluda et al., 2015). Stronger increases in cortisol may, in line with the STaR model, cause larger effects on extinction memories. However, the TSST could be considered more of an event that needs to be processed and by increasing the amount of new information on the second test day, it could interfere with the consolidation of the extinction memory (Cadle and Zoladz, 2015), thereby probably weakening any stress hormone effects that were previously observed in pharmacological or physiological studies. The current results also contradict previous pharmacological studies that cortisol administration before extinction training (Merz et al., 2018) or exposure therapy (de Quervain et al., 2017) reduce the return of fear and facilitate extinction consolidation. The current findings therefore call for a slight refinement of the predictions of the STaR model on pre-extinction stress concerning the stress induction method.

Extending previous stress studies, we investigated generalization of the extinction memory to a novel context, which is highly relevant in clinical practice. In line with previous work on cortisol administration (Merz et al., 2018), psychosocial stress did not influence extinction memory retrieval in the novel context. When visually inspecting our SCR data, however, stress might reduce fear renewal in the novel context (Fig. 4a), specifically in men (Figure S6), but this descriptive difference between the stress and control group was not statistically significant and should therefore be interpreted with caution. Yet, it could imply that, under specific conditions, stress might benefit the generalization of the extinction memory across contexts.

One of these specific conditions could be the participant's sex. Sex hormones impact fear and extinction memory processes (Merz, Kinner, et al., 2018), also in interaction with stress hormones (Merz, 2023). Indeed, our SCR data suggest that stress might improve extinction memory generalization to the original threat context in men, but not in women. Previous studies have almost solely tested men, and two studies that assessed both sexes, found an effect on extinction retrieval only in men (Bentz et al., 2013) or only in women using oral contraceptives (Peyrot et al., 2024). As our study was not powered to detect sex-moderated effects, further systematic comparisons between men and women are needed, also considering sex hormone status and hormonal contraceptives usage in women (Jentsch et al., 2022; Peyrot et al., 2024), which should be included in adequate preregistered models. Another explanation for sex-specific effects is that men typically show higher cortisol responding to psychosocial stress than women (Merz and Wolf, 2017; Figure S2). If a stress effect is mainly driven by cortisol, it would be interesting to investigate whether sex differences disappear when similar levels of cortisol are induced in both sexes.

Correlational analyses revealed that stronger cortisol responding during the TSST and control intervention, but not exercise, relate to smaller differential SCRs in the extinction context after reinstatement. This finding aligns with positive findings of cortisol on memory

[1] A sensitivity analyses revealed that the subgroup analyses with $N = 60$ has 95 % power to detect effect sizes of at least Cohen's $f = 0.27$.

consolidation (e.g., Schwabe et al., 2022). Yet, this correlation was not present in the acquisition or novel context, indicating a context-specific relation between cortisol and extinction memory retrieval, and contrasting literature on cortisol effects on contextual generalization (Bahtiyar et al., 2020). Nevertheless, this correlation also slightly confirms predictions of the STaR model regarding beneficial cortisol effects on extinction consolidation, albeit not overcoming the context-dependency of the extinction memory. Hence, other mechanisms than cortisol may be involved in generalization of extinction memories to multiple contexts.

Contrary to our expectations, vigorous-intensity physical exercise did not influence extinction memory consolidation or generalization across contexts in the full sample, but only after reinstatement in men. To test the context-specific predictions of the STaR model, we employed exercise before extinction. A meta-analysis on episodic memory, however, found that exercise effects are smallest when employed before vs. after learning (Loprinzi et al., 2019). Accordingly, null effects also occurred when exercise was performed prior to exposure therapy (Jacquart et al., 2017; Voorendonk et al., 2021). Previous studies on extinction learning therefore implemented exercise after learning (Crombie et al., 2021a, 2023; Jentsch et al., 2023; Keyan and Bryant, 2019b) and found evidence for a strengthening effect of exercise on memory. Repeating the current study, including contextual generalization, while employing exercise after extinction could provide additional valuable insights, in particular regarding translational relevance.

Another potential explanation for a lack of extinction generalization to other contexts in the exercise group may be an overly strong SNS activation in response to the exercise relative to the stress intervention, especially for sAA in women. While noradrenergic arousal is needed for successful extinction processes, too much arousal might impair later extinction retrieval (Bierwirth and Stockhorst, 2022). Besides, heart rate dynamics might play an important role in regulating fear and extinction expression, possibly via the ventromedial prefrontal cortex (Battaglia et al., 2025). The STaR model does not specify the role of SNS activity, but the current findings hint at an important boundary condition that may have affected women more strongly than men. Future research should determine the optimal level of SNS activation, for example by testing different intensities of physical exercise, while also taking potential sex differences into account. Nonetheless, this research should consider that moderate to high intensities are needed to reliably induce a cortisol response (Hill et al., 2008). Hence, we chose a vigorous exercise intensity that was also used in previous studies from our lab (Jentsch and Wolf, 2020; Jentsch et al., 2023).

Testing different exercise intensities is also important because our vigorous-intensity exercise intervention did not increase positive emotions and was rated as strenuous, which is not surprising considering the high intensity (Ekkekakis et al., 2011). Importantly, however, the exercise intervention also did not induce negative emotions, in contrast to the TSST. Nevertheless, it raises questions regarding the (emotional) tolerability of this intervention in clinical practice. Perhaps moderate-intensity exercise can be beneficial for extinction memories while inducing positive emotions, thereby being a more suitable intervention in clinical practice. Moreover, women showed a stronger cortisol increase compared to men (Figure S2), again underlining the need to account for sex differences.

Furthermore, participants were willing to undergo an exercise intervention and may therefore differ from the general or clinical population in attitude towards exercise or exercise frequency, potentially influencing our results. For example, fitness level can impact memory performance and the effectiveness of an exercise intervention on memory (e.g., Loprinzi et al., 2023). Although self-reported physical activity was not related to any outcome measure (S2.8.3), more objective fitness level measures could offer useful insights into whether specific individuals benefit more from exercise as an adjunct to extinction learning. Relatedly, we tested a healthy student sample to examine the effects of stress and exercise and the underlying mechanisms in a controlled laboratory setting. Clinical samples may, however, differ in several regards from our sample, such as lower cognitive abilities, or deficits in extinction learning or recall (e.g., Colvonen et al., 2019). To assess translational value, further research needs to test these questions in clinical samples, e.g., in patients with anxiety or stress-related disorders.

To influence extinction memory retrieval, fear and extinction learning need to be successful. If extinction learning is incomplete, exercise may further improve fear rather than extinction memory consolidation (Jentsch et al., 2023). In this study, we matched the number of extinction trials with those for fear acquisition, but this did not prevent extinction learning from being incomplete. Yet, fear responding to CS+ clearly diminished over extinction trials, indicating that an extinction memory was formed, even though extinction learning was not fully complete at the end of extinction training. Future research should consider increasing the number of extinction trials even further to establish full extinction learning. However, when the extinction memory is too strong and no return of fear occurs, extinction-related interventions cannot have any additional benefits. We thus need to have some room to improve the extinction memory, while also ensuring that sufficient extinction learning occurs.

Another limitation of the current study is that we solely measured physiological fear responses with SCRs and pupil dilation, and no behavioral or self-report responses. Some studies only found effects of exercise on extinction memories when assessing US expectancy ratings (Crombie et al., 2023, 2021a). Hence, assessing ratings in further studies would provide useful information as to whether stress hormones elicit different effects on these measures than on physiology. It should be noted, however, that trial-by-trial ratings may influence physiological responding and draw attention to possible CS-US contingencies, therefore altering the learning process itself (Lonsdorf et al., 2017). Thus, combining multiple measures within one study is not always ideal.

To conclude, after successful activation of the SNS and the HPA axis, pre-extinction psychosocial stress and vigorous-intensity physical exercise improved extinction retrieval and generalization across contexts, but only after reinstatement and only in men but not in women. These findings shed light on potential boundary conditions of the stress effects as predicted by the STaR model, such as sex differences and optimal SNS activation, that call for further experiments and refinements of the model. Moreover, more studies on physical exercise are needed to determine whether and how it can be a suitable alternative to stress for improving extinction memory strength and its generalization across contexts. For translational purposes, lower intensity exercise as well as shifting the timing of physical exercise from pre-to post-extinction should be prioritized.

CRediT authorship contribution statement

Lianne N. Wolsink: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Investigation, Formal analysis, Data curation. **Christian J. Merz:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Valerie L. Jentsch:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **Oliver T. Wolf:** Writing – review & editing, Funding acquisition, Conceptualization.

Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) within SFB 1280 Extinction Learning (grant no. 316803389 - SFB1280; project A09).

Declaration of Competing Interest

We declare no competing interests.

Acknowledgements

We would like to thank Katharina Beck for help with study preparation and piloting, Daniel Bienert, Kira Buschkämper, and Lisa Kahl for help with data collection, and Tobias Otto for technical support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.psyneuen.2026.107793.

References

Allen, A.P., Kennedy, P.J., Cryan, J.F., Dinan, T.G., Clarke, G., 2014. Biological and psychological markers of stress in humans: focus on the trier social stress test. Neurosci. Biobehav. Rev. 38, 94–124. https://doi.org/10.1016/j.neubiorev.2013.11.005.

Aust, F., Barth, M., 2024. papaja: Prepare reproducible APA journal articles with R Markdown. R. Package Version. ⟨https://github.com/crsh/papaja⟩.

Azar, A., Hubert, T., Adams, T.G., Cisler, J.M., Crombie, K.M., 2024. Exercise and Fear and Safety Learning. In: Cisler, J.M., Crombie, K.M., Adams, T.G. (Eds.), Exercise and Mental Health. Current Topics in Behavioral Neurosciences, 67. Springer, Cham. https://doi.org/10.1007/7854_2024_494.

Bahtiyar, S., Karaca, K.G., Henckens, M.J., Roozendaal, B., 2020. Norepinephrine and glucocorticoid effects on the brain mechanisms underlying memory accuracy and generalization. Mol. Cell. Neurosci. 108, 103537. https://doi.org/10.1016/j.mcn.2020.103537.

Basso, J.C., Suzuki, W.A., 2017. The effects of acute exercise on mood, cognition, neurophysiology, and neurochemical pathways: a review. Brain Plast. 2, 127–152. https://doi.org/10.3233/BPL-160040.

Battaglia, S., Andero, R., Thayer, J.F., 2025. Translational cross-species evidence of heart-related dynamics in threat learning. Neurosci. Biobehav. Rev., 106273 https://doi.org/10.1016/j.neubiorev.2025.106273.

Bentz, D., Michael, T., Wilhelm, F.H., Hartmann, F.R., Kunz, S., Rohr, I.R.R. von, Quervain, D.J.-F. de, 2013. Influence of stress on fear memory processes in an aversive differential conditioning paradigm in humans. Psychoneuroendocrinology 38, 1186–1197. https://doi.org/10.1016/j.psyneuen.2012.12.018.

Bierwirth, P., Stockhorst, U., 2022. Role of noradrenergic arousal for fear extinction processes in rodents and humans. Neurobiol. Learn. Mem. 194, 107660. https://doi.org/10.1016/j.nlm.2022.107660.

Borg, G., 1998. Borg's perceived exertion and pain scales. Hum. Kinet.

Bouton, M.E., 2004. Context and behavioral processes in extinction. Learn. Mem. (Cold Spring Harb. N. Y.) 11 (5), 485–494. https://doi.org/10.1101/lm.78804.

Brainard, D.H., Vision, S., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Cadle, C.E., Zoladz, P.R., 2015. Stress time-dependently influences the acquisition and retrieval of unrelated information by producing a memory of its own. Front. Psychol. 6, 910. https://doi.org/10.3389/fpsyg.2015.00910.

Colvonen, P.J., Straus, L.D., Acheson, D., Gehrman, P., 2019. A review of the relationship between emotional learning and memory, sleep, and PTSD. Curr. Psychiatry Rep. 21 (1), 2. https://doi.org/10.1007/s11920-019-0987-2.

Craske, M.G., Hermans, D., Vervliet, B., 2018. State-of-the-art and future directions for extinction as a translational model for fear and anxiety. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 373, 20170025. https://doi.org/10.1098/rstb.2017.0025.

Crombie, K.M., Sartin-Tarm, A., Sellnow, K., Ahrenholtz, R., Lee, S., Matalamaki, M., Almassi, N.E., Hillard, C.J., Koltyn, K.F., Adams, T.G., Cisler, J.M., 2021b. Exercise-induced increases in Anandamide and BDNF during extinction consolidation contribute to reduced threat following reinstatement: preliminary evidence from a randomized controlled trial. Psychoneuroendocrinology 132, 105355. https://doi.org/10.1016/j.psyneuen.2021.105355.

Crombie, K.M., Sartin-Tarm, A., Sellnow, K., Ahrenholtz, R., Lee, S., Matalamaki, M., Adams, T.G., Cisler, J.M., 2021a. Aerobic exercise and consolidation of fear extinction learning among women with posttraumatic stress disorder. Behav. Res. Ther. 142, 103867. https://doi.org/10.1016/j.brat.2021.103867.

Crombie, K.M., Azar, A., Botsford, C., Heilicher, M., Moughrabi, N., Gruichich, T.S., Schomaker, C.M., Dunsmoor, J.E., Cisler, J.M., 2023. Aerobic exercise after extinction learning reduces return of fear and enhances memory of items encoded during extinction learning. Ment. Health Phys. Act. 24. https://doi.org/10.1016/j.mhpa.2023.100510.

Dawson, M.E., Schell, A.M., Filion, D.L., 2017. The electrodermal system. In: Cacioppo, In.J.T., Tassinary, L.G., Berntson, G. (Eds.), Handbook of psychophysiology. Cambridge university press, pp. 217–243. https://doi.org/10.1017/9781107415782.010.

Ekkekakis, P., Parfitt, G., Petruzzello, S.J., 2011. The pleasure and displeasure people feel when they exercise at different intensities. Sports Med. 41, 641–671. https://doi.org/10.2165/11590680-000000000-00000.

Finke, J.B., Roesmann, K., Stalder, T., Klucken, T., 2021. Pupil dilation as an index of Pavlovian conditioning. A systematic review and meta-analysis. Neurosci. Biobehav. Rev. 130, 351–368. https://doi.org/10.1016/j.neubiorev.2021.09.005.

Fitzgerald, P.J., Seemann, J.R., Maren, S., 2014. Can fear extinction be enhanced? A review of pharmacological and behavioral findings. Brain Res. Bull. 105, 46–60. https://doi.org/10.1016/j.brainresbull.2013.12.007.

Godin, G., 2011. The Godin-Shephard leisure-time physical activity questionnaire. Health Fit. J. Can. 4, 18–22. https://doi.org/10.14288/HFJC.V4I1.82.

Hermann, A., Stark, R., Milad, M.R., Merz, C.J., 2016. Renewal of conditioned fear in a novel context is associated with hippocampal activation and connectivity. Soc. Cogn. Affect. Neurosci. 11, 1411–1421. https://doi.org/10.1093/scan/nsw047.

Hill, E.E., Zack, E., Battaglini, C., Viru, M., Viru, A., Hackney, A.C., 2008. Exercise and circulating cortisol levels: the intensity threshold effect. J. Endocrinol. Investig. 31, 587–591. https://doi.org/10.1007/BF03345606.

Jacquart, J., Roquet, R.F., Papini, S., Powers, M.B., Rosenfield, D., Smits, J.A.J., Monfils, M.-H., 2017. Effects of acute exercise on fear extinction in rats and exposure therapy in humans: null findings from five experiments. J. Anxiety Disord. 50, 76–86. https://doi.org/10.1016/j.janxdis.2017.05.010.

Jentsch, V.L., Wolf, O.T., 2020. Acute physical exercise promotes the consolidation of emotional material. Neurobiol. Learn. Mem. 173, 107252. https://doi.org/10.1016/j.nlm.2020.107252.

Jentsch, V.L., Pötzl, L., Wolf, O.T., Merz, C.J., 2022. Hormonal contraceptive usage influences stress hormone effects on cognition and emotion. Front. Neuroendocrinol. 67, 101012. https://doi.org/10.1016/j.yfrne.2022.101012.

Jentsch, V.L., Wolf, O.T., Otto, T., Merz, C.J., 2023. The impact of physical exercise on the consolidation of fear extinction memories. Psychophysiology 60, e14373. https://doi.org/10.1111/psyp.14373.

Joëls, M., Baram, T.Z., 2009. The neuro-symphony of stress. Nat. Rev. Neurosci. 10, 459–466. https://doi.org/10.1038/nrn2632.

Keyan, D., Bryant, R.A., 2019b. Acute exercise-induced enhancement of fear inhibition is moderated by BDNF Val66Met polymorphism. Transl. Psychiatry 9, 131. https://doi.org/10.1038/s41398-019-0464-z.

Keyan, D., Bryant, R.A., 2019a. The capacity for acute exercise to modulate emotional memories: a review of findings and mechanisms. Neurosci. Biobehav. Rev. 107, 438–449. https://doi.org/10.1016/j.neubiorev.2019.09.033.

Kirschbaum, C., Pirke, K.M., Hellhammer, D.H., 1993. The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology 28, 76–81. https://doi.org/10.1159/000119004.

Lonsdorf, T.B., Menz, M.M., Andreatta, M., Fullana, M.A., Golkar, A., Haaker, J., Merz, C. J., 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.

Loprinzi, P.D., Blough, J., Crawford, L., Ryu, S., Zou, L., Li, H., 2019. The temporal effects of acute exercise on episodic memory function: systematic review with meta-analysis. Brain Sci. 9, 87. https://doi.org/10.3390/brainsci9040087.

Loprinzi, P.D., Roig, M., Tomporowski, P.D., Javadi, A.H., Kelemen, W.L., 2023. Effects of acute exercise on memory: Considerations of exercise intensity, post-exercise recovery period and aerobic endurance. Mem. Cogn. 51, 1011–1026. https://doi.org/10.3758/s13421-022-01373-4.

Mathôt, S., Fabius, J., Van Heusden, E., Van der Stigchel, S., 2018. Safe and sensible preprocessing and baseline correction of pupil-size data. Behav. Res. Methods 50, 94–106. https://doi.org/10.3758/s13428-017-1007-2.

McHugh, R.K., Whitton, S.W., Peckham, A.D., Welge, J.A., Otto, M.W., 2013. Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. J. Clin. Psychiatry 74, 13979. https://doi.org/10.4088/JCP.12r07757.

Meir Drexler, S., Hamacher-Dang, T.C., Wolf, O.T., 2017. Stress before extinction learning enhances and generalizes extinction memory in a predictive learning task. Neurobiol. Learn. Mem. 141, 143–149. https://doi.org/10.1016/j.nlm.2017.04.002.

Meir Drexler, S., Merz, C.J., Wolf, O.T., 2018. Pre-extinction stress prevents context-related renewal of fear. Behav. Ther. 49, 1008–1019. https://doi.org/10.1016/j.beth.2018.03.001.

Meir Drexler, S., Merz, C.J., Jentsch, V.L., Wolf, O.T., 2019. How stress and glucocorticoids timing-dependently affect extinction and relapse. Neurosci. Biobehav. Rev. 98, 145–153. https://doi.org/10.1016/j.neubiorev.2018.12.029.

Merz, C.J., 2023. How different factors in combination change fear extinction learning: the case of sex and stress hormones. Fear Extinction: From Basic Neuroscience to Clinical Implications. Springer International Publishing, Cham, pp. 179–191.

Merz, C.J., Wolf, O.T., 2017. Sex differences in stress effects on emotional learning. J. Neurosci. Res. 95, 93–105. https://doi.org/10.1002/jnr.23811.

Merz, C.J., Wolf, O.T., 2022. How stress hormones shape memories of fear and anxiety in humans. Neurosci. Biobehav. Rev. 142, 104901. https://doi.org/10.1016/j.neubiorev.2022.104901.

Merz, C.J., Hamacher-Dang, T.C., Stark, R., Wolf, O.T., Hermann, A., 2018. Neural underpinnings of cortisol effects on fear extinction. Neuropsychopharmacology 43, 384–392. https://doi.org/10.1038/npp.2017.227.

Merz, C.J., Kinner, V.L., Wolf, O.T., 2018. Let's talk about sex … differences in human fear conditioning. Curr. Opin. Behav. Sci. 23, 7–12. https://doi.org/10.1016/j.cobeha.2018.01.021.

Milad, M.R., Wright, C.I., Orr, S.P., Pitman, R.K., Quirk, G.J., Rauch, S.L., 2007. Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. Biol. Psychiatry 62, 446–454. https://doi.org/10.1016/j.biopsych.2006.10.011.

Otto, T., Wolf, O.T., Merz, C.J., 2023. EDA-Analysis App (Version v3). Zenodo. https://doi.org/10.5281/zenodo.15525784.

Peyrot, C., Duplessis-Marcotte, F., Provencher, J., Marin, M.F., 2024. Understanding sex differences in extinction retention: pre-extinction stress and sex hormone status. Psychoneuroendocrinology 169, 107161. https://doi.org/10.1016/j.psyneuen.2024.107161.

Ponce, P., Del Arco, A., Loprinzi, P., 2019. Physical activity versus psychological stress: effects on salivary cortisol and working memory performance. Medicina 55, 119. https://doi.org/10.3390/medicina55050119.

Powers, M.B., Medina, J.L., Burns, S., Kauffman, B.Y., Monfils, M., Asmundson, G.J.G., Diamond, A., McIntyre, C., Smits, J.A.J., 2015. Exercise augmentation of exposure therapy for PTSD: rationale and pilot efficacy data. Cogn. Behav. Ther. 44, 314–327. https://doi.org/10.1080/16506073.2015.1012740.

de Quervain, D., Schwabe, L., Roozendaal, B., 2017. Stress, glucocorticoids and memory: implications for treating fear-related disorders. Nat. Rev. Neurosci. 18, 7–19. https://doi.org/10.1038/nrn.2016.155.

de Quervain, D.J.-F., Wolf, O.T., Roozendaal, B., 2019. Glucocorticoid-induced enhancement of extinction—from animal models to clinical trials. Psychopharmacology 236, 183–199. https://doi.org/10.1007/s00213-018-5116-0.

Rose, J., Otto, T., Dittrich, L., 2008. The Biopsychology-Toolbox: a free, open-source Matlab- toolbox for the control of behavioral experiments. J. Neurosci. Methods 175, 104–107. https://doi.org/10.1016/j.jneumeth.2008.08.006.

Schwabe, L., Hermans, E.J., Joëls, M., Roozendaal, B., 2022. Mechanisms of memory under stress. Neuron 110, 1450–1467. https://doi.org/10.1016/j.neuron.2022.02.020.

Skoluda, N., Strahler, J., Schlotz, W., Niederberger, L., Marques, S., Fischer, S., Thoma, M.V., Spoerri, C., Ehlert, U., Nater, U.M., 2015. Intra-individual psychological and physiological responses to acute laboratory stressors of different intensity. Psychoneuroendocrinology 51, 227–236. https://doi.org/10.1016/j.psyneuen.2014.10.002.

Tanner, M.K., Hake, H.S., Bouchet, C.A., Greenwood, B.N., 2018. Running from fear: exercise modulation of fear extinction. Neurobiol. Learn. Mem. 151, 28–34. https://doi.org/10.1016/j.nlm.2018.03.021.

Tarvainen, M.P., Niskanen, J.P., Lipponen, J.A., Ranta-Aho, P.O., Karjalainen, P.A., 2014. Kubios HRV—Heart rate variability analysis software. Comput. Methods Prog. Biomed. 113, 210–220. https://doi.org/10.1016/j.%20cmpb.2013.07.024.

Vervliet, B., Baeyens, F., van den Bergh, O., Hermans, D., 2013. Extinction, generalization, and return of fear: a critical review of renewal research in humans. Biol. Psychol. 92, 51–58. https://doi.org/10.1016/j.biopsycho.2012.01.006.

Voorendonk, E.M., Sanches, S.A., Mojet, M., de Jongh, A., van Minnen, A., 2021. Combining a single session of prolonged exposure with physical activity in patients with PTSD: the effect of sequence. Ment. Health Phys. Act. 21, 100417. https://doi.org/10.1016/j.mhpa.2021.100417.