Cognitive Therapy and Research (2024) 48:704–719
https://doi.org/10.1007/s10608-023-10463-9

ORIGINAL ARTICLE

Endogenous Cortisol Levels, Sleep or Vigilance: Which Factors Contribute to Better Exposure Therapy Outcomes in the Morning?

Marie Roxanne Sopp[1] · Sarah K. Schäfer[1,2] · Tanja Michael[1] · Monika Equit[1] · Diana S. Ferreira de Sá[1] · Johanna Lass-Hennemann[1]

Accepted: 29 December 2023 / Published online: 20 February 2024

Abstract

Background Research suggests that exposure therapy delivered in the morning is more successful than delivered in the evening, which is often explained by higher diurnal endogenous cortisol levels. However, this "morning exposure effect" might also be explained by other factors such as sleep or vigilance.

Methods The current study aimed to disentangle these effects by assessing the impact of video-based exposure therapy delivered in the morning or in the evening, whilst considering pre-exposure sleep quality, vigilance, and cortisol levels. To this end, 80 snake fearful individuals were randomly assigned to receive exposure treatment in the morning or evening.

Results Contrary to previous findings, groups did not differ in their pre-post and post-follow up decrease of snake anxiety. However, higher vigilance was found to be associated with a greater pre-post and post-follow-up decrease in snake anxiety. Moreover, pre-exposure sleep efficiency moderated the post-follow-up decrease in snake anxiety across groups: In individuals with high pre-exposure sleep efficiency, those receiving exposure in the morning were estimated to show a stronger decrease in snake anxiety than those receiving exposure in the evening. The opposite pattern was found in individuals with low pre-exposure sleep efficiency.

Conclusions The results of this study illustrate that diurnal effects on exposure therapy might be more complex than previously assumed.

Trial Registration The study was prospectively preregistered at the German Clinical Trial Register (https://drks.de/search/en/trial/DRKS00016183).

Keywords Sleep · Exposure therapy · Cortisol · Vigilance · Exposure enhancement

Introduction

Although exposure therapy is considered the gold standard in the treatment of anxiety and stressor-related disorders, there is a substantial proportion of patients, who do not profit from treatment, and relapse is a common problem (Michael et al., 2019). Consequently, research on novel strategies that can enhance the success of exposure therapy has received much interest in the last years.

One line of research focuses on possible pharmacological agents as boosters of exposure therapy. These new therapeutic approaches are based on the idea that the pharmacological agents enhance the learning processes underlying exposure therapy. One pharmacological agent that has been proposed to enhance the success of exposure therapy is cortisol (Bentz et al., 2010). Several studies have shown that exogenous administration of cortisol prior to exposure therapy enhances therapeutic gains of exposure therapy in patients with height phobia (de Quervain et al., 2011), spider phobia (Soravia et al., 2014; but see Raeder et al., 2019 for contrary findings), social phobia (Soravia et al., 2006), posttraumatic stress disorder (Yehuda et al., 2015), and alcohol use disorder (Soravia et al., 2021).

Marie Roxanne Sopp and Sarah K. Schäfer shares first authorship.

✉ Johanna Lass-Hennemann
lass-hennemann@mx.uni-saarland.de

[1] Division of Clinical Psychology and Psychotherapy, Department of Psychology, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany

[2] Division of Clinical Child and Adolescent Psychology and Psychotherapy, Technische Universität Braunschweig, Brunswick, Germany

Cortisol is secreted in response to stress and can be administered exogenously (Bentz et al., 2010). However, cortisol secretion also shows natural fluctuation across the day, with a peak in the morning and low levels during the evening and night. Thus, another line of research has focused on the effects of endogenous cortisol levels on the success of exposure therapy. Pace-Schott et al. (2013) found that extinction learning (as an analog for exposure therapy) was more successful in the morning (when endogenous cortisol levels are high) than in the evening. Lass-Hennemann and Michael (2014) transferred these findings to a clinical sample: They found that spider phobic patients who were treated in the morning showed better treatment outcomes than patients who were treated in the evening. In line with these findings, Meuret et al. (2015) showed that higher cortisol levels during exposure sessions conducted at different daytimes were associated with enhanced clinical improvement in a multi-session in-vivo exposure protocol for panic disorder and agoraphobia. In a further study in patients with panic disorder, the same authors were able to show that cortisol mediated the effect of time of day on treatment outcome, providing a link between earlier exposure sessions and greater clinical improvement (Meuret et al., 2016).

These findings gave rise to a simple clinical hypothesis: Exposure sessions conducted in the morning are more effective than exposure sessions conducted in the afternoon or evening ("morning exposure effect"), because the high endogenous cortisol in the morning will enhance memory consolidation for new non-anxiety related material and thus lead to better clinical outcomes. Even though there is some evidence linking high endogenous cortisol levels to higher therapeutic gains during exposure therapy (Meuret et al., 2016), the research on the "morning exposure effect" is just at the very beginning and there are also other potential mechanisms that may account for better therapy outcomes in the morning.

In contrast to accounts linking diurnal variations in cortisol levels to daytime effects on exposure therapy, there is an alternative assumption, positing that the temporal proximity to awakening is the critical factor boosting exposure effects in the morning (Nissen et al., 2017). This assumption is based on the synaptic homeostasis hypothesis (Tononi & Cirelli, 2006). According to this hypothesis, new learning experiences result in a continuous increase of synaptic connections. Without downregulation, such a continuous increase would lead to a saturation of synaptic networks, preventing subsequent learning. Hence, to avoid saturation, synaptic connections are downregulated during sleep, a process that is referred to as synaptic downscaling. Based on this hypothesis, it has been suggested that the capacity for learning is highest immediately after awakening and decreases continuously throughout the day. Empirical findings support this assumption (Kaida et al., 2015; Mander et al., 2011) and confirm that preceding sleep enhances extinction learning (Straus et al., 2017). Moreover, Zuj et al. (2016) found that extinction learning in patients with more severe post-traumatic stress disorder is less successful after prolonged wakefulness than immediately after awakening.

Another factor that has received little attention so far but that may account for superior effects of exposure therapy during the morning is vigilance. Cognitive psychologists and neuroscientists define vigilance as the ability to sustain attention to a task for a period of time (Parasuraman, 2000). Vigilance has been shown to be higher in the morning and to decline over the course of the day (Harrison et al., 2007; Riley et al., 2017). Previous research indicates that reduced vigilance co-occurs with reduced learning and emotion processing (Helton & Russell, 2011; Schwarz et al., 2013; Wang et al., 2012). Psychotherapy in general, but especially exposure therapy relies on emotion processing and learning (Lass-Hennemann et al., 2018). Higher vigilance levels in the morning as compared to the evening might thus—fully or partially—account for greater therapeutic gains in the morning.

Taken together, several factors may contribute to the "morning exposure effect": High endogenous cortisol levels, sleep, and vigilance. However, up to date there are no studies assessing the different factors in one study to disentangle the importance of the proposed mechanisms. Moreover, despite compelling evidence, a study employing virtual exposure exercises for spider phobia in which exposures and testing of spider fear were performed both in the morning and evening (as controls for circadian vs. sleep effects) did not find superior exposure outcomes in the morning (Pace-Schott et al., 2012). Furthermore, a recent study found that higher cortisol levels during exposure were linked to reduced—rather than enhanced—symptom decline in patients with social anxiety disorder (Kuhlman et al., 2020). In addition, a meta-analysis was not able to confirm a significant association between cortisol levels during exposure and symptom improvement questioning the central role of cortisol in the "morning exposure effect" (Fischer & Cleare, 2017).

In light of these ambiguities, we conducted an experimental study to further characterize the impact of daytime on exposure therapy and the involvement of cortisol, sleep, and vigilance. High snake fearful individuals were randomly assigned to receive one session of video-exposure treatment in the morning or in the evening. Symptoms of snake phobia were assessed prior to treatment, after treatment and at two follow-ups. Furthermore, we assessed time since awakening, sleep quality (i.e., total sleep time and sleep efficiency), and vigilance prior to exposure as well as endogenous cortisol levels prior to and during exposure. We expected exposure in the morning to be more successful than exposure in the evening. Moreover, we explored the extent to which these effects are accounted for by endogenous cortisol levels, time

since awakening, sleep quality and vigilance. Finally, we conducted explorative analyses to assess whether the group effects are moderated by this set of variables.

Methods

Participant Characteristics

High snake anxious participants were recruited via social media and flyers posted at Saarland University, Germany. Participation was restricted to healthy, non-smoking participants with a body mass index between 20 and 25 and an age range between 18 and 40 years. In order to minimize menstrual cycle effects on learning mechanisms, free-cycling women participated in the follicular phase of their menstrual cycle (on the first 10 days of after the start of their menstruation). 46.6% ($n = 28$) of the female participants were free cycling, the remaining 53.3% ($n = 32$) reported to take oral contraceptives. Furthermore, people with a recent history of systemic or oral cortisol therapy and who were pregnant, or lactating were excluded from participation. We also excluded people with any Axis I disorder (other than snake phobia), severe acute or chronic disease (e.g., lung or cardiovascular diseases) and current pharmacological treatment (except of oral contraceptives and L-thyroxine for hypothyroidism) or psychotherapy. We also asked participants to refrain from physical exercise, alcohol, and caffeinated drinks (factors known to influence cortisol levels) within 3 h prior to exposure session. Eighty participants were included in the study. Of those, 9 participants had to be completely excluded from analysis: eight due to insufficient data ($n = 8$) and one participant had to be excluded due to extremely high (possibly artificially increased) endogenous cortisol levels ($n = 1$).

Sample Size Calculation

The required sample size was determined a-priori using G*Power version 3.1 (Faul et al., 2009). Based on previous data of our group on a similar video-based exposure procedure (Ihmig et al., 2020; Schäfer et al., 2018), we aimed at detecting an at least small effect ($d = 0.25$) with a power of 0.80 and alpha level set to 0.05. Building on previous data of our group (Ihmig et al., 2020; Lass-Hennemann & Michael, 2014), we estimated correlations between repeated measurements at $r = 0.60$, which resulted in a sample size of 72. As we expected a dropout rate of 10%, we aimed to recruit 80 participants (i.e., 40 for both experimental groups).

Procedure and Design

The study was preregistered at the German Clinical Trials Register https://drks.de/search/en/trial/DRKS00016183. The study took place at the Division of Clinical Psychology and Psychotherapy at Saarland University, Germany. Participation included a telephone screening, an online questionnaire and four appointments at our laboratory: a pretreatment interview clarifying study eligibility and assessing symptoms (behavioral approach test), a 1-h video treatment session, an assessment one week after treatment (one-week follow up), and a follow-up assessment 4 weeks after treatment (four-week follow up). Figure 1 presents a schematic illustration of the study procedure. Participants were randomized by an independent person to the "exposure in the morning" or the "exposure in the evening" group.

Screening Phase

Telephone Screening

Participants who volunteered to participate in the study were contacted via telephone. During the telephone screening, participants were informed about the study and important inclusion and exclusion criteria were checked. Fear of snakes was assessed with the fear of snakes screening (Schlangenangst-Screening [SCANS], Reinecke et al., 2009) which consists of 4 questions (fear of snakes, physiological fear reactions to snakes, avoidance of snakes and clinical distress) that were rated on a seven-point scale from 0 to 6. Participants who scored at least 10 points on the fear of snakes screening were invited for further diagnostic assessments.

Online Questionnaires

After the telephone screening potential participants received a link to the online questionnaire via email. The online questionnaire consisted of a battery of questionnaires assessing fear of snakes (Snake Questionnaire [SNAQ]) and general psychopathological symptoms (Patient Health Questionnaire [PHQ-D]). These questionnaires were administered to ensure that participants had a substantial fear of snakes and no other clinically relevant psychopathological symptoms. Furthermore, we assessed trait anxiety (State-Trait Anxiety Inventory [STAI-T]) and anxiety sensitivity (Anxiety Sensitivity Index [ASI-3]), depressive symptoms (Beck's Depression Inventory II [BDI-II]), chronotype (reduced Morningness-Eveningness-Questionnaire [rMEQ]), and sleep quality (Pittsburgh Sleep Quality Index [PSQI]) to control for a-priori between-group differences and as potential moderators.

Pretreatment Interview

Participants, who met the cut off criteria in the online questionnaires, were invited for the pretreatment interview. To finally identify high snake anxious individuals and exclude any other Axis I disorder we conducted a structured

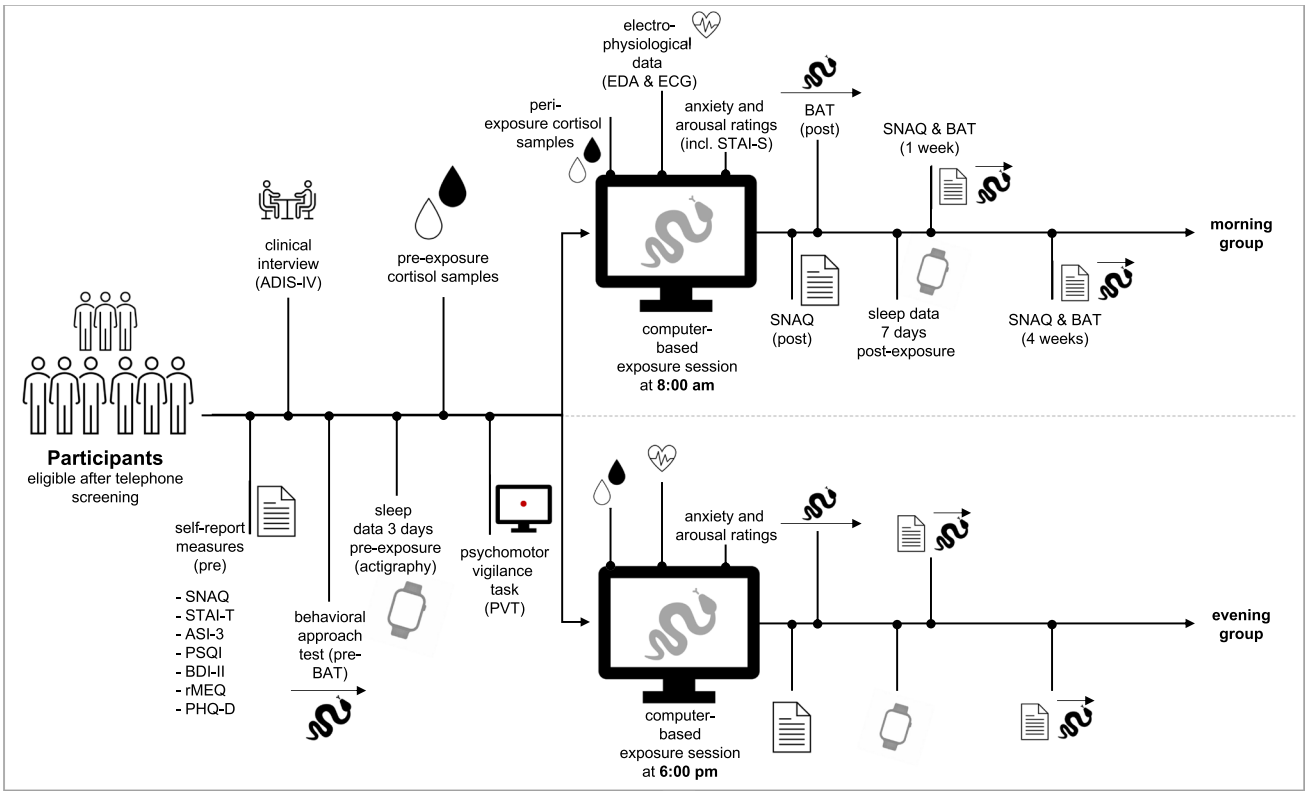

Fig. 1 Schematic illustration of the study procedure. *Note* Assessments relevant to our analyses took place at four larger time points: pre-exposure, post-exposure, at 1- and 4-week follow-up. Moreover, we have peri-exposure data available for anxiety and arousal ratings, electrophysiological data as well as cortisol. *ADIS-IV* Anxiety Disorders Interview Schedule-IV, *SNAQ* Snake Questionnaire, *STAI-T/S* State-Trait Anxiety Inventory – trait/state version, *ASI* Anxiety Sensitivity Index 3, *PSQI* Pittsburgh Sleep Quality Index, *BDI-II* Beck Depression Inventory II, *rMEQ* reduced Morningness-Eveningness-Questionnaire, *PHQ-D* Patient Health Questionnaire, *BAT* Behavioral Approach Test, *PVT* Psychomotor Vigilance Task, *EDA* electrodermal activity, *ECG* electrocardiography

interview (modified German version of the Anxiety Disorders Interview Schedule-IV [ADIS-IV, German version: MINI-DIPS], Margraf, 1994). In order to be classified as high snake anxious, participants had to fulfill all diagnostic criteria for specific phobia of the animal type (relating to snakes) except for the "clinical distress" criteria, in detail they had to reach a fear score ≥ 4 and an avoidance score ≥ 3 in the section 'Specific Phobia' in the ADIS-IV. After completing the clinical interview, the behavioral approach test (BAT) was conducted. Participants were only included if they reached a BAT score of score ≤ 4. At the end of the pre-treatment interview participants received instructions for the assessment of saliva samples and for the use of the actigraph (see Materials and Measures).

Treatment Session

Participants arrived for the one-hour video-based exposure session at 08.00 a.m. (morning exposure group) or at 06.00 p.m. (evening exposure group). After participants arrived, they handed in the saliva samples (from the cortisol awakening response). Then they were seated in a closed laboratory room with a 27" LCD monitor (60 Hz refresh rate; viewing distance: approx. 65 cm) in front of them and were prepared for physiological measurements (electrodermal activity, electrocardiogram). The experimenter was seated in a closed room next to the participants, with participants being aware of the experimenter being present all the time. Participants first underwent the Psychomotor Vigilance Task (PVT). After the PVT, participants were asked to fill out the state scale of the State-Trait Anxiety Inventory (STAI-S). Afterwards, the one-hour video-based exposure session was conducted. During the experiment, the experimenter checked on participants' compliance with the exposure rationale via camera (i.e., visual focus on the screen). Participants provided saliva samples every 15 min during the exposure session and subjective stress and arousal ratings were assessed. Experimenters prompted each saliva sample collection and checked whether the samples were collected correctly via camera. After the exposure session participants filled out the STAI-S again. Electrodes for physiological measurements were detached, saliva samples were stored, and the

BAT was conducted again. At the end of the exposure session, appointments for the two follow-up assessments were scheduled.

Follow-Up Assessments

The follow-up assessments took place between 03.00 p.m. and 04.00 p.m. for each participant. The first follow-up was conducted one week after the therapy session. Upon arrival at the laboratory, participants were asked to return the actigraphs. Subsequently, participants' behavioral and subjectively experienced fear was assessed with the BAT and the SNAQ. The second follow-up assessment was conducted four weeks after the exposure session also including the BAT and the SNAQ.

Materials and Measures

Behavioral Approach Test (BAT)

To test the participants' fear and avoidance of living snakes, we used the well-established BAT. The BAT (adapted from Lass-Hennemann & Michael, 2014) involved the following procedure: Standing in front of a closed room that contained a snake, each participant was asked to open the door and approach a living corn snake of approximately one-meter length, which was placed in a sealed transparent plastic container on a table at the far end of the room (approximately 6 m from the door). Next, if possible, the patient was to remove the lid, insert a hand, and touch the snake for at least 20 s. When the participant had reached for and touched the snake, or when the participant decided to stop the approach, the remaining distance between the participant and the snake was noted. In detail, the BAT comprised 13 steps: 0—refuses to enter the test room,1—stops 5 m from the container, 2—stops 4 m from the container, 3—stops 3 m from the container, 4—stops 2 m from the container, 5—stops 1 m from the container, 6—stops close to the table with the container, 7—touches the container, 8—removes the lid, 9—puts a hand in the container, 10—touches the snake with one forefinger (in the plastic container), 11—touches the snake for less than 20 s (while the experimenter holds the snake in his/her hands), and 12—touches the snake for at least 20 s (while the experimenter holds the snake in his/her hands).

Psychomotor Vigilance Task (PVT)

Participants' vigilance prior to the treatment session was assessed with the PVT. The PVT is a well-validated and widely used sustained-attention reaction-time based task. During the 5-min PVT, participants' sustained or vigilant attention is measured by recording reaction times to visual stimuli that occur at random inter-stimulus-intervals ranging from 2000 to 10,000 ms (Roach et al., 2006). Participants received the instructions to react as quickly as possible to the onset of the stimulus (red circle presented on black background). In order to maximize their performance, participants received feedback on their response time after each stimulus. The mean reaction time across all trials was calculated and used as outcome measure (Groch et al., 2013).

Video Based Exposure Session

At the beginning of the treatment session, participants were introduced to the exposure procedure and rationale, which was based on the principles of the one-session exposure treatment by Öst (1989). Its key aspects are controlled exposure to the fear-provoking stimulus and changes of fearful cognitions. Every exposure trial starts with a question on the content of the video clip (e.g., what color has the snake?) in order to set a cognitive focus for the following clip. The 20 90-s-clips, all taken from TV documentaries, show detailed shots of different snakes. After every clip, the question asked prior to the clip is presented again and participants are asked to choose the right answer by choosing between four multiple-choice answers. In case of a correct answer, the participant was praised. In case of an incorrect response, an instruction reminding the participant to concentrate on the videos in order to reduce their fear was presented. After ten clips, the questions focused on positive emotional features instead of cognitive aspects to change the participant's attitude towards snakes. In detail, participants were instructed to identify positive aspects of the snakes in each video. Following each video, participants were asked to key in the identified positive aspects. Participants were constantly complimented on identifying positive facets. After every fourth video clip in the exposure session, participants are asked to rate their subjective fear and arousal levels on 10-point scales from "1 = not at all" to "10 = strongly".

Actigraphs

Actigraphs were worn for a total of 10 days (3 nights prior to the exposure session and 7 days after the exposure session). Total sleep time (TST) and sleep efficiency (SE) during the night preceding exposure were assessed using actigraphy. Actigraphy (wGT3X-BT; Aggio et al., 2015) was used to detect the amount of participants' movements with a built-in motion sensor. ActiLife software and manufacturer algorithms for detecting sleep based on 60-s epochs were used to generate summary statistics for participants' sleep. Specifically, whether an epoch was scored as 'wake' or 'sleep' was determined by comparing activity counts for the epoch in question and those immediately surrounding it

to a threshold value using the Cole–Kripke algorithm (Cole et al., 1992). Pre-defined manual sleep scoring rules were used to adjust automatic scoring in light of participants' sleep timings, which they recorded in a sleep log (ActiGraph accelerometers do not have event markers). Specifically, the participants reported their bedtimes (lights off) just before going to sleep, and then in the morning their final awakening time and the time they got up (lights on). This information was manually entered in the ActiLife software to adjust the autoscored sleep timings, e.g., distinguishing sedentary wake-time behavior from the periods asleep. Total sleep time was determined by summing the epochs that were scored as sleep. SE was determined by dividing the amount of time spent asleep (in min) by the total amount of time in bed (in min). Time in bed was determined using the sleep log. Data from the night preceding the exposure session (pre-exposure sleep) and the mean of the three nights preceding the exposure session (baseline sleep) were used for all subsequent analyses. Data of the seven nights following the exposure session will be reported elsewhere.

Endogenous Cortisol

To assess the basal cortisol reaction (cortisol awakening response, CAR), participants provided three saliva samples (awake, + 30 min, + 45 min) during the morning of the treatment session. Participants in the evening group provided one more sample at 03.00 p.m. Furthermore, the cortisol response during the exposure session was assessed with seven saliva samples (before the beginning of exposure, 15 min after start of exposure sessions, + 30 min, + 45 min, + 60 min, + 75 min, + 90 min).

Saliva samples were collected using Salivette tubes (Sarstedt AG). After thawing the saliva samples for biochemical analysis, the fraction of free cortisol in saliva was determined using a time resolved immunoassay with fluorometric detection as described in detail elsewhere (Dressendörfer et al., 1992). The area under the curve with respect to ground was calculated during awakening ($AUC_{G\text{-}CAR}$) and exposure ($AUC_{G\text{-}EXP}$), with the $AUC_{G\text{-}CAR}$ reflecting the total cortisol output in response to awakening during the first hour after awakening (Pruessner et al., 2003).

Physiological Measurements

Throughout the one-hour video-based exposure session electrodermal activity (EDA) and electrocardiography (ECG) were acquired. Due to technical errors and bad signal quality the majority of data was not analysable. Because there was no central hypothesis regarding physiological data, we decided to discard the physiological data from analysis.

Snake Questionnaire (SNAQ)

The Snake Questionnaire is one of the most widely used measures to assess specific fears of snakes (Klieger, 1987; German version: Hamm, 2006). It consists of 30 items with positive (e.g., I enjoy watching snakes in zoos) and negative statements about snakes (e.g., I am terrified by the thought of touching a harmless snake) which are answered on a two-point scale (agree/disagree). The sum scores of the SNAQ reach from 0 to 30, while higher scores indicate more fear of snakes.

STAI-T and STAI-S

The German version of the State-Trait Anxiety Inventory (Spielberger, 1970; German version: Laux et al., 1981) was used to assess trait anxiety (STAI-T) as well as short-term changes in state anxiety before and after exposure (STAI-S). Both STAI scales are brief self-report questionnaires consisting of 20 items each. Participants are asked to rate each item on a 4-point Likert scale. The sum scores of both scales range from 20 to 80, while lower scores are indicators of low (state or trait) anxiety and higher scores indicate high (state or trait) anxiety.

ASI-3

The Anxiety Sensitivity Scale is a self-report measure designed to assess fearful cognitions about physiological anxiety symptoms (Taylor et al., 2007; German version: Kemper et al., 2009). The 16-item scale is answered on a 5-point Likert scale. The sum scores of the ASI-3 range between 0 and 72, while higher scores indicate higher anxiety sensitivity. The ASI-3 is administered to control for a-priori between-group differences.

PSQI

The Pittsburgh Sleep Quality Index is a self-report measure, which assesses sleep quality over the last four weeks (Buysse et al., 1989; German version: Riemann & Backhaus, 1996). Nineteen individual items generate seven "component" scores: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction. Each item is rated on a 4-point Likert scale. The global score is the sum of the component scores and can range between 0 (very good overall sleep quality) to 21 (very poor overall sleep quality). The PSQI is administered to assess sleep quality and to control for a-priori

differences in sleep quality between the two experimental groups.

BDI-II

The BDI-II is a 21-item self-report measure designed to assess the severity of depressive symptoms over the last two weeks with scores ranging from 0 to 63 (German version: Hautzinger et al., 2006). Higher total scores indicate more severe depressive symptoms. Scores larger than 17 are considered to be clinically relevant depressive symptoms and participants with scores higher than 17 were excluded from the study. Thus, the BDI-II was administered to exclude participants with clinically relevant depressive symptoms and to control for a-priori differences between the two experimental groups.

rMEQ

The reduced Morningness-Eveningness-Questionnaire (German version: Randler, 2013) assesses the degree to which persons are active and alert at certain times of day. The 5-item questionnaire asks for preferences in sleep and waking times, and subjective "peak" times at which respondents feel their best. To obtain a global score, each item is totaled, with the sum score ranging between 5 and 24.

Patient Health Questionnaire D

The Patient Health Questionnaire D (German version: Löwe et al., 2002) is a short and economic instrument for assessing symptoms of mental health disorders. Its newest version relies on the DSM-IV-TR criteria and has good psychometric properties (Löwe et al., 2004).

Data Analyses

Potential baseline differences as well as differences in cortisol levels between groups were tested by means of unpaired *t*-tests. Changes of arousal/anxiety ratings during the exposure session were examined by means of mixed ANOVAs.

In order to test our hypotheses, a series of multilevel models was fitted separately for BAT and SNAQ scores. We conducted two sets of analyses: A pre-post analysis including all assessments (pre-exposure, post-exposure, one-week follow-up, four-week follow-up) was focused on pre-post symptom change (time centered at pre-exposure). A post-follow-up analysis including only post-exposure, one-week follow-up, and four-week follow-up was run separately, focusing on the maintenance of symptom change from post-exposure to the follow-up assessments (time centered at post-exposure). In a first step, we constructed a baseline model, comprising the random and fixed effect of Time. Subsequently, we investigated group effects by including Group and the interaction between Time and Group as fixed effects. In order to test potential confounding effects of sex, we repeated these analyses including only female participants. The subsample of male participants was too limited ($N = 11$) to allow for separate analyses. We report descriptive statistics of outcome measures for female and male participants in the Supplementary Material.

In addition, we aimed to test the predictive effects of cortisol levels during exposure, psychomotor vigilance, time since awakening (TSW), and sleep quality for symptom change. To this end, we added the respective predictor and the Time × Predictor interaction to the baseline model and tested the improvement of model fit (χ^2 difference test). For all models that yielded a significant improvement of model fit beyond the baseline model, we planned to evaluate fit indices of models including the respective predictor and the model including Group to conclude which factors bears the strongest predictive value. Given that none of the analyses yielded a significant effect of Group, such comparisons were not necessary. Finally, we aimed to investigate moderator effects. Given that the Group factor was highly correlated with TSW and cortisol levels during exposure, these variables were excluded from moderator analyses to avoid issues arising from multicollinearity. Mean arousal level during the exposure session was investigated as an additional moderator since baseline analysis yielded unexpected group differences. All potential moderator variables were found to approximate a normal distribution, KS Test $p > 0.08$. In order to investigate potential moderator effects of sleep quality and psychomotor vigilance, we evaluated whether a model including the interaction between the respective moderator, Group and Time improved model fit beyond the baseline model. All Level-2 predictors were grand-mean centered (Kreft et al., 1995). The Level-1 predictor Time was centered at baseline (Singer & Willett, 2014). All multilevel models were fit using maximum likelihood estimation with the *lme4* package (Bates, 2010) in *R* (Team, 2022). Significant interactions were probed using simple slopes techniques implemented in the *reghelper* package (Hughes et al., 2022). Slopes were estimated at ± 2 *SD* above/below the mean of the respective moderator variable. The alpha level was set to 0.05 for all analyses.

BAT scores were missing for 11 participants at post-assessment or follow-up assessments. One participant showed a mean PVT reaction time over 3 interquartile ranges above the upper quartile and was thus excluded from all analyses including the PVT. Due to data loss, actigraphy data of 34 participants (Evening group: $n = 19$, Morning group: $n = 15$) were not available for analysis. TSW was not documented by three participants. As a result, degrees of freedom vary across analyses.

Results

Sample Characteristics and Baseline Differences

As summarized in Table 1, groups did not differ in age, sex, Cortisol Awakening Response ($AUC_{G\text{-}CAR}$) nor on any symptom or trait measures. They similarly did not differ in snake anxiety at baseline (BAT and SNAQ scores). As anticipated, they differed significantly with respect to TSW and cortisol levels during exposure ($AUC_{G\text{-}EXP}$) with the morning group showing higher $AUC_{G\text{-}EXP}$ scores (see Fig. 2) and a shorter TSW. They further differed in psychomotor vigilance, with the morning group unexpectedly showing higher scores, indicating reduced vigilance as opposed to the evening group. Finally, they had comparable baseline and pre-exposure TST and SE.

Arousal and Anxiety Changes During Exposure

As anticipated, state anxiety levels significantly increased from pre- to post-intervention, $F(1, 67) = 29.92$, $p < 0.001$, see Fig. 3. No significant effect of Group, $F(1, 67) = 0.48$, $p = 0.491$, or Group × Time, $F(1, 67) = 0.04$, $p = 0.846$, was evident. Anxiety levels increased during exposure, indicating successful activation of fear memory, followed by a decline, which may indicate successful extinction, $F(5, 32) = 4.85$, $p = 0.002$. Groups were not found to differ in

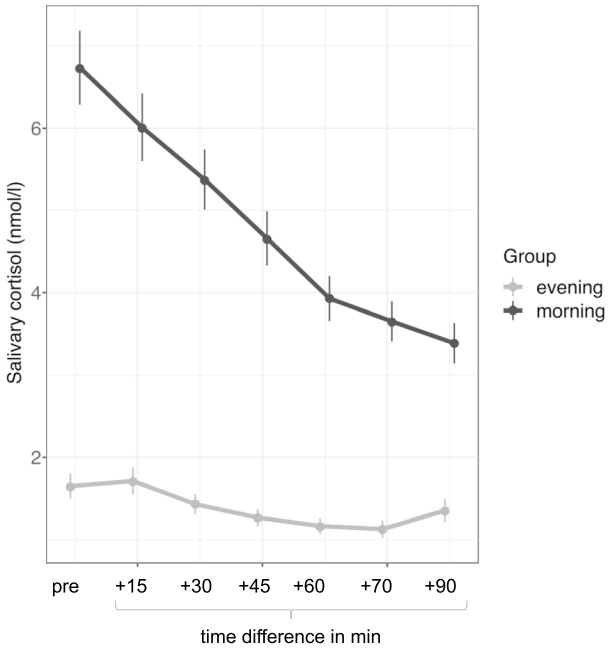

Fig. 2 Salivary cortisol in nanomole per liter measured every 15 min during the exposure session. *Note* Time refers to time of measurement with assessment taking place at prior to exposure (pre) and every 15 min (+15 min, +30 min, +45 min, +60 min, +75 min, +90 min) during the exposure session. Error bars represent the standard error of the mean

Table 1 Demographic and psychometric characteristics of the morning and evening group

Variables (range and units of measurement)	Morning group ($n = 36$)		Evening group ($n = 35$)		Comparison
	M	SD	M	SD	
Sex	♀31/♂5		♀29/♂6		$\chi^2 = 0.14$, $p = 0.705$
Chronotypes (based on rMEQ$^+$)	E: 6/N: 27/M: 3		E: 9/N: 20/M: 6		$\chi^2 = 2.63$, $p = 0.269$
Age (years)	22.03	3.35	22.57	3.60	$t = -0.66$, $p = 0.512$
Baseline SNAQ (0–30)	19.56	4.21	20.91	3.45	$t = -1.49$, $p = 0.142$
Baseline BAT (0–12)	2.19	1.55	2.51	1.40	$t = -0.91$, $p = 0.365$
STAI-T (20–80)	34.08	6.14	37.31	8.10	$t = -1.90$, $p = 0.062$
PSQI (0–21)	4.11	2.25	4.63	2.51	$t = -0.90$, $p = 0.370$
BDI-II (0–36)	3.94	3.82	4.60	3.64	$t = -0.74$, $p = 0.462$
rMEQ (5–24)	13.83	2.51	14.14	3.83	$t = -0.40$, $p = 0.688$
ASI-3 (0–72)	16.94	10.72	20.29	13.92	$t = -1.14$, $p = 0.260$
$AUC_{G\text{-}CAR}$ (nmol/L*min)	500.69	344.00	506.11	355.45	$t = -0.07$, $p = 0.948$
TSW (min)	98.32	29.31	690.06	30.90	$t = -81.02$, $p < 0.001$
$AUC_{G\text{-}EXP}$ (nmol/L*min)	430.29	346.01	117.35	115.38	$t = 5.02$, $p < 0.001$
Baseline TST (min)	425.03	58.03	396.00	43.53	$t = 1.64$, $p = 0.111$
Baseline SE (%)	88.07	6.35	85.64	6.27	$t = 1.14$, $p = 0.263$
Pre-exposure TST (min)	347.43	80.04	350.13	70.27	$t = -0.11$, $p = 0.915$
Pre-exposure SE (%)	87.08	7.35	85.38	7.80	$t = 0.68$, $p = 0.502$
PVT (ms)	317.94	54.76	291.56	30.40	$t = 2.50$, $p = 0.015$

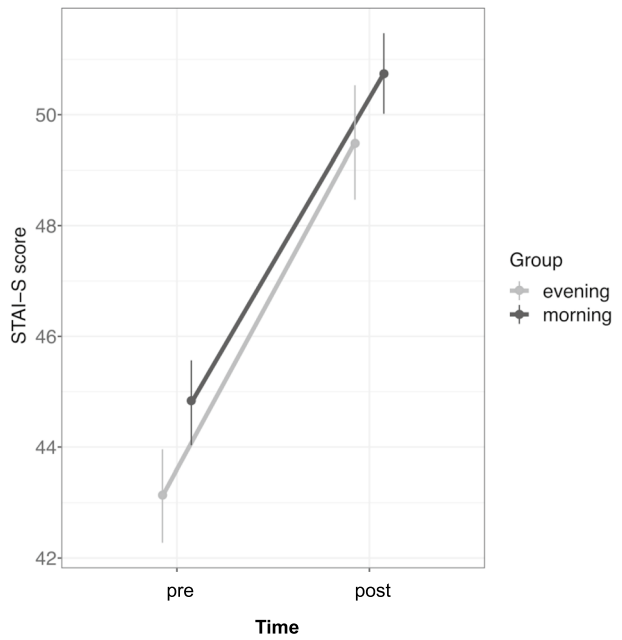

Fig. 3 Change of STAI-S scores from pre- to post-exposure in the morning and evening group. *STAI-S* state-trait anxiety inventory—state version. Error bars represent the standard error of the mean. *The STAI-S has a range* from 20 to 80, while lower scores are indicators of lower state anxiety

overall level, $F(5, 32) = 3.11$, $p = 0.087$, or change across time, $F(5, 32) = 1.15$, $p = 0.355$, see Fig. 4a. Arousal levels also increased, indicating successful activation of fear memory, followed by a decline, potentially indicating successful habituation/extinction, $F(5, 32) = 4.05$, $p = 0.006$. The morning group showed slightly higher scores than the evening group, $F(5, 32) = 7.14$, $p = 0.011$, however both groups showed a similar change across time, $F(5, 32) = 0.90$, $p = 0.495$, see Fig. 4b.

E = Evening type, N = Neither type, M = Morning type, + = Cut-off values were taken from Randler (2013), SNAQ = Snake Anxiety Questionnaire (SNAQ, German version), BAT = Behavioural Approach Test, STAI-T = State-trait anxiety inventory – trait version, PSQI = Pittsburgh Sleep Quality Index, BDI-II = Beck Depression Inventory II, rMEQ = reduced Morningness–Eveningness Questionnaire, ASI-3 = Anxiety Sensitivity Index 3, AUC_G = Area under the curve with respect to ground, CAR = Cortisol Awakening Response, EXP = Cortisol during exposure, TSW = Time since awakening, TST = Total sleep time, SE = Sleep efficiency, PVT = Psychomotor vigilance task

Pre-post Changes in Symptoms

Note that an overview of intercepts and slopes as well as the estimated variance accounted for by each significant model

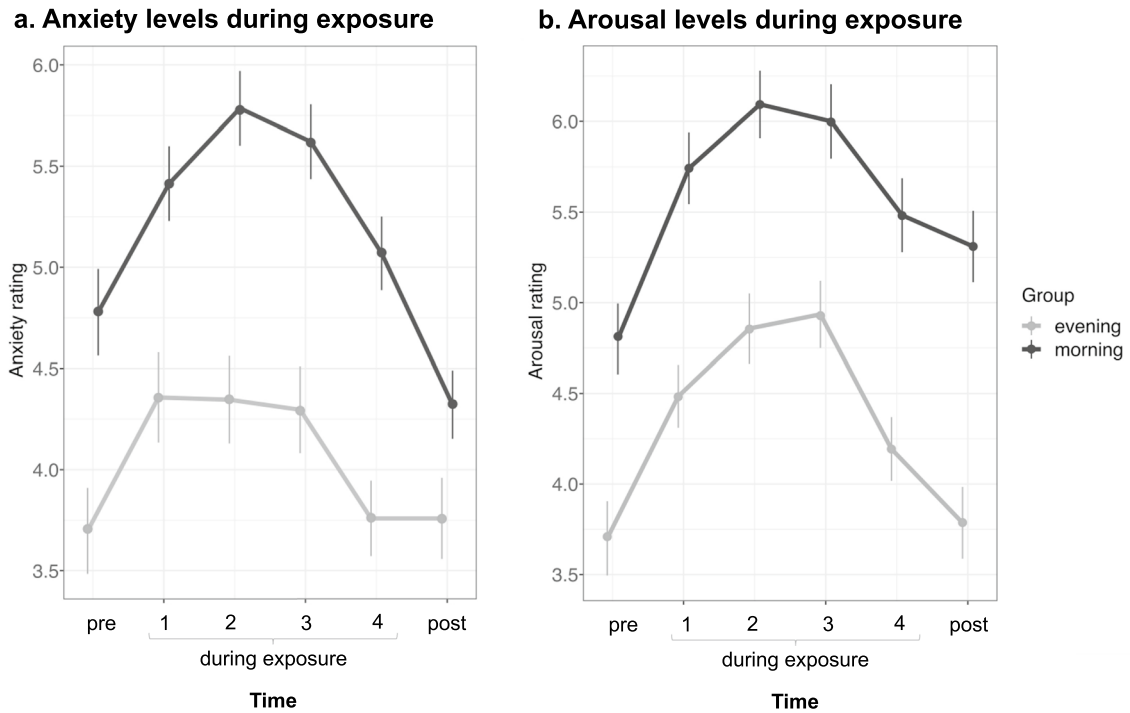

Fig. 4 Change of anxiety and arousal ratings during exposure in the morning and evening group. **a** represents anxiety ratings, **b** presents arousal ratings. Anxiety and arousal were rated on 10-point scales from "1 = not at all" to "10 = strongly". Time refers to time of measurement with assessment taking place prior to exposure (pre) and after every 4th video clip (1, 2, 3, 4) during the exposure session as well as post-exposure. Error bars represent the standard error of the mean

is provided in the Supplementary Material. Models examining group effects in the overall sample are reported regardless of significance.

Group Effects on Pre-post Changes in Symptoms

Repeated assessments of BAT scores were non-independent as reflected in an ICC of 0.85. A significant effect of Time emerged in the baseline model, reflecting a significant incline in BAT scores and hence increased approaching the phobic stimulus across time, $B = 1.11$, 95% CI [0.88, 1.33], $p < 0.001$, see Fig. 5a. Including the Group variable as predictor of symptom change did not result in a significant improvements of model fit, $\chi^2_{diff}(2) = 5.04$, $p = 0.080$, indicating that both groups showed a similar linear increase across time. Repeating these analyses including only female participants similarly revealed a significant effect of Time in the baseline model, reflecting a significant incline in BAT scores and hence increased approaching the phobic stimulus across time, $B = 1.09$, 95% CI [0.86, 1.32], $p < 0.001$. Including the Group variable as predictor of symptom change did not result in a significant improvements of model fit, $\chi^2_{diff}(2) = 1.62$, $p = 0.444$, indicating that both groups showed a similar linear increase across time.

Repeated assessments of SNAQ scores were non-independent as reflected in an ICC of 0.71. A significant effect of Time emerged in the baseline model, reflecting a significant decrease in SNAQ scores and hence a decrease of anxious cognitions about the phobic stimulus across time, $B = -1.63$, 95% CI [− 2.05, − 1.21], $p < 0.001$, see Fig. 5b. Including the Group variable as predictor of symptom change did not result in significant improvements of model fit, $\chi^2_{diff}(2) = 3.13$, $p = 0.209$, indicating that both groups showed a similar linear increase across time. Repeating these analyses including only female participants similarly revealed a significant effect of Time in the baseline model, reflecting a significant decline in SNAQ scores and hence a decrease of anxious cognitions about the phobic stimulus across time, $B = -1.65$, 95% CI [− 2.10, − 1.20], $p < 0.001$. Including the Group variable as predictor of symptom change did not result in a significant improvements of model fit, $\chi^2_{diff}(2) = 1.50$, $p = 0.472$, indicating that both groups showed a similar linear decrease across time.

Cortisol Levels, TSW, Vigilance, and Sleep Quality (TST and SE) as Predictors of Pre-post Changes in Symptoms

Including $AUC_{G\text{-}EXP}$ scores as predictor of BAT scores did not result in significant improvements of model fit, $\chi^2_{diff}(2) = 1.04$, $p = 0.594$. Nor did including TSW, $\chi^2_{diff}(2) = 4.23$, $p = 0.120$, baseline TST, $\chi^2_{diff}(2) = 1.64$, $p = 0.441$, baseline SE, $\chi^2_{diff}(2) = 0.29$, $p = 0.865$,

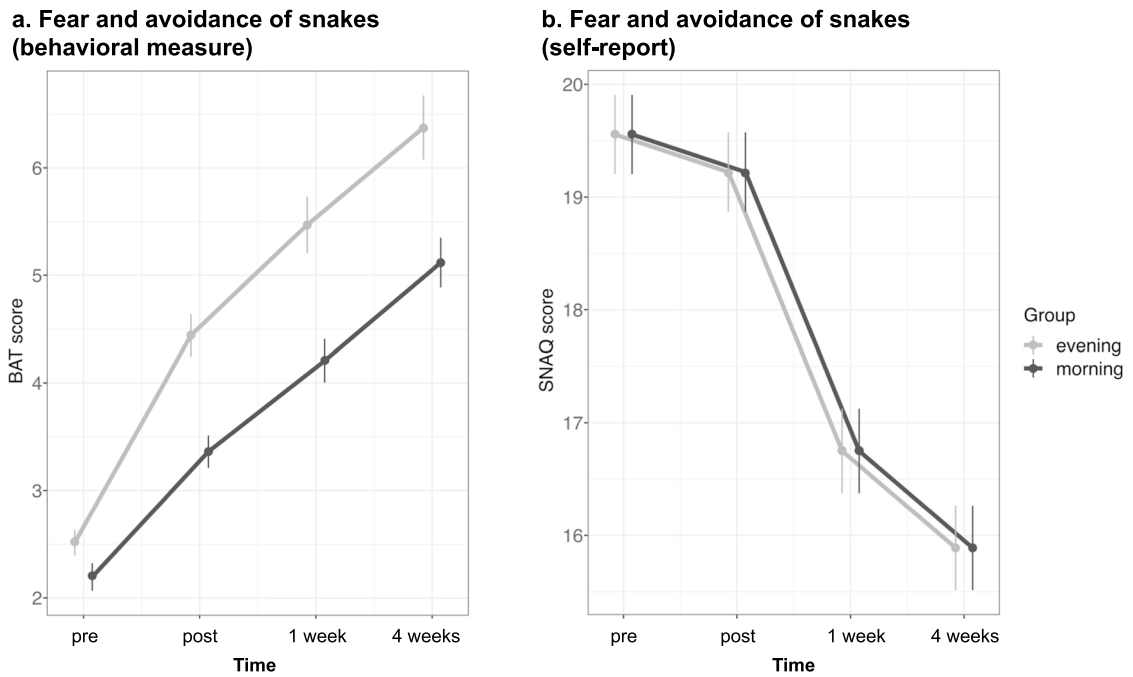

Fig. 5 Changes in BAT and SNAQ scores from pre-exposure to follow-up assessments. **a** represents scores from the Behavior Approach Test (range from 0 to 12, the higher, the better), **b** presents ratings from the Snake Questionnaire (range from 0 to 30, the lower, the better). Time refers to time of measurement with assessment taking place prior to exposure (pre), post exposure (post) and at the first (1 week) and second follow-up (4 weeks). Error bars represent the standard error of the mean. *BAT* Behavioral Approach Test, *SNAQ* Snake Questionnaire

pre-exposure TST, $\chi^2_{diff}(2) = 0.70$, $p = 0.706$, or pre-exposure SE, $\chi^2_{diff}(2) = 0.41$, $p = 0.817$. Including vigilance levels as predictor did improve model fit, $\chi^2_{diff}(2) = 9.24$, $p < 0.001$. Inspection of the model revealed that individuals with a higher vigilance (lower reaction time) showed a stronger increase of BAT scores across time, $B = -0.01$, 95% CI [− 0.02, 0.00], $p = 0.003$.

We repeated the analyses with SNAQ scores as outcome and did not find a significant improvement of model fit by introducing $AUC_{G\text{-}EXP}$, $\chi^2_{diff}(2) = 0.24$, $p = 0.888$, TSW, $\chi^2_{diff}(2) = 3.50$, $p = 0.174$, baseline TST, $\chi^2_{diff}(2) = 5.16$, $p = 0.076$, baseline SE, $\chi^2_{diff}(2) = 3.15$, $p = 0.207$, pre-exposure TST, $\chi^2_{diff}(2) = 0.85$, $p = 0.652$, pre-exposure SE, $\chi^2_{diff}(2) = 4.93$, $p = 0.085$, or vigilance, $\chi^2_{diff}(2) = 4.52$, $p = 0.104$, as predictors.

Vigilance and Sleep Quality as Moderators of Pre-post Changes in Symptoms

Neither vigilance, $\chi^2_{diff}(6) = 11.65$, $p = 0.070$, baseline TST, $\chi^2_{diff}(6) = 7.17$, $p = 0.306$, baseline SE, $\chi^2_{diff}(6) = 6.97$, $p = 0.324$, nor pre-exposure TST, $\chi^2_{diff}(6) = 8.32$, $p = 0.216$, were found to moderate the effects of Group on BAT scores as indicated by lack of significant improvement of model fit. The model including pre-exposure SE as moderator improved model fit, $\chi^2_{diff}(6) = 12.79$, $p = 0.047$, however the improvement was not related to the Group × Time × SE interaction but reflected that higher BAT scores at baseline were linked to greater pre-exposure SE in the night before exposure therapy in the evening but not in the morning group, Group × SE: $B = 0.17$, 95% CI [0.06, 1.21], $p < 0.001$. The model including arousal levels during exposure as moderator showed improved model fit, $\chi^2_{diff}(6) = 18.62$, $p = 0.005$. The improvement was related to a significant main effect of Arousal, indicating that higher arousal during exposure was linked to lower BAT scores, $B = -0.34$, 95% CI [− 0.52, − 0.15], $p < 0.001$. Group-related effects remained unchanged.

We repeated the analyses with SNAQ scores as outcome and did not find any significant moderation effects of baseline TST, $\chi^2_{diff}(6) = 7.05$, $p = 0.317$, baseline SE, $\chi^2_{diff}(6) = 6.84$, $p = 0.336$, pre-exposure TST, $\chi^2_{diff}(6) = 4.25$, $p = 0.643$, pre-exposure SE, $\chi^2_{diff}(6) = 8.17$, $p = 0.226$, or vigilance: $\chi^2_{diff}(6) = 6.27$, $p = 0.394$. The model including arousal levels during exposure as moderator showed improved model fit, $\chi^2_{diff}(6) = 17.78$, $p = 0.007$. The improvement was related to a significant main effect of Arousal, indicating that higher arousal during exposure was linked to higher SNAQ scores, $B = 0.72$, 95% CI [0.28, 1.15], $p < 0.001$. In addition, a main effect of Group emerged, indicating higher SNAQ scores in the evening than in the morning group, $B = 0.72$, 95% CI [0.28, 1.15], $p < 0.001$.

Post-follow-up Maintenance of Symptom Changes

Note that an overview of intercepts and slopes as well as the estimated variance accounted for by each significant model is provided in the Supplementary Material. Models examining group effects in the overall sample are reported regardless of significance.

Group Effects on Post-follow-up Maintenance of Symptom Changes

Repeated assessments of BAT scores were non-independent as reflected in an ICC of 0.91. A significant effect of Time emerged in the baseline model, reflecting a significant incline in BAT scores and hence an increased approaching the phobic stimulus across time, $B = 0.93$, 95% CI [0.67, 1.18], $p < 0.001$, see Fig. 5a. Including the Group variable as predictor of symptom change did not result in significant improvements of model fit, $\chi^2_{diff}(2) = 5.14$, $p = 0.077$, indicating that both groups showed a similar linear increase across time. Repeating these analyses including only female participants similarly revealed a significant effect of Time in the baseline model, reflecting a significant incline in BAT scores and hence increased approaching the phobic stimulus across time, $B = 0.88$, 95% CI [0.61, 1.15], $p < 0.001$. Including the Group variable as predictor of symptom change did not result in a significant improvements of model fit, $\chi^2_{diff}(2) = 2.42$, $p = 0.298$, indicating that both groups showed a similar linear increase across time.

Repeated assessments of SNAQ scores were non-independent as reflected in an ICC of 0.81. A significant effect of Time emerged in the baseline model, reflecting a significant decrease in SNAQ scores and hence a decrease of anxious cognitions about the phobic stimulus across time, $B = -1.87$, 95% CI [− 2.34, − 1.39], $p < 0.001$, see Fig. 5b. Including the Group variable as predictor of symptom change did not result in significant improvements of model fit, $\chi^2_{diff}(2) = 0.82$, $p = 0.665$, indicating that both groups showed a similar linear increase across time. Repeating these analyses including only female participants similarly revealed a significant effect of Time in the baseline model, reflecting a significant decline in SNAQ scores and hence a decrease of anxious cognitions about the phobic stimulus across time, $B = -1.74$, 95% CI [− 2.28, − 1.20], $p < 0.001$. Including the Group variable as predictor of symptom change did not result in a significant improvements of model fit, $\chi^2_{diff}(2) = 0.66$, $p = 0.718$, indicating that both groups showed a similar linear decrease across time.

Cortisol Levels, TSW, Vigilance, and Sleep Quality (TST and SE) as Predictors of Post-follow-up Changes in Symptoms

Including $AUC_{G\text{-}EXP}$ scores as predictor of BAT scores did not result in a significant improvement of model

fit, $\chi^2_{diff}(2) = 0.97$, $p = 0.616$. Nor did including TSW, $\chi^2_{diff}(2) = 4.28$, $p = 0.118$, baseline TST, $\chi^2_{diff}(2) = 2.61$, $p = 0.271$, baseline SE, $\chi^2_{diff}(2) = 0.58$, $p = 0.749$, pre-exposure TST, $\chi^2_{diff}(2) = 2.66$, $p = 0.265$, or pre-exposure SE, $\chi^2_{diff}(2) = 1.14$, $p = 0.567$. However, including vigilance levels as predictor did improve model fit, $\chi^2_{diff}(2) = 9.07$, $p = 0.011$. Inspection of the model revealed that individuals with a higher vigilance (lower reaction time) showed a stronger increase of BAT scores across time, $B = -0.01$, 95% CI [− 0.02, 0.00], $p = 0.011$.

We repeated the analyses with SNAQ scores as outcome and did not find a significant improvement of model fit by introducing $AUC_{G\text{-}EXP}$, $\chi^2_{diff}(2) = 0.35$, $p = 0.842$, TSW, $\chi^2_{diff}(2) = 0.84$, $p = 0.657$, baseline TST, $\chi^2_{diff}(2) = 4.26$, $p = 0.119$, baseline SE, $\chi^2_{diff}(2) = 3.27$, $p = 0.195$, pre-exposure TST, $\chi^2_{diff}(2) = 0.17$, $p = 0.919$, pre-exposure SE, $\chi^2_{diff}(2) = 3.77$, $p = 0.152$, or vigilance, $\chi^2_{diff}(2) = 1.64$, $p = 0.440$, as predictors.

Vigilance and Sleep Quality as Moderators of Post-follow-up Changes in Symptoms

Neither vigilance, $\chi^2_{diff}(6) = 12.20$, $p = 0.058$, baseline TST, $\chi^2_{diff}(6) = 10.17$, $p = 0.118$, baseline SE, $\chi^2_{diff}(6) = 7.54$, $p = 0.274$, nor pre-exposure TST, $\chi^2_{diff}(6) = 8.67$, $p = 0.192$, were found to moderate Group effects on BAT scores as indicated by lack of significant improvement of model fit. The model including pre-exposure SE as moderator improved model fit, $\chi^2_{diff}(6) = 13.94$, $p = 0.030$, and yielded a significant Group × Time × SE interaction, $B = -0.09$, 95% CI [− 0.18, − 0.00], $p = 0.047$. Decomposing this interaction revealed that, for individuals with high pre-exposure SE (mean + 2 SD; see Supplementary Material for further information on the distribution of pre-exposure SE), a significant slope was estimated in the morning exposure condition, $B = 1.17$, 95% CI [0.28, 2.11], $p = 0.016$ (indicating symptom reduction from post- to follow-up), whereas a non-significant slope was estimated in the evening exposure condition, $B = 0.26$, 95% CI [− 0.90, 1.41], $p = 0.661$ (indicating no substantial symptom reduction from post- to follow-up). By contrast, for individuals with low pre-exposure SE (mean − 2 SD), a significant positive slope was estimated in the evening exposure condition, $B = 2.32$, 95% CI [1.30, 3.35], $p < 0.001$ (indicating symptom reduction from post- to follow-up), whereas a non-significant slope was estimated in the morning exposure condition, $B = 0.61$, 95% CI [− 0.39, 1.61], $p = 0.230$ (indicating no substantial symptom reduction from post- to follow-up). The model including arousal levels during exposure as moderator showed improved model fit, $\chi^2_{diff}(6) = 15.31$, $p = 0.018$. The improvement was related to a significant main effect of Arousal, indicating that higher arousal during exposure was linked to lower BAT scores, $B = -0.39$, 95% CI [− 0.66, − 0.11], $p = 0.006$. Group-related effects remained unchanged.

We repeated the analyses with SNAQ scores as outcome and did not find any significant moderation effects of baseline TST, $\chi^2_{diff}(6) = 10.99$, $p = 0.089$, baseline SE, $\chi^2_{diff}(6) = 7.54$, $p = 0.274$, pre-exposure TST, $\chi^2_{diff}(6) = 3.04$, $p = 0.803$, or pre-exposure SE, $\chi^2_{diff}(6) = 7.01$, $p = 0.320$, vigilance: $\chi^2_{diff}(6) = 7.47$, $p = 0.280$. The model including arousal levels during exposure as moderator showed improved model fit, $\chi^2_{diff}(6) = 16.72$, $p = 0.010$. The improvement was related to a significant main effect of Arousal, indicating that higher arousal during exposure was linked to higher SNAQ scores, $B = 1.05$, 95% CI [0.55, 1.54], $p < 0.001$.

Discussion

The current study set out to replicate previous research showing that exposure therapy is more effective in the morning than in the evening, while shedding further light on the involvement of cortisol levels, sleep, and vigilance. In contrast to previous research, we did not find that a video-based exposure session was more effective in the morning than in the evening. Both behavioral and subjective assessments of snake fear were found to decrease from pre- to post-intervention and from post- to follow-up. Controlling for baseline differences in arousal during the exposure session, revealed significantly higher subjectively experienced snake fear in the evening as opposed to the morning group. However, since this effect was not qualified by a significant interaction between Group and Time, we refrain from interpreting it in terms of intervention effects. Interestingly, we did find indications that vigilance and pre-exposure sleep efficiency may be involved in modulating daytime effects on exposure therapy. On the one hand, we found that vigilance levels were higher in the evening group and greater vigilance predicted a greater post-exposure increase of BAT scores and further increase of BAT scores in the follow-up period across both groups. On the other hand, we found that morning as opposed to evening exposure was associated with a stronger increase of BAT scores in the follow-up period, however this effect was only estimated for individuals with high pre-exposure sleep efficiency and estimated inversely for individuals with low pre-exposure sleep efficiency. However, this effect was only found in a restricted subsample for which actigraphy data was available ($n = 16$ in the evening group and $n = 21$ in the morning group). Moreover, sleep efficiency was generally rather high. Neither baseline sleep quality nor cortisol levels during exposure were found to predict treatment-related changes in behavioral or subjective snake fear.

Our findings on vigilance indicate that—contrary to our assumption—vigilance levels were higher in the evening than in the morning. Although some studies show that vigilance is higher in the morning, other studies indicate that this effect varies according to chronotype (Harrison et al., 2007; Riley et al., 2017). That is, evening types may show higher vigilance in the evening as opposed to the morning and morning types may show higher vigilance in the morning as opposed to the evening (Correa et al., 2014; Venkat et al., 2020). Our result could thus indicate that our participants who were largely recruited amongst university students were tested at their non-optimal time of day when assigned to the morning group. In order to further explore this possibility, we examined the distribution of morning and evening types and found that 16.7% of participants (evening types) were tested at their non-optimal time in the morning, whereas only 5.7% of participants (i.e., morning types) were tested at their non-optimal time in the evening. This disproportionate misalignment may have caused the baseline difference in PVT performance. Additionally, vigilance was found to predict the increase of BAT scores across time. Taken together, chronobiological factors may have prevented us from replicating the previous findings of Lass-Hennemann and Michael (2014).

Our second finding indicates that a superior effect of morning exposure may be present, but only in individuals with high pre-exposure sleep efficiency and only in the follow-up period. These results suggest that the "morning exposure effect" may be hampered by insufficient nighttime sleep. As with the aforementioned possible interference of chronobiological factors, the moderation thus stresses that individual factors must be taken into account. For instance, getting up early in the morning to attend morning exposure therapy may affect preceding sleep quality, especially in younger populations with a tendency towards eveningness. The anticipation of exposure therapy in the morning could also cause difficulties falling asleep, thereby affecting pre-exposure sleep efficiency. Such individual factors should be considered when scheduling appointments with patients. In addition, it seems worthwhile to provide patients with psychoeducation and tools to improve sleep quality (e.g., sleep directed hypnosis; Friesen et al., 2023) in order to boost pre-exposure sleep efficiency (and thereby morning vigilance levels).

Beyond these considerations, there are further explanations that may account for our failure to replicate the "morning exposure effect" as well as the previously reported correlations between cortisol levels and exposure therapy outcome (Lass-Hennemann & Michael, 2014; Meuret et al., 2015, 2016). First, some studies did not show a significant association between cortisol levels during exposure and symptom change or even showed an inverse association (Kuhlman et al., 2020). Kuhlman et al. (2020) argue that these mixed findings are related to the fact that endogenous cortisol during treatment does not only reflect daytime variations but also cortisol reactivity, which may be linked to less symptom improvement throughout exposure therapy (see also Rauch et al., 2017). To explore this possibility, we conducted separate analyses with pre-exposure cortisol levels (t_0) as predictor of symptoms. However, none of these analyses yielded a significant result. Moreover, in contrast to the cortisol reactivity hypothesis, there are several studies showing that patients do not experience a stress-related endogenous cortisol reaction to exposure therapy (Gustafsson et al., 2008; Kellner et al., 2012; Lass-Hennemann & Michael, 2014; Siegmund et al., 2011). Second, we examined video-based in virtuo exposure whereas previous research examined in vivo exposure (Lass-Hennemann & Michael, 2014; Meuret et al., 2015, 2016). We opted for the video-based approach, as in vivo exposure trials are not able to achieve full blinding, because the involved psychotherapists are often not blind to study hypothesis (Lass-Hennemann & Michael, 2014). Although allowing us to test effects under highly standardized conditions, this approach may have dampened exposure effects and thereby the potential of finding daytime differences. Moreover, we examined high snake anxious individuals, whereas our previous study examined individuals with spider phobia (Lass-Hennemann & Michael, 2014). Thus, anxiety levels may not have been sufficiently high to detect any daytime effects on exposure. However, it is important to note that we did find significant, albeit small, exposure effects both in terms of symptom changes and anxiety/arousal ratings during exposure. Finally, it is important to consider potential confounding effects of sex, since the morning acrophase of testosterone may interact with cortisol in generating the morning exposure effect (see e.g., Hutschemaekers et al., 2020). In order to explore this possibility, we repeated our analyses while including only female participants. These analyses did not yield any significant group-related effects, thus paralleling our results presented above.

Overall, it is important to point out that effects of vigilance and pre-exposure sleep efficiency were only evident for behavioral but not for subjective fear indices. However, the BAT is considered the gold standard in the assessment of phobic fear and has been reported as the primary outcome measure in many studies on treatment of specific phobias (Lambe et al., 2023). Moreover, our findings are in line with previous studies showing effects only for behavioral or self-report outcome measures (e.g., de Quervain et al., 2011; Lass-Hennemann & Michael, 2014) and could indicate a lack of agreement between these measures (Reinecke et al., 2009). In addition, effects of pre-exposure sleep efficiency were only evident in the follow-up period and not in our pre-post analyses. Though speculative, this finding could indicate that effects of pre-exposure sleep only emerge over

time when intervention effects are diminished by the time lag between intervention and testing (for similar findings see Soravia et al., 2014).

Additionally, several limitations of our study must be considered. First, we used actigraphy rather than polysomnography for the assessment of pre-exposure sleep, which is known to overestimate sleep duration and does not allow differentiating between different sleep stages (Marino et al., 2013). However, actigraphy also has some advantages as it allows for a non-invasive and economic assessment of sleep quality in natural sleep settings. Due to practical considerations, we only collected actigraphy data for three nights prior to the exposure session. Future studies should consider assessing baseline sleep for a minimum of seven nights to improve reliability (Aili et al., 2017).

Second, even though we included men and women in our study, the vast majority of our participants self-identified as women. Although more women suffer from snake phobia than men, the sex ratio in epidemiological studies is not as unequally distributed as in our study (Oosterink et al., 2009). One major strength of our study is that it was preregistered with an a-priori sample size calculation. However, we failed to reach the desired sample size in several subanalyses, which may have limited statistical power. This concern especially applies to our analyses of actigraphy data. Our findings of the moderator analyses thus have to be interpreted with caution and require replication in adequately powered samples. In an effort to improve statistical power, we conducted exploratory analyses examining subjective SE (calculated based on sleep logs) as a moderator of symptom improvement. These analyses did not reveal any significant effects. While this seems to disconfirm our actigraphy-based findings, it is important to note the low level of agreement between actigraphy-based and sleep log-based assessment of sleep quality (Girschik et al., 2012; McCall & McCall, 2012). Hence, inconsistent effects may be related to lack of agreement between SE measures rather than poor reliability of our findings in actigraphy-based analyses.

In summary, our preregistered experimental study aimed to replicate the "morning exposure effect" and is (to our knowledge) the first study, which systematically assessed different potential factors contributing to the "morning exposure effect". Further research is needed to confirm our findings and generalize them to the wider population of individuals with anxiety disorders. Such research should aim to improve shortcomings of our study, while taking into account the predictors and moderators that we identified, namely, vigilance levels and pre-exposure sleep efficiency. Such studies may also consider investigating daytime effects within patients by varying daytime between repeated exposure session (Meuret et al., 2015, 2016). Though preliminary, our study shines further light on the intricate relations between daytime effects, vigilance, and sleep, suggesting that clinicians should take all factors that are linked to these processes (e.g., chronotype, difficulties falling asleep) into account when scheduling individual exposure sessions. Simply scheduling exposure sessions in the morning does not seem to be sufficient to achieve optimal treatment effects.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10608-023-10463-9.

Funding Open Access funding enabled and organized by Projekt DEAL. The study was funded by the German Sleep Research Society and Start-up funding from Saarland University.

Data Availability Data will be provided by the authors upon reasonable request.

Declarations

Conflict of Interest Marie Roxanne Sopp, Sarah K. Schäfer, Tanja Michael, Monika Equit, Diana S. Ferreira de Sá, Johanna Lass-Hennemann declare that they have no conflict of interest.

Ethical Approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Saarland University (No.: 18-19).

Informed Consent All participants gave their written informed consent.

References

Aggio, D., Smith, L., Fisher, A., & Hamer, M. (2015). Association of light exposure on physical activity and sedentary time in young people. *International Journal of Environmental Research and Public Health, 12*(3), 2941–2949.

Aili, K., Åström-Paulsson, S., Stoetzer, U., Svartengren, M., & Hillert, L. (2017). Reliability of actigraphy and subjective sleep measurements in adults: The design of sleep assessments. *Journal of Clinical Sleep Medicine, 13*(1), 39–47.

Bates, D. M. (2010). *lme4: mixed-effects modeling with R*. Springer.

Bentz, D., Michael, T., Dominique, J.-F., & Wilhelm, F. H. (2010). Enhancing exposure therapy for anxiety disorders with glucocorticoids: From basic mechanisms of emotional learning to clinical applications. *Journal of Anxiety Disorders, 24*(2), 223–230.

Buysse, D. J., Reynolds, C. F., III., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213.

Cole, R. J., Kripke, D. F., Gruen, W., Mullaney, D. J., & Gillin, J. C. (1992). Automatic sleep/wake identification from wrist activity. *Sleep, 15*(5), 461–469.

Correa, Á., Molina, E., & Sanabria, D. (2014). Effects of chronotype and time of day on the vigilance decrement during simulated driving. *Accident Analysis & Prevention, 67*, 113–118.

de Quervain, D. J. -F., Bentz, D., Michael, T., Bolt, O. C., Wiederhold, B. K., Margraf, J., & Wilhelm, F. H. (2011). Glucocorticoids enhance extinction-based psychotherapy. *Proceedings of the National Academy of Sciences, 108*(16), 6621–6625.

Dressendörfer, R., Kirschbaum, C., Rohde, W., Stahl, F., & Strasburger, C. (1992). Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. *The Journal of Steroid Biochemistry and Molecular Biology, 43*(7), 683–692.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160.

Fischer, S., & Cleare, A. J. (2017). Cortisol as a predictor of psychological therapy response in anxiety disorders—Systematic review and meta-analysis. *Journal of Anxiety Disorders, 47*, 60–68.

Friesen, E., Sopp, M. R., Cordi, M. J., Rasch, B., & Michael, T. (2023). Sleep-directed hypnosis improves subjective sleep quality but not extinction memory after exposure to analog trauma. *Cognitive Therapy and Research, 47*, 1–14.

Girschik, J., Fritschi, L., Heyworth, J., & Waters, F. (2012). Validation of self-reported sleep against actigraphy. *Journal of Epidemiology, 22*(5), 462–468.

Groch, S., Wilhelm, I., Diekelmann, S., & Born, J. (2013). The role of REM sleep in the processing of emotional memories: Evidence from behavior and event-related potentials. *Neurobiology of Learning and Memory, 99*, 1–9.

Gustafsson, P. E., Gustafsson, P. A., Ivarsson, T., & Nelson, N. (2008). Diurnal cortisol levels and cortisol response in youths with obsessive-compulsive disorder. *Neuropsychobiology, 57*(1–2), 14–21.

Hamm, A. (2006). *Spezifische Phobien*. Hogrefe.

Harrison, Y., Jones, K., & Waterhouse, J. (2007). The influence of time awake and circadian rhythm upon performance on a frontal lobe task. *Neuropsychologia, 45*(8), 1966–1972.

Hautzinger, M., Keller, F., & Kühner, C. (2006). Das Beck Depressionsinventar II: Deutsche Bearbeitung und Handbuch zum BDI II [The Beck Depression Inventory II: German version of the BDI II]. HarcourtTest Services.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research, 212*, 429–437.

Hughes, J., Beiner, D., & Hughes, M. J. (2022). Package 'reghelper'. *R package version, 1*(1).

Hutschemaekers, M. H. M., de Kleine, R. A., Davis, M. L., Kampman, M., Smits, J. A. J., & Roelofs, K. (2020). Endogenous testosterone levels are predictive of symptom reduction with exposure therapy in social anxiety disorder. *Psychoneuroendocrinology, 115*, 104612.

Ihmig, F. R., Neurohr-Parakenings, F., Schäfer, S. K., Lass-Hennemann, J., & Michael, T. (2020). On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals. *PLoS ONE, 15*(6), e0231517.

Kaida, K., Niki, K., & Born, J. (2015). Role of sleep for encoding of emotional memory. *Neurobiology of Learning and Memory, 121*, 72–79.

Kellner, M., Wiedemann, K., Yassouridis, A., & Muhtz, C. (2012). Non-response of cortisol during stressful exposure therapy in patients with obsessive-compulsive disorder—Preliminary results. *Psychiatry Research, 199*(2), 111–114.

Kemper, C. J., Ziegler, M., & Taylor, S. (2009). Überprüfung der psychometrischen Qualität der deutschen Version des Angstsensitivitätsindex-3. *Diagnostica, 55*(4), 223–233.

Klieger, D. M. (1987). The Snake Anxiety Questionnaire as a measure of ophidophobia. *Educational and Psychological Measurement, 47*(2), 449–459.

Kreft, I. G., De Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30*(1), 1–21.

Kuhlman, K. R., Treanor, M., Imbriano, G., & Craske, M. G. (2020). Endogenous in-session cortisol during exposure therapy predicts symptom improvement: Preliminary results from a scopolamine-augmentation trial. *Psychoneuroendocrinology, 116*, 104657.

Lambe, S., Bird, J. C., Loe, B. S., Rosebrock, L., Kabir, T., Petit, A., Mullhall, S., Jenner, L., Aynsworth, C., Murphy, E., Jones, J., Powling, R., Chapman, K., Dudley, R., Morrison, A., Regan, E. O., Yu, L., Clark, D., Waite, F., & Freeman, D. (2023). The Oxford agoraphobic avoidance scale. *Psychological Medicine, 53*(4), 1233–1243. https://doi.org/10.1017/S0033291721002713

Lass-Hennemann, J., & Michael, T. (2014). Endogenous cortisol levels influence exposure therapy in spider phobia. *Behaviour Research and Therapy, 60*, 39–45.

Lass-Hennemann, J., Tuschen-Caffier, B., & Michael, T. (2018). Expositionsverfahren. *Lehrbuch der Verhaltenstherapie, Band 1: Grundlagen, Diagnostik, Verfahren und Rahmenbedingungen psychologischer Therapie* (pp. 411–424).

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. (1981). Stai. *State-trait-angstinventar.* Beltz Test Gmbh.

Löwe, B., Kroenke, K., Herzog, W., & Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: Sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders, 81*(1), 61–66.

Löwe, B., Spitzer, R. L., Zipfel, S., & Herzog, W. (2002). *Ge-sundheitsfragebogen für Patienten (PHQ-D). Manual undTestunterlagen* (2nd ed.). Pfizer.

Mander, B. A., Santhanam, S., Saletin, J. M., & Walker, M. P. (2011). Wake deterioration and sleep restoration of human learning. *Current Biology, 21*(5), R183–R184.

Marino, M., Li, Y., Rueschman, M. N., Winkelman, J. W., Ellenbogen, J. M., Solet, J. M., Dulin, H., Berkman, L. F., & Buxton, O. M. (2013). Measuring sleep: Accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep, 36*(11), 1747–1755. https://doi.org/10.5665/sleep.3142

McCall, C., & McCall, W. V. (2012). Comparison of actigraphy with polysomnography and sleep logs in depressed insomniacs. *Journal of Sleep Research, 21*(1), 122–127.

Meuret, A. E., Trueba, A. F., Abelson, J. L., Liberzon, I., Auchus, R., Bhaskara, L., Ritz, T., & Rosenfield, D. (2016). Timing matters: Endogenous cortisol mediates benefits from early-day psychotherapy. *Psychoneuroendocrinology, 74*, 197–202.

Meuret, A. E., Trueba, A. F., Abelson, J. L., Liberzon, I., Auchus, R., Bhaskara, L., Ritz, T., & Rosenfield, D. (2015). High cortisol awakening response and cortisol levels moderate exposure-based psychotherapy success. *Psychoneuroendocrinology, 51*, 331–340. http://doi.org/10.1016/j.psyneuen.2014.10.008

Michael, T., Schanz, C. G., Mattheus, H. K., Issler, T., Frommberger, U., Köllner, V., & Equit, M. (2019). Do adjuvant interventions improve treatment outcome in adult patients with posttraumatic stress disorder receiving trauma-focused psychotherapy? A systematic review. *European Journal of Psychotraumatology, 10*(1), 1634938.

Nissen, C., Kuhn, M., Hertenstein, E., & Landmann, N. (2017). Sleep-related interventions to improve psychotherapy. In *Cognitive neuroscience of memory consolidation* (pp. 381–400).

Oosterink, F. M., De Jongh, A., & Hoogstraten, J. (2009). Prevalence of dental fear and phobia relative to other fear and phobia subtypes. *European Journal of Oral Sciences, 117*(2), 135–143.

Öst, L.-G. (1989). One-session treatment for specific phobias. *Behaviour Research and Therapy, 27*(1), 1–7.

Pace-Schott, E. F., Spencer, R. M., Vijayakumar, S., Ahmed, N. A., Verga, P. W., Orr, S. P., Pitman, R. K., & Milad, M. R. (2013). Extinction of conditioned fear is better learned and recalled in the morning than in the evening. *Journal of Psychiatric Research, 47*(11), 1776–1784. https://doi.org/10.1016/j.jpsychires.2013.07.027

Pace-Schott, E. F., Verga, P. W., Bennett, T. S., & Spencer, R. M. (2012). Sleep promotes consolidation and generalization of extinction learning in simulated exposure therapy for spider fear. *Journal of Psychiatric Research, 46*(8), 1036–1044.

Parasuraman, R. (2000). The attentive brain: issues and prospects. In R. Parasuraman (Ed.), *The attentive brain* (pp. 3–16). MIT Press.

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology, 28*(7), 916–931.

Raeder, F., Merz, C. J., Tegenthoff, M., Wolf, O. T., Margraf, J., & Zlomuzica, A. (2019). Post-exposure cortisol administration does not augment the success of exposure therapy: A randomized placebo-controlled study. *Psychoneuroendocrinology, 99*, 174–182.

Randler, C. (2013). German version of the reduced Morningness-Eveningness Questionnaire (rMEQ). *Biological Rhythm Research, 44*(5), 730–736.

Rauch, S. A., King, A. P., Liberzon, I., & Sripada, R. K. (2017). Changes in salivary cortisol during psychotherapy for posttraumatic stress disorder: A pilot study in 30 veterans. *The Journal of Clinical Psychiatry, 78*(5), 2490.

Reinecke, A., Hoyer, J., Becker, E., & Rinck, M. (2009). Two short-screenings measuring fear of snakes: Reliability and validity by contrast with the SNAQ. *Klin. Diagn. Eval, 3*, 221–239.

Riemann, D., & Backhaus, J. (1996). *Behandlung von Schlafstörungen: ein psychologisches Gruppenprogramm*. Beltz, Psychologie-Verlag-Union.

Riley, E., Esterman, M., Fortenbaugh, F. C., & DeGutis, J. (2017). Time-of-day variation in sustained attentional control. *Chronobiology International, 34*(7), 993–1001.

Roach, G. D., Dawson, D., & Lamond, N. (2006). Can a shorter psychomotor vigilance task be usedas a reasonable substitute for the ten-minute psychomotor vigilance task? *Chronobiology International, 23*(6), 1379–1387.

Schäfer, S. K., Ihmig, F. R., Lara, H. K. A., Neurohr, F., Kiefer, S., Staginnus, M., Lass-Hennemann, J., & Michael, T. (2018). Effects of heart rate variability biofeedback during exposure to fear-provoking stimuli within spider-fearful individuals: Study protocol for a randomized controlled trial. *Trials, 19*(1), 1–11. https://doi.org/10.1186/s13063-018-2554-2

Schwarz, J. F., Popp, R., Haas, J., Zulley, J., Geisler, P., Alpers, G. W., Osterheider, M., & Eisenbarth, H. (2013). Shortened night sleep impairs facial responsiveness to emotional stimuli. *Biological Psychology, 93*(1), 41–44. https://doi.org/10.1016/j.biopsycho.2013.01.008

Siegmund, A., Köster, L., Meves, A. M., Plag, J., Stoy, M., & Ströhle, A. (2011). Stress hormones during flooding therapy and their relationship to therapy outcome in patients with panic disorder and agoraphobia. *Journal of Psychiatric Research, 45*(3), 339–346.

Singer, J. D., & Willett, J. B. (2014). Growth curve modeling. *Wiley StatsRef: Statistics Reference Online*.

Soravia, L. M., Heinrichs, M., Aerni, A., Maroni, C., Schelling, G., Ehlert, U., Roozendaal, B., & de Quervain, D. J.-F. (2006). Glucocorticoids reduce phobic fear in humans. *Proceedings of the National Academy of Sciences, 103*(14), 5585–5590. https://doi.org/10.1073/pnas.0509184103

Soravia, L. M., Heinrichs, M., Winzeler, L., Fisler, M., Schmitt, W., Horn, H., Dierks, T., Strik, W., Hofmann, S. G., & de Quervain, D. J.-F. (2014). Glucocorticoids enhance in vivo exposure-based therapy of spider phobia. *Depression and Anxiety, 31*(5), 429–435. https://doi.org/10.1002/da.22219

Soravia, L. M., Moggi, F., & de Quervain, D.J.-F. (2021). Effects of cortisol administration on craving during in vivo exposure in patients with alcohol use disorder. *Translational Psychiatry, 11*(1), 6.

Spielberger, C. D. (1970). *Manual for the state-trait anxietry, inventory*. Consulting Psychologist.

Straus, L. D., Acheson, D. T., Risbrough, V. B., & Drummond, S. P. (2017). Sleep deprivation disrupts recall of conditioned fear extinction. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2*(2), 123–129.

Taylor, S., Zvolensky, M. J., Cox, B. J., Deacon, B., Heimberg, R. G., Ledley, D. R., Abramowitz, J. S., Holaway, R. M., Sandin, B., Stewart, S. H., Coles, M., Eng, W., Daly, E. S., Arrindell, W. A., Bouvard, M., & Cardenas, S. J. (2007). Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. *Psychological Assessment*, *19*(2), 176. https://doi.org/10.1037/1040-3590.19.2.176

Team, R. C. (2022). R: A language and environment for statistical computing (version 4.2. 0)[Computer software](4.2. 0). R Foundation for Statistical Computing.

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews, 10*(1), 49–62.

Venkat, N., Sinha, M., Sinha, R., Ghate, J., & Pande, B. (2020). Neurocognitive profile of morning and evening chronotypes at different times of day. *Annals of Neurosciences, 27*(3–4), 257–265.

Wang, Y., Li, X., Huang, J., Cao, X., Cui, J., Zhao, Q., Wang, Y., Shum, D. H. K., & Chan, R. C. K. (2012). Relationship between prospective memory and vigilance: Evidence from ERP. *Chinese Science Bulletin, 57*, 4057–4063. https://doi.org/10.1007/s11434-012-5306-9

Yehuda, R., Bierer, L. M., Pratchett, L. C., Lehrner, A., Koch, E. C., Van Manen, J. A., Flory, J. D., Makotkine, I., & Hildebrandt, T. (2015). Cortisol augmentation of a psychological treatment for warfighters with posttraumatic stress disorder: Randomized trial showing improved treatment retention and outcome. *Psychoneuroendocrinology, 51*, 589–597. https://doi.org/10.1016/j.psyneuen.2014.08.004

Zuj, D. V., Palmer, M. A., Hsu, C. M. K., Nicholson, E. L., Cushing, P. J., Gray, K. E., & Felmingham, K. L. (2016). Impaired fear extinction associated with PTSD increases with hours-since-waking. *Depression and Anxiety, 33*(3), 203–210.