Cognitive Bias Modification-Interpretation in the Context of Social Anxiety: Effects on Stress-Relevant Cognitive and Psychophysiological Markers

Marius Kunna [1], Fanny A. Dietel [2], Felix Würtz [1], Simon E. Blackwell [3], Dirk Adolph[1], Elske Salemink [4], Oliver T. Wolf [5], Jürgen Margraf [1], and Marcella L. Woud [3, 1]

[1] Mental Health Research and Treatment Center, Faculty of Psychology, Ruhr University Bochum, Massenbergstraße 9-13, 44787 Bochum – Germany

[2] Department of Clinical Psychology and Psychotherapy, University of Osnabrueck, Lise-Meitner-Straße 3, 49076 Osnabrück – Germany

[3] Department of Clinical Psychology and Experimental Psychopathology, Georg-Elias-Müller-Institute of Psychology, Georg-August-University Göttingen, Kurze-Geismar-Str.1, 37073 Göttingen – Germany

[4] Department of Clinical Psychology, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht – Netherlands

[5] Department of Cognitive Psychology, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum – Germany

PLEASE NOTE THAT THIS IS A PREPRINT VERSION OF AN UNPUBLISHED MANUSCRIPT

Author Note

Marius Kunna (Marius.Kunna@rub.de) https://orcid.org/0000-0002-3333-8046

Fanny A. Dietel (Fanny.Dietel@uni-osnabrueck.de) https://orcid.org/0000-0002-9911-6155

Felix Würtz (Felix.Wuertz@rub.de) https://orcid.org/0000-0003-1627-9432

Simon E. Blackwell (Simon.Blackwell@uni-goettingen.de) https://orcid.org/0000-0002-3313-7084

Dirk Adolph (Dirk.Adolph@rub.de) https://orcid.org/0000-0001-9068-9910

Elske Salemink (E.Salemink@uu.nl) https://orcid.org/0000-0002-6885-2598

Oliver T. Wolf (Oliver.T.Wolf@ruhr-uni-bochum.de) https://orcid.org/0000-0002-9320-2124

Jürgen Margraf (Juergen.Margraf@rub.de) https://orcid.org/0000-0001-5207-7016

Marcella L. Woud (Marcella.Woud@uni-goettingen.de) https://orcid.org/0000-0002-4974-505X

Preregistration and data are available at https://osf.io/nqab6/overview?view_only=9f69a149ae9848fca322ca716f0deb1c

We have no known conflicts of interest to disclose.

This project was funded by an Emmy Noether Research Grant (Project Identifier: 442163275) awarded to Marcella L. Woud.

Correspondence concerning this article should be addressed to Marcella L. Woud, Department of Clinical Psychology and Experimental Psychopathology, Georg-Elias-Müller-Institute of Psychology, Georg-August-University Göttingen, Kurze-Geismar-Str.1, 37073 Göttingen – Germany. E-mail address: marcella.woud@uni-goettingen.de.

Abstract

Cognitive Bias Modification for Interpretation (CBM-I) is designed to alter interpretation biases (IBs) and may have potential for reducing stress reactivity in individuals with social anxiety. However, evidence for such transfer effects remains inconsistent and largely restricted to specific cognitive or self-report outcomes. Physiological responses, such as heart rate, heart rate variability, and salivary cortisol, have received disproportionately limited attention in CBM-I research, despite their relevance to the social anxiety symptom spectrum. Moreover, previous studies have rarely employed comprehensive experimental designs that directly compare positive, negative, and control training conditions. To address these gaps, the present study compared the effects of three CBM-I training conditions (positive, negative, neutral) on changes in IBs and on self-reported and physiological responses to a standardized laboratory stressor (anagram task). The sample included $N = 87$ individuals with moderate levels of social anxiety. Results showed that CBM-I successfully modified IB in a condition-congruent direction. However, stress-related outcomes changed similarly over time across all conditions. Self-reported stress and heart rate increased, whereas heart rate variability decreased. Cortisol levels remained unchanged. These findings suggest that although CBM-I can modify IB, such changes may not readily translate into reduced acute stress reactivity. Factors such as training intensity, alignment between training and stressor, and the sensitivity of stress markers may moderate this transfer. Future research should refine CBM-I protocols and measurement approaches to better elucidate the mechanisms linking cognitive change with stress physiology.

Keywords: Social anxiety, interpretation bias, CBM-I, cortisol, heart rate, heart rate variability, post-event processing

Cognitive Bias Modification-Interpretation in the Context of Social Anxiety: Effects on Stress-Relevant Cognitive and Psychophysiological Markers

Anxiety disorders are the most prevalent mental health conditions, affecting about 4% of the global population (World Health Organization, 2023). Social anxiety disorder (SAD) is among the most common (Kessler et al., 2005) and is characterized by intense fear of social evaluation, often accompanied by physiological symptoms such as rapid heartbeat, blushing, or sweating (American Psychiatric Association, 2013). A range of theoretical accounts has been proposed to characterize the etiological and maintaining processes of social anxiety disorder. For example, the cognitive model proposed by Clark and Wells (1995) posits that social situations trigger dysfunctional assumptions in individuals with SAD, heightening their perception of social threat and prompting negative automatic thoughts. These thoughts, in turn, give rise to physiological, behavioral, and cognitive symptoms, and foster a self-perception as a “social object”, thereby reinforcing the experience of social threat. Subsequent behavior, such as post-event processing (PEP), that is, excessive rumination following social events (Morrison & Heimberg, 2013), then contributes to the maintenance of SAD. In line with the model, interpretation biases (IBs) are a key factor in SAD. IBs refer to the tendency to process ambiguous, disorder-relevant information in a biased, that is, threatening manner (Hirsch et al., 2016; Mathews & MacLeod, 2005; Woud, 2023). For example, an audience member’s intense gaze during a presentation may be interpreted as judgmental rather than attentive. Individuals with SAD tend to favor such negative interpretations, triggering physiological arousal and reinforcing a vicious cycle between interpretation bias and symptoms (Chen et al., 2020). Consequently, IBs are a central focus in research, with numerous studies linking them to anxiety symptoms (for a recent meta-analysis, see Würtz et al., in press).

Recent research has moved beyond correlational studies of IBs, actively manipulating them via Cognitive Bias Modification for Interpretation (CBM-I) to explore their causal role in social anxiety (e.g., MacLeod & Mathews, 2012; Woud, 2023). Such approaches involve inducing and reducing IBs via computerized trainings and comparing the

effects of these manipulations on symptoms of anxiety. For example, Mathews and Mackintosh's (2000) CBM-I paradigm systematically alters interpretation by repeatedly presenting ambiguous scenarios resolved positively or negatively (e.g., Salemink et al., 2007, 2009; Woud et al., 2012, 2013, 2018, 2021). To illustrate, in a study by Salemink et al. (2009), participants with high trait anxiety and negative social anxiety-related interpretive styles received either daily positive CBM-I or sham training (control condition) for eight days. Positive CBM-I induced training-congruent changes in interpretation from pre- to post-training, with participants interpreting new ambiguous stimuli more positively than controls. Effects generalized to lower state and trait anxiety and reduced general psychopathology, though social anxiety and stress vulnerability did not differ between groups. Importantly, this study did not employ a full experimental design comparing the effects of a positive, negative, and control condition. To the best of our knowledge, no other study has done so to date. However, inclusion of positive, negative and neutral control conditions is necessary if we wish not only to draw conclusions about the effects of experimentally induced positive vs. relatively negative IBs but also to check that both positive and negative training have shifted IBs in the desired direction. Similar results as Salemink et al. (2009) were obtained by Nowakowski et al. (2015), showing that positive CBM-I increased positive interpretations of ambiguous social scenarios in participants with elevated social anxiety; however, these effects did not transfer to self-reported social anxiety or stress responses. In conclusion, despite the large number of studies examining the effects of CBM-I, transfer effects to symptoms are inconsistent (for network meta-analyses and reviews, see e.g., Chen et al., 2020; Fodor et al., 2020; Liu et al., 2017; Jones & Sharpe, 2017).

Interestingly, alongside these inconsistencies, CBM-I research has largely neglected PEP, a key maintaining factor in SAD. PEP contributes to the development of maladaptive self-representations, which, in turn, reinforce negative IBs, anxiety, and avoidance (Brozovich & Heimberg, 2008), thereby perpetuating a vicious cycle. To illustrate, Badra et al. (2017) examined the relationship between PEP and IBs in individuals with low and high social anxiety and found that PEP mediated the association between social anxiety group

and negative IBs, suggesting a mechanism maintaining SAD. They concluded that increased PEP may facilitate negative IBs in individuals with high levels of social anxiety, though their cross-sectional design prevents causal inferences and leaves open the possibility that negative IBs also amplify PEP. These findings underscore the need for experimental research to clarify directionality. Mechanistically, CBM-I may reduce negative IBs and indirectly attenuate PEP, disrupting this vicious cycle.

Beyond cognitive symptoms, psychophysiological markers of stress reactivity, such as heart rate (HR) and heart rate variability (HRV), remain largely understudied as well. This gap is particularly significant given that physiological responses are a core component of SAD, not only as a symptom of anxiety itself (e.g., increased HR and decreased HRV; Kreibig, 2010), but also as a perceptible manifestation (e.g., sweating) that can potentially be subject to negative evaluation by others (American Psychiatric Association, 2013). Accordingly, CBM-I research should also examine these responses to better capture its effects across the full spectrum of social anxiety symptoms (Abado et al., 2023). A recent systematic review identified seven psychophysiological correlates of IBs, including HR and HRV (Collins et al., 2022). Further, the review reports evidence generally supporting the idea that positive CBM-I training has the potential to attenuate psychophysiological markers of stress reactivity as measured by HR or HRV. However, previous results are not always consistent. The study by van Bockstaele et al. (2020), for example, examined the effects of cognitive load during CBM-I on IBs and stress reactivity. Healthy students were randomized to a positive or negative single-session CBM-I training with or without cognitive load. After training, a stress inducing anagram task was employed. HR and HRV was recorded before, during, and after this task. Results showed that positive CBM-I improved stress recovery as evidenced by changes in HR but not HRV (for additional studies on CBM-I effects on HR/HRV, see e.g., Meeten, 2017; Nowakowski et al., 2015; Rozenman et al., 2020).

Cortisol, a hormone released in response to stress (Kirschbaum et al., 1993), is another relevant marker of stress reactivity. However, studies in SAD have yielded inconsistent findings, from heightened responses (Condren, 2002) to no differences from

controls (Klumbies et al., 2014), likely due to methodological variability. Understanding cortisol remains important for clarifying the biological mechanisms of SAD. To the best of our knowledge, only two studies have examined CBM-I effects on cortisol: Hollocks et al. (2016) and Turton et al. (2018), the latter using a single-session CBM-I in women with anorexia nervosa. Participants were randomized to positive (aimed at reducing fear of social rejection) or control conditions, and salivary cortisol was measured before and after training and after a disorder-relevant stressor, namely a test meal. No significant effects were found, which the authors attributed to the test meal being too distal from the training target, underscoring the need for more proximal, content-aligned paradigms.

To conclude, prior work has shown promising effects in the context of CBM-I and SAD, yet transfer effects are inconsistent and limited to certain self-reported and cognitive markers of anxiety, and full experimental designs are lacking. Accordingly, the present study adopted a conservative and rigorous proof-of-principle approach and compared the effects of three conditions: a positive condition, aiming at increasing positive IBs; a negative condition, aiming at reducing positive IBs; and a neutral control condition. Changes in IBs were assessed pre- and post-CBM-I. Following training, self-reported and psychophysiological markers of stress reactivity (HR, HRV, and cortisol) were measured before, during, and after a social anxiety-relevant stressor (anagram task). To modify IBs in the positive and negative direction, the sample consisted of individuals with moderate levels of social anxiety, lowering the risk of strong existing positive or negative IBs (for a similar sample selection approach, see Clarke et al., 2017). Further, this selection increased the relevance of the trained cognitions, encouraging participant engagement. We expected condition-congruent effects on IBs within (i.e., pre- vs. post-training) and between (i.e., post-training differences) CBM-I training conditions, that is, those who were trained positively would show a stronger increase in positive IBs from pre- to post-training, relative to participants in the negative and control condition, and vice versa for those trained negatively. The second aim was to investigate transfer effects to stress reactivity markers measured via self-reported and psychophysiological outcomes. We expected to find condition-congruent

effects within, that is, pre- vs. (during vs.) post-stressor, and between conditions, that is, post-stressor differences. For example, those who were trained positively should show less pronounced self-reported and psychophysiological stress reactions. The third aim was to explore potential transfer effects to PEP, expecting that those who were trained positively would show lower levels of PEP after the stress task compared to the other two conditions.

Method

Design

This experimental lab-based study used a between-subjects design and included three CBM-I training conditions: positive, negative, neutral. Prior to data collection, the study protocol was registered on the Open Science Framework (https://osf.io/nqab6/). All deviations from the protocol are reported in the corresponding sections.

Recruitment, Screening, and Participants

Various outlets were used to advertise the study, for example, flyers and postings on social media. Participants were selected based on a set of inclusion and exclusion criteria, which were assessed as part of a larger online validation project (for more details, see https://osf.io/eyw5g/). To take part in the lab session, the following inclusion criteria had to be fulfilled: aged 18 to 35 years, motivated by studies showing that social anxiety mainly develops in adolescence or, at the latest, in early adulthood and because we aimed for a homogeneous sample in terms of age (Solmi et al., 2022); to cover a moderate range of social anxiety, a score between 28 and 52 on the Social Phobia and Anxiety Inventory (SPAI-G; Fydrich, 2016; for more details on cut-offs, see Randomization); normal or corrected to normal vision; willing and able to complete all study procedures. The exclusion criteria were as follows: scoring >15 and/or a score of >1 on the suicidality item on the 16-item Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR; Rush et al., 2003); psychopharmacological treatment (current or in the past); severe underweight or overweight (<17,5 kg/m² or ≥30 kg/m²); regular use of alcohol (women: >1 glass/day in 5 days/week; men: >2 glasses/day in 5 days/week); regular use of nicotine (>5 cigarettes/month); regular use of illegal drugs (>1x/month); current pregnancy; prior

participation in studies using the anagram task; participation in CBM-I training studies within the last 2 years; careless answering in the online screening (e.g., incorrectly answering attention check questions). Finally, in the online screening, participants were informed that a high level of proficiency in the German language was a prerequisite for participation.

In total, 394 participants completed the online screening. Of these, 87 took part in the lab study (for a participant flowchart, see Figure 1). For sample characteristics, see Table 1.

Online Screening Measures

SPAI-G

Levels of social anxiety were measured using the SPAI-G (Fydrich, 2016), the German adaptation of the original Social Phobia and Anxiety Inventory (Turner et al., 1989). The SPAI-G assesses cognitive, somatic, and behavioral responses to fear in social situations. All 22 items are rated by the participants on a 7-point Likert scale (0 = “never”; 6 = “always”), with 13 items consisting of multiple items mostly focusing on different interaction partners in social situations. To generate the SPAI-G score, items with multiple items are averaged before summing up all 22 items. This results in scores between 0 and 132. Internal consistency in the present sample was $\alpha = 0.63$ [0.50, 0.73]. However, in the screening study from which this sample was drawn, internal consistency was $\alpha = 0.96$ [0.96, 0.97], suggesting that the lower values may partly reflect the restricted range of this sample.

QIDS-SR

The QIDS-SR (Rush et al., 2003) is a self-report questionnaire covering different aspects of depressive symptomatology (using the German version, Rush et al., 2003). All 16 items are rated on a 4-point Likert scale with 0 reflecting the lowest and 3 the highest symptom severity. To obtain the total score, a sum is calculated, using only the highest score on any of the four sleep items (1-4), item 5, the highest score on any of the four appetite/weight items (6-9), items 10 to 14, and the highest score on either of the two psychomotor items (15 & 16), resulting in scores between 0 and 27. Internal consistency in the present sample was $\alpha = 0.54$ [0.38, 0.67]. However, in the screening study from which

this sample was drawn, internal consistency was α = α = 0.78 [0.74, 0.81], suggesting that the lower values may partly reflect the restricted range of this sample.

CBM-I

The CBM-I training was based on the well-established paradigm by Mathews and Mackintosh (2000; for additional studies see e.g., Salemink et al., 2007, 2009; Woud et al., 2012, 2013, 2018, 2021). Participants completed one session of their respective condition (i.e., positive, negative or neutral), consisting of ten blocks, each including eight trials (i.e., 80 trials in total). In the positive and negative training conditions, seven trials in each block were "active" trials, and one trial served as a "filler" to mask the training's goal. Each active trial consisted of an ambiguous, socially relevant scenario (e.g., "You are giving a speech at a wedding and the audience faces indicate…") and concluded with a final word fragment. Participants were instructed to read the scenario descriptions and imagine themselves as vividly as possible in these situations. In their own pace, participants continued to the word fragment, which they were required to complete. Completing the word fragment, in turn, resolved the scenario's ambiguity. According to the respective training condition, the word fragment was always positive (e.g., "int_r_st" / "interest") in the positive condition or negative (e.g., "bor_d_m" / "boredom") in the negative condition. In each block, two active trials were followed by a comprehension question that reinforced the intended interpretation of the scenario. Participants were provided with feedback indicating whether their response was correct. Filler trials also described ambiguous but socially neutral situations (e.g., "You go to a party and look around first. You see an unusual dish in the kitchen. Since you don't know it, you try it, and it tastes…"; "d_licio_s"). The neutral control condition also included 80 trials and followed the same procedure but scenarios were not socially relevant. Instead, they were related to general facts or descriptions of everyday (non-social) scenarios (e.g., "You turn the kettle on and wait for the water to boil. You get a teabag out of the tin, which you put into a mug, and pour the boiling water onto the teabag. Next, you add the…"; word fragment: "m_ lk"; solution: "milk"). The neutral scenarios were translated from Yiend et al. (2014).

Error feedback was provided across all the conditions: If a word fragment was answered incorrectly, “incorrect” appeared on the screen and participants had to enter a new letter until the word fragment was completed correctly.

Main Outcome: Encoding Recognition Task

The Encoding Recognition Task (ERT; Salemink & van den Hout, 2010) comprises an initial encoding phase, followed by a recognition phase. During encoding, participants are presented with ten scenarios and are instructed to read and memorize them. Each scenario consists of a title (e.g., “The students”) and a short description of an ambiguous, social situation (e.g., “You are walking alone along the street and see a group of students chatting. As you walk past the group, they start to...”), and ends with a word fragment (e.g., “g_ggl_”) participants are required to complete by typing in its first missing letter. Importantly, word fragments are designed such that completing them does not resolve the scenarios’ ambiguity. Once the word fragment is completed correctly (e.g., “giggle”), a simple comprehension question follows (e.g., “Did you walk down the street alone?”) which participants answer with “yes” or “no”. During the recognition phase, the titles from the encoding phase are presented again (e.g., “The students”), alongside a positive (e.g., “As you walk past, you realize that they are watching a funny dog video and therefore giggle.”) and negative (e.g. “As you walk past, you realize that they are talking about you and making fun of you.”) interpretation, that is, “targets” (with randomized presentation). Participants then rate on a 4-point Likert scale how close in meaning each target is to the original description that was presented with the corresponding title in the previous phase of the task (1 = “not at all similar”; 4 = “very similar”). Bias scores are calculated by subtracting the mean ratings of negative targets from the mean ratings of positive targets, with a positive score reflecting a stronger positive interpretational processing style and vice versa. Since the ERT was administered twice, there were two versions (i.e., version A and version B), and their order (AB / BA) was counterbalanced (see Supplements for more detailed analyses). Internal consistency in the present sample was as follows: version A pre-training $\alpha = 0.75$ [0.63, 0.85], post-training $\alpha = 0.88$ [0.81, 0.93]; version B pre-training $\alpha = 0.67$ [0.49, 0.80],

post-training α = 0.83 [0.74, 0.89] (for detailed internal consistency values, see Supplementary Table 1).

Stress Induction: Anagram Task

An anagram task was used to induce psychosocial stress (based on van Bockstaele et al., 2020). The task was administered on a computer and was introduced as a verbal intelligence test. In total, 15 anagrams were presented in a predetermined order. All anagrams were solvable, yet extremely difficult within the time limit. Each anagram had to be solved within 20 seconds, with a countdown appearing after 10 seconds. The next anagram appeared once a response was entered or time expired. After the last anagram, participants received negative feedback, regardless of their actual performance: “You solved X of 15 anagrams. Your performance is well below average. On average, participants solve X+4 anagrams. Your performance corresponds to a percentile rank of 13%. This means that 87% were able to solve as many anagrams as you or more than you”. Additionally, a bogus, normal distribution graph was displayed marking participants’ percentile rank. To increase stress, participants were told that they would be recorded via a webcam (positioned in front of them) during the task. It was further explained that the recording would be analyzed later, and that this evaluation would also contribute to their overall performance score.

Stress Reactivity

Self-Report

An adapted version of the self-report questionnaire used by Becker et al. (2016) was employed to measure participants’ stress reactivity (SR). Specifically, participants had to rate the extent to which they currently experienced seven different feelings (“I feel… tense / sad / anxious / confident / relaxed / happy / relieved”) using a 5-point Likert scale (1 = “not at all”; 5 = “very”). The mean ratings of all scales were used as an overall score of self-reported SR, with positive items (i.e., confident, relaxed, happy, relieved) being inverted, such that higher scores reflect higher levels of self-reported SR. Internal consistency in the present sample was as following: baseline: α = 0.67 [0.56, 0.77]; pre-anagram task: α = 0.66 [0.54, 0.76]; post-anagram task: α = 0.81 [0.74, 0.86]; follow-up: α = 0.76 [0.68, 0.83].

HR and HRV

HR and HRV were assessed to index psychophysiological levels of arousal. An Einthoven Lead-II electrocardiogram (ECG) with a sampling rate of 1000 Hz (digitization 16 bit) was recorded continuously during the study. Therefore, Ag/AgCl electrodes were placed on the participant's lowest left rib within the axillar line as well as on the left and right clavicle. During acquisition, a BrainAmp ExG amplifier (Brain Products GmbH, Gilching, Germany) and BrainVision Recorder (Vers. 1.23.0003, Brain Products GmbH, Gilching, Germany) was used, and a 50 Hz notch filter was applied. In the offline preprocessing, signals underwent band-pass filtering (5–35 Hz, 24 dB/oct). Ectopic beats and recording artifacts were identified and corrected. After detrending, interbeat intervals were processed using an end-tapered Hamming window and subsequently transformed into the frequency domain via Fast Fourier transform. Using Kubios HRV software (Version 2.1), HRV parameters were then extracted (for details on HRV, see Adolph et al., 2025). Mean HR and HRV of specific 5-minute assessment time points were selected for the analysis. HR will be described and analyzed in bpm, and HRV in both the natural log transformed root mean square of successive inter-beat-interval differences (ln RMSSD) and the natural log transformed power in high-frequency range (ln HF-Power).

Salivary Cortisol

Stress levels were also assessed via the stress hormone cortisol. Using Salivette® sampling devices (Sarstedt, Nümbrecht, Germany), saliva was collected at several time points: baseline, pre-anagram, post-anagram, follow-up (for timings, see Supplementary Figure 2). Samples were stored in a freezer at -20 °C until analysis at the Genetic Psychology Lab at Ruhr University Bochum using a Chemoluminescence Immunoassay (CLIA; IBL International, Hamburg, Germany). Cortisol levels will be described and analyzed in nmol/l saliva. Intra- and inter-assay coefficients of variation were below 10%.

Post-Event Processing

To assess levels of state post-event processing (PEP) post-anagram task, the self-report questionnaire by Helbig-Lang et al. (2016) was employed. The PEP comprises three

items (e.g., “At the moment, I worry I could have left a bad impression”), reflecting negative post-event thoughts, particularly about perceived mistakes or embarrassment. Participants indicate their level of agreement to each statement using a 7-point Likert scale (0 = “do not agree”; 6 = “completely agree”). The mean scores across all items were used as overall score, with higher scores indicating higher levels of PEP. Internal consistency in the present sample was at post-anagram task α = 0.84 [0.76, 0.89] and at follow-up α = 0.65 [0.50, 0.76].

Randomization

A stratified randomization for the conditions’ assignment was used (positive vs. negative vs. neutral), with each participant being randomly assigned to one of the equally sized strata. The randomization sequences were generated using http://random.org and the randomization scheme itself was implemented in R (R Core Team, 2022) with fixed block lengths of 3. Participants were stratified by levels of social anxiety (low / high), age (younger / older), and gender/contraception (women & no pill / women & pill / men or diverse). Strata for social anxiety were determined by SPAI-G scores and the range we applied for recruitment (i.e., scores between 28 and 52). These scores represent the 25th and the 75th percentile of a sample of German medical students (Baldauf et al., 2014). The median score (i.e., 39) from this sample was used as the stratification’s cut-off (i.e., ≤39 vs. >39). Strata for age were determined by mean age (26.50) of the planned age range (between 18 and 35 years old, i.e., <27 vs ≥27).

Procedure

This study was approved by the local ethics committee of Ruhr University Bochum, Germany (ethical approval number: 582), and observed all local (data protection) laws. The data collection was conducted from October 2022 until May 2024 at Ruhr University Bochum. Participants who met all inclusion criteria following the online screening were invited to the lab session. To control for cortisol fluctuations, there were two possible timeslots for the lab sessions[1], that is, between 12:00 p.m. and 02:30 p.m. or 03:00 p.m. and

[1] Testing time was only considered as a factor in analyses that included salivary cortisol.

05:30 p.m. After obtaining written informed consent, participants started with the baseline assessment, including the first collection of saliva samples and the first self-reported SR ratings. Next, electrodes were applied and baseline ECG was recorded. This was followed by the pre-training IB assessment, the CBM-I training, and the post-training IB assessment. After that, the second collection of saliva samples and self-reported SR ratings followed as well as the second recording phase of ECG (i.e., pre-anagram). Next, the anagram task was completed, followed by the third collection of saliva samples and self-reported SR ratings, a third ECG recording, and the first application of the PEP questionnaire (i.e., post-anagram). This was followed by a 20-minute waiting period. After that, peak cortisol levels were assessed via the fourth saliva sample, alongside the fourth set of self-reported SR ratings, ECG recordings, and the second administration of the PEP questionnaire (i.e., follow-up, referred to as “recovery” in the pre-registration). Subsequently, participants’ awareness of the study’s aim was assessed. Finally, all were debriefed, and those in the negative CBM-I condition completed a brief positive CBM-I session to counteract induced negative IBs. Participants received 40€ or course credit for completing both the online screening and the lab session (for a study flow chart, see Supplementary Figure 2).

Power Analyses

The review of meta-analyses by Jones and Sharpe (2017) was used for a-priori power analyses since it specifically examined the effects of CBM-I in the context of emotional psychopathology. Overall effect sizes for changes in IBs varied between moderate and large effects. Hence, when adopting a rather conservative approach, a moderate effect size was expected (Cohen’s d = 0.70). The power calculation was based on pairwise comparisons between conditions, planned as follow-ups to initial ANOVAs, indicating that n = 45 participants per condition would provide ~80% power at p = .017 (Bonferroni-corrected). Therefore, the target sample size was N = 135. Due to difficulties in recruitment and limited funding, however, the study had to be terminated before reaching the target sample size (final overall sample: N = 87; with: positive n = 27, negative n = 31, neutral control n = 29).

Data Analyses

All experimental and analysis scripts and the anonymized datasets are available on the Open Science Framework (https://osf.io/nqab6/overview?view_only=9f69a149ae9848fca322ca716f0deb1c). Prior to data analyses, the assumptions of each statistical method were examined. Outliers were identified following Field et al. (2012). That is, data above the third quartile plus 1.5 times the interquartile range or data below the first quartile minus 1.5 times the interquartile range were considered as outliers. Analyses were calculated with and without the respective participants, as well as with and without participants who identified the real purpose of the study (see Supplements). Any deviations from the preregistration, for example due to assumption violations, are reported in footnotes and supplements together with the alternative methods applied. We used the standard $p < .05$ for significances and report effect sizes with the respective 95% confidence intervals. All included variables were assessed electronically – hence, missing data only occurred in case of technical issues.

A series of mixed analyses of variance (ANOVAs) were conducted to examine the effects of CBM-I training condition on different outcomes over time. Across all models, group (i.e., positive, negative, neutral) was included as between-subjects factor, and time (specific to each outcome, see below) as within-subjects factor. We expected to find time x group interactions, which would then be further analyzed via (Bonferroni-corrected) *t*-tests. Changes in IBs were analyzed with a mixed ANOVA including the two timepoints pre- and post-training. To investigate the effects of the stress task on self-reported SR and salivary cortisol, two separate mixed ANOVAs compared pre- and post-anagram task assessments. Additional mixed ANOVAs compared pre-anagram task and follow-up assessments, to test if participants recovered from stress. To examine stress reactivity in HR and HRV, two separate mixed ANOVAs included three timepoints (i.e., pre-, during- and post-anagram task). Additional mixed ANOVAs were conducted to assess recovery in HR and HRV after the anagram task (pre-anagram task vs. follow-up). Group differences in PEP post-anagram

task were analyzed using a one-way ANOVA. To explore changes in PEP, a mixed ANOVA was conducted using the timepoints post-anagram task and follow-up.

Results

Results of sensitivity analyses including examination of the effects of participant inclusion/exclusion can be found in footnotes and the supplements.

Baseline Sample Characteristics

No significant differences emerged between the three groups on any baseline variable, that is, age, gender, relationship status, SPAI-G, QIDS-SR, IBs, self-reported SR, salivary cortisol, HR or HRV (see Table 1).

Table 1

Baseline Variables (Full Sample and per Group)

Variable	Full Sample (N = 87)	Training group			Group comparisons
		Positive (n = 27)	Negative (n = 31)	Neutral (n = 29)	
Sociodemographic measures					
Age (years), *M* (*SD*)	22.29 (3.51)	23.30 (3.43)	21.55 (3.52)	22.14 (3.46)	$F(2, 84) = 1.87$, $p = .161$
Gender, *n* (%)					$\chi^2(2, N = 87) = 0.24$, $p = .885$
Women	60 (69.0)	19 (70.4)	22 (71.0)	19 (65.5)	
Men	27 (31.0)	8 (29.6)	9 (29.0)	10 (34.5)	
Nationality, *n* (%)					[a]
German	78 (89.7)	23 (85.2)	28 (90.3)	27 (93.1)	
Turkish	2 (2.3)	1 (3.7)	0 (0.0)	1 (3.4)	
Ukrainian	2 (2.3)	1 (3.7)	0 (0.0)	1 (3.4)	
Other	5 (5.7)	2 (7.4)	3 (9.7)	0 (0.0)	
Main occupation, *n* (%)					[a]
Student	78 (89.7)	23 (85.2)	27 (87.1)	28 (96.6)	
Employed	7 (8.0)	4 (14.8)	2 (6.5)	1 (3.4)	
Unemployed	2 (2.3)	0 (0.0)	2 (6.5)	0 (0.0)	
Relationship status, *n* (%)					$\chi^2(2, N = 87) = 2.35$, $p = .309$[b]
Single	50 (57.5)	14 (51.9)	16 (51.6)	20 (69.0)	
Partnership	35 (40.2)	12 (44.4)	14 (45.2)	9 (31.0)	
Married	2 (2.3)	1 (3.7)	1 (3.2)	0 (0.0)	
Symptom measures, *M* (*SD*)					
SPAI-G	40.37 (6.50)	42.39 (5.06)	39.33 (7.85)	39.61 (5.85)	$F(2, 84) = 1.95$, $p = .149$
QIDS-SR	4.92 (2.54)	5.33 (3.03)	4.68 (1.85)	4.79 (2.73)	$F(2, 84) = 0.53$, $p = .592$
Interpretation bias, *M* (*SD*)					
Encoding Recognition Task	0.16 (0.75)	0.21 (0.71)	0.00 (0.82)	0.29 (0.71)	$F(2, 84) = 1.18$, $p = .311$
Psychophysiological measures, *M* (*SD*)					

Self-reported stress reactivity	2.16 (0.42)	2.13 (0.34)	2.11 (0.42)	2.24 (0.48)	$F(2, 84) = 0.77$, $p = .467$
Heart rate (bpm)	80.89 (9.27)[c]	82.86 (9.86)[d]	79.40 (8.83)[e]	80.51 (9.14)[d]	$F(2, 80) = 1.00$, $p = .371$
HRV: ln RMSSD (ms)	3.52 (0.40)[c]	3.49 (0.38)[d]	3.50 (0.45)[e]	3.57 (0.37)[d]	$F(2, 80) = 0.33$, $p = .721$
HRV: ln HF-Power (ms^2)	6.27 (0.86)[c]	6.21 (0.80)[d]	6.12 (0.94)[e]	6.48 (0.80)[d]	$F(2, 80) = 1.34$, $p = .266$
Salivary cortisol (nmol/l)	4.08 (2.29)	3.97 (2.33)	3.76 (2.07)	4.53 (2.47)	$F(2, 84) = 0.91$, $p = .405$

Note. SPAI-G: Social Phobia and Anxiety Inventory, German version; QIDS-SR: Quick Inventory of Depressive Symptomatology, Self-Report; HRV: Heart rate variability; ln RMSSD: Natural log transformed root mean square of successive inter-beat-interval differences; ln HF-Power: Natural log transformed power in high-frequency range.
[a] Due to highly unbalanced distributions, no inferential statistics were conducted.
[b] Relationship status was coded as either "single" or "partnership" (including both participants in a relationship and those who were married).
[c] $n = 83$. [d] $n = 27$. [e] $n = 29$.

IBs

Results of the mixed ANOVA showed significant main effects of group and time. Most importantly, the group x time interaction was significant (for detailed statistics, see Table 2)[2]. Post-hoc between-group comparisons (Bonferroni-corrected) revealed that post-training both the positive ($p < .001$, $d = –1.49$ [–1.97, –1.01]) and the neutral group ($p < .001$, $d = –1.08$ [–1.54, –0.62]) reported stronger positive, socially-relevant interpretations than the negative group. However, there was no significant difference between the neutral and the positive group ($p = .171$, $d = –0.42$ [–0.85, 0.01]). Within-group comparisons (Bonferroni-corrected) revealed an increase in positive, socially-relevant interpretations in the positive ($p < .001$, $d = 0.56$ [0.33, 0.79]) and in the neutral group ($p = .043$, $d = 0.22$ [0.01, 0.44]). In contrast, the negative group showed a significant decrease ($p = .004$, $d = –0.32$ [–0.54, –0.10]) in positive interpretations from pre- to post-training. Overall, these findings suggest that the training was generally effective, with the positive training increasing and the negative training decreasing

[2] Removing three outliers from these analyses did not lead to significant changes in mixed ANOVA results. However, in the neutral group, the (Bonferroni-corrected) within-group comparison did not reach significance anymore ($p = .064$, $d = 0.21$ [–0.01, 0.43]). We additionally conducted nonparametric ANOVAs with and without outliers using the aligned rank transform to account for deviations from normality in one cell. In contrast to the previous analysis (i.e., ANOVA without outliers), in both analyses, the previously observed main effect of time was no longer significant. Bonferroni-corrected post-hoc tests showed no changes in within-group comparisons, while the negative vs. neutral group comparison at post-training no longer reached significance.

positive interpretations, albeit without a statistically significant difference between the positive and neutral conditions (see Figure 2).[3]

Stress Reactivity

Self-Report

The mixed ANOVA examining changes in self-reported SR pre- to post-anagram task revealed no significant main effect of group and no significant group x time interaction. However, there was a significant main effect of time, with means indicating that all participants reported higher levels of stress post- compared to pre-anagram task. Comparing self-reported SR pre-anagram task vs. follow-up revealed a similar pattern, that is, neither the main effect of group nor the group x time interaction was significant. However, the main effect of time was significant, indicating that all participants reported lower levels of stress at follow-up compared to pre-anagram task (for detailed statistics, see Table 2). In sum, these results indicate that self-reported SR increased after the anagram task and decreased at follow-up across all participants (see Figure 3).

HR and HRV

Due to poor data quality, data from two participants (entire datasets) and from four additional timepoints across different participants were excluded from the respective analyses. Outliers in HR and HRV data were identified, but exclusion did not lead to significantly different results. Therefore, the results of the sample with outliers are reported here.

[3] We further examined the group x time interaction via comparing change scores between the conditions via a one-way ANOVA, which yielded a significant group effect on change in IBs, $F(2, 84) = 17.12$, $p < .001$, $\eta^2 = .29$ [.13, .43]. Bonferroni-corrected post-hoc tests showed that the negative group changed more strongly in IBs compared to both the neutral ($p = .002$, $d = –0.77$ [–1.22, –0.33]) and the positive group ($p < .001$, $d = –1.26$ [–1.73, –0.79]). The difference between the neutral and the positive group was not significant ($p = .079$, $d = –0.49$ [–0.93, –0.06]). Removing three outliers from these analyses did not lead to significant changes in results.

Table 2

Means and Standard Deviations of All Measured Variables and Results Across All Mixed Analyses of Variance (ANOVA)

Variable	Training group			Contrast	Mixed ANOVA		
	Full sample (N = 87)	Positive (n = 27)	Negative (n = 31)	Neutral (n = 29)		Main effect: Group	Main effect: Time	Interaction effect
Interpretation bias: Encoding Recognition Task, *M* (*SD*)								
Pre-training	0.16 (0.75)	0.21 (0.71)	0.00 (0.82)	0.29 (0.71)	Pre vs. post	$F(2, 84) = 12.81$, $p < .001$, $\eta^2_G = .18$ [.05, .32]	$F(1, 84) = 6.72$, $p = .011$, $\eta^2_G = .02$ [.00, .12]	$F(2, 84) = 17.12$, $p < .001$, $\eta^2_G = .10$ [.01, .22]*
Post-training	0.36 (1.00)	1.01 (0.77)	–0.43 (0.89)	0.60 (0.72)				
Self-reported stress reactivity, *M* (*SD*)								
Pre-anagram	2.26 (0.48)	2.20 (0.50)	2.36 (0.47)	2.21 (0.47)	Pre vs. post	$F(2, 84) = 0.59$, $p = .554$	$F(1, 84) = 91.56$, $p < .001$, $\eta^2_G = .22$ [.08, .36]	$F(2, 84) = 0.23$, $p = .797$
Post-anagram	2.87 (0.68)	2.87 (0.74)	2.93 (0.66)	2.82 (0.68)				
Follow-up	2.13 (0.56)	2.21 (0.58)	2.09 (0.57)	2.11 (0.54)	Pre vs. follow-up	$F(2, 84) = 0.17$, $p = .846$	$F(1, 84) = 5.03$, $p = .028$, $\eta^2_G = .01$[.00, .10]	$F(2, 84) = 2.33$, $p = .103$
Heart rate (bpm), *M* (*SD*)								
Pre-anagram	77.42 (9.32)[a]	79.69 (9.20)[c]	75.41 (8.63)[e]	77.31 (9.91)[e]	Pre vs. during vs. post[e]	$F(2, 80) = 1.87$, $p = .161$	$F(1.49, 119.47) = 16.65$, $p < .001$, $\eta^2_G = .01$ [.00, .06]	$F(2.99, 119.47) = 0.05$, $p = .985$
Anagram	78.99 (10.00)[b]	81.50 (8.92)[d]	76.96 (9.74)[f]	78.71 (10.98)[e]				
Post-anagram	76.19 (8.70)[a]	78.44 (8.29)[c]	74.04 (7.78)[e]	76.24 (9.66)[e]	Pre vs. follow-up	$F(2, 82) = 1.48$, $p = .233$	$F(1, 82) = 15.41$, $p < .001$, $\eta^2_G = .01$ [.00, .08]	$F(2, 82) = 0.38$, $p = .685$
Follow-up	75.85 (8.98)[a]	78.11 (8.98)[c]	74.26 (8.47)[e]	75.32 (9.36)[e]				
Heart rate variability: ln RMSSD (ms), *M* (*SD*)								
Pre-anagram	3.67 (0.41)[a]	3.58 (0.35)[c]	3.71 (0.43)[e]	3.71 (0.44)[e]	Pre vs. during vs. post[e]	$F(2, 80) = 1.10$, $p = .337$	$F(1.80, 144.39) = 10.21$, $p < .001$, $\eta^2_G = .01$[.00, .06]	$F(3.61, 144.39) = 0.36$ $p = .815$
Anagram	3.61 (0.42)[b]	3.51 (0.36)[d]	3.61 (0.44)[f]	3.69 (0.44)[e]				
Post-anagram	3.72 (0.40)[a]	3.65 (0.33)[c]	3.74 (0.44)[e]	3.78 (0.42)[e]				

Follow-up	3.75 (0.41)[a]	3.65 (0.32)[c]	3.77 (0.46)[e]	3.81 (0.42)[e]	Pre vs. follow-up	$F(2, 82) = 1.17$, $p = .316$	$F(1, 82) = 8.86$, $p = .004$, $\eta^2_G = .01$ [.00, .09]	$F(2, 82) = 0.32$, $p = .726$
Heart rate variability: ln HF-Power (ms^2), *M* (*SD*)								
Pre-anagram	6.47 (0.82)[a]	6.32 (0.70)[c]	6.46 (0.90)[e]	6.62 (0.84)[e]	Pre vs. during vs. post[e]	$F(2, 80) = 1.28$, $p = .284$	$F(1.87, 149.95) = 6.83$, $p = .002$, $\eta^2_G = .01$[.00, .05]	$F(3.75, 149.95) = 0.38$, $p = .810$
Anagram	6.37 (0.88)[b]	6.20 (0.84)[d]	6.29 (0.93)[f]	6.60 (0.86)[e]				
Post-anagram	6.58 (0.83)[a]	6.49 (0.69)[c]	6.51 (0.94)[e]	6.75 (0.83)[e]	Pre vs. follow-up	$F(2, 82) = 1.17$, $p = .316$	$F(1, 82) = 4.00$, $p = .049$, $\eta^2_G = .01$ [.00, .07]	$F(2, 82) = 0.11$, $p = .897$
Follow-up	6.58 (0.83)[a]	6.44 (0.56)[c]	6.54 (0.99)[e]	6.76 (0.87)[e]				
Salivary cortisol (nmol/l), *M* (*SD*)								
Pre-anagram	2.89 (1.65)	3.02 (1.62)	2.64 (1.88)	3.02 (1.44)	Pre vs. post	$F(2, 84) = 0.29$, $p = .747$	$F(1, 84) = 23.59$, $p < .001$, $\eta^2_G = .01$ [.00, .08]	$F(2, 84) = 1.30$, $p = .279$
Post-anagram	2.65 (1.46)	2.76 (1.42)	2.52 (1.71)	2.67 (1.21)				
Follow-up	2.38 (1.29)	2.55 (1.60)	2.18 (1.01)	2.44 (1.26)	Pre vs. follow-up	$F(2, 84) = 0.50$, $p = .606$	$F(1, 84) = 18.36$, $p < .001$, $\eta^2_G = .03$ [.00, .18]	$F(2, 84) = 0.13$, $p = .878$
Post-event processing, *M* (*SD*)								
Post-anagram	2.75 (1.62)	2.75 (1.63)	2.70 (1.74)	2.80 (1.51)	Post	$F(2, 84) = 0.03$, $p = .969$	–	–
Follow-up	1.49 (1.12)	1.56 (1.25)	1.38 (1.12)	1.54 (1.02)	Post vs. follow-up	$F(2, 84) = 0.10$, $p = .906$	$F(1, 84) = 101.49$, $p < .001$, $\eta^2_G = .17$ [.05, .32]	$F(2, 84) = 0.08$, $p = .920$

Note. ln RMSSD: Natural log transformed root mean square of successive inter-beat-interval differences; ln HF-Power: Natural log transformed power in high-frequency range.
[a] $n = 85$. [b] $n = 83$. [c] $n = 27$. [d] $n = 26$. [e] $n = 29$. [f] $n = 28$.
[e] Greenhouse-Geisser corrected.

HR. The mixed ANOVA examining HR across the anagram task (i.e., pre-, during-, post-anagram task) revealed no significant main effect of group and no significant group x time interaction. However, there was a significant main effect of time. Post-hoc comparisons (Bonferroni-corrected) indicated that HR of all participants significantly increased from pre- to during-anagram task (p = .045, d = –0.28 [–0.50, –0.05]) and decreased from during- to post-anagram task (p < .001, d = 0.62 [0.38, 0.86]) and from pre- to post-anagram task (p < .001, d = 0.46 [0.23, 0.69]). Comparing HR pre-anagram task vs. follow-up revealed a similar pattern, that is, neither the main effect of group nor the group x time interaction was significant. However, the main effect of time was significant, with means indicating that all participants showed lower HR at the follow-up timepoint compared to pre-anagram task (for detailed statistics, see Table 2). To summarize, HR increased during-anagram task, decreased post-anagram task, and fell below baseline at follow-up across all participants.

HRV. The mixed ANOVA examining HRV (i.e., both ln RMSSD and ln HF-Power) over the course of the task (i.e., pre-, during-, post-anagram task) revealed no significant main effect of group and no significant group x time interaction. However, there were significant main effects of time for both indices. Post-hoc comparisons (Bonferroni-corrected) indicated that only ln RMSSD showed a significant increase from pre- to post-anagram task (p = .049, d = –0.27 [–0.50, –0.05]), but not ln HF-Power (p = .120, d = d = –0.23 [–0.45, –0.01]). Neither ln RMSSD (p = .160, d = 0.22 [0.00, 0.44]) nor ln HF-Power (p = .456, d = 0.16 [–0.06, 0.38]) showed significant changes from pre- to during-anagram task. However, from during- to post-anagram task, both indices significantly increased (ln RMSSD: p < .001, d = –0.56 [–0.79, –0.32]; ln HF-Power: p < .001, d = –0.47 [–0.70, –0.23]). Comparing HRV pre-anagram task vs. follow-up revealed a similar pattern, that is, neither the main effect of group nor the group x time interaction was significant. However, the main effect of time was significant, with means indicating that average HRV levels were higher at follow-up compared to pre-anagram task levels for all participants (for means and detailed statistics, see Table 2). In sum, these results indicate that HRV increased, particularly from during- to post-anagram task, and was highest at follow-up across all participants.

Salivary Cortisol

The mixed ANOVA examining changes in cortisol pre- to post-anagram task revealed no significant main effect of group and no significant group x time interaction. However, a significant main effect of time was found, with means indicating that, in contrast to our expectations, salivary cortisol levels decreased (rather than increased) over time. A similar pattern emerged for the pre-anagram task vs. follow-up comparison; both the main effect of group and the group x time interaction did not reach significance. However, a main effect of time was observed, with means suggesting that salivary cortisol levels decreased over time. In sum, these results indicate that salivary cortisol levels consistently decreased over time.

Testing time was included as a factor in both analyses and showed consistent main effects. Cortisol levels were significantly higher in early compared to late sessions, $F(1, 81) = 7.64$, $p = .007$, $\eta^2_G = .05$ [.00, .16] (pre- vs. post-anagram task), and $F(1, 81) = 10.72$, $p = .002$, $\eta^2_G = .06$ [.00, .19] (pre-anagram task vs. follow-up). However, these effects did not interact with group or timepoint, indicating that diurnal variation did not confound the results.

PEP

Results of the one-way ANOVA comparing PEP scores post-anagram task showed no significant effect of group, indicating that PEP scores did not differ between conditions. The mixed ANOVA examining changes in PEP post-anagram task to follow-up revealed neither a significant main effect of group and nor a significant group x time interaction. However, a significant main effect of time was found, with means indicating that PEP scores decreased over time across all participants.

Exploratory Analyses

Anagram task

Overall, participants solved $M = 33.33\%$ ($SD = 13.37$) of all fifteen anagrams. No significant differences, $F(2, 84) = 1.56$, $p = .217$, emerged in the percentage of correctly solved anagrams between the positive ($M = 36.80\%$, $SD = 14.70$), negative ($M = 32.90\%$, $SD = 10.50$) and neutral group ($M = 30.60\%$ $SD = 14.50$).

Interrelationships Among Measures

Baseline data were used to examine correlations and the relationships among tasks and measures. None of the correlations were significant, except for the following four: SPAI-G and ln HF-Power; HR and ln RMSSD; HR and ln HF-Power; ln RMSSD and ln HF-Power (for an overview, see Table 3). Correlation coefficients indicate that higher levels of social anxiety were associated with increased HRV levels (i.e., ln HF-Power), while higher HR was associated with lower levels of both HRV indices. Notably, the two HRV indices showed a strong positive correlation.

Table 3

Correlations Between Baseline Variables

	1.	2.	3.	4.	5.	6.	7.
1. SPAI-G[a]	–						
2. QIDS-SR[a]	–.02	–					
3. IBs[a]	–.09	.16	–				
4. Self-reported SR[a]	.05	.21	–.16	–			
5. Salivary cortisol[a]	.09	.09	–.07	–.03	–		
6. Heart rate[b]	–.07	–.03	.11	–.03	.11	–	
7. Heart rate variability (ln RMSSD)[b]	.20	.09	–.08	.13	–.03	–.62***	–
8. Heart rate variability (ln HF-Power)[b]	.23*	.06	–.02	.08	.01	–.36***	.89***

Note. SPAI-G: Social Phobia and Anxiety Inventory, German version; QIDS-SR: Quick Inventory of Depressive Symptomatology Self-Report; IBs: Pre-training interpretation biases; Self-reported SR: Self-reported stress reactivity; ln RMSSD: Natural log transformed root mean square of successive inter-beat-interval differences; ln HF-Power: Natural log transformed power in high-frequency range.
[a]n = 87. [b]n = 83.
* p < .05. *** p < .001.

Discussion

The present study aimed to provide a rigorous proof-of-principle investigation into the mechanisms of cognitive bias modification for interpretation (CBM-I) in the context of social anxiety. Specifically, we examined the effects of three training conditions (positive, negative, and neutral) on interpretation biases (IBs) in individuals with moderate levels of social anxiety. A second aim was to assess whether CBM-I effects would transfer to stress reactivity, measured via self-reported and psychophysiological responses before, during, and after a stress task (i.e., an anagram task). Finally, we examined whether the training would influence levels of post-event processing (PEP) following the stress task.

Overall, results for the first aim were as expected. That is, post-hoc tests revealed that the positive group reported more positive socially-relevant interpretations than the negative group after the training. In addition, the positive group showed a within-group increase in positive, while the negative group showed a within-group decrease in positive interpretations from pre- to post-training. The sham training group showed a modest increase in positive interpretations, albeit not statistically different from the positive group. Given its neutral content, this likely reflects time or repeated testing effects and underscores the value of a neutral group as a benchmark for optimally evaluating the training's effect (see Blackwell et al., 2017). However, when statistical outliers were excluded, the increase in positive interpretations in the neutral group was no longer statistically significant (see Supplements), suggesting that the observed effect may not be reliable.

The second aim was to examine CBM-I transfer to stress reactivity, given that IBs are linked to psychophysiological responses. We assessed self-reported stress, HR, HRV, and cortisol, but contrary to expectations, no differential changes were found between the three training groups pre- to post-anagram task. Instead, we mainly observed main effects of time, with all participants showing higher self-reported stress and increased HR from pre- to post-anagram task. Although the study ended before reaching the targeted sample size, the achieved sample was sufficient to address the primary questions. Robust group differences in IBs, particularly between positive and negative groups, demonstrate that the intervention effectively modified biases. The lack of effects on stress-related outcomes therefore most likely reflects limited generalization of induced IBs beyond bias measures, rather than insufficient statistical power. Second, the anagram task may have been a limiting factor. Self-reported stress was relatively low, suggesting that the task did not induce a strong enough stress response. Hence, without a robust physiological response, detecting potential modulation from CBM-I is difficult. Third, the stress induced by the anagram task may not have aligned with CBM-I content. While the training focused on social scenarios, the anagram task emphasized performance stress, lacking a genuine socio-evaluative component.

The potential mismatch between the cognitive content targeted during training and the cognitive-emotional processes elicited by the stress task warrants more detailed discussion and underscores the complexity of designing studies that accurately capture transfer effects. While it is often assumed that CBM-I is most effective when there is a clear match between the trained material and stressor, this assumption does not consistently hold across studies. Notably, even though Turton et al. (2018) employed a disorder-specific stressor, a test meal tailored to individuals with anorexia nervosa, they likewise found no evidence of transfer effects. Regarding HR and HRV, several studies with modest sample sizes have observed transfer effects of CBM-I on stress-related outcomes. For example, van Bockstaele et al. (2020) conducted a single-session CBM-I intervention with *N* = 71 participants and used an anagram task to induce stress afterwards. While no differences emerged in self-reported stress reactivity between the positive and the negative condition directly after the anagram task, participants who received positive CBM-I showed improved stress recovery, as indicated by HR measures, compared to negative CBM-I. Similarly, Rozenman et al. (2020) found that participants in the CBM-I group exhibited enhanced physiological stress regulation, that is, reduced electrodermal activity and HR, compared to the control group. Notably, this study involved a 12-session CBM-I protocol and used a speech task as stressor. Along these lines, Nowakowski et al. (2015) found that positive CBM-I led to reduced HR during recovery from stress; however, a comparable recovery effect was also observed in a control group that completed a 50:50 training. In sum, these findings suggest that under certain conditions CBM-I does reduce psychophysiological stress responses.

Finally, despite theoretical and correlational evidence suggesting a link between IBs and PEP in social anxiety (e.g., Badra et al., 2017; Brozovich & Heimberg, 2008), the present study did not find differential changes in PEP as a function of CBM-I condition. However, immediately after the stressor (i.e., the anagram task), PEP levels were comparable to those observed in a clinical sample diagnosed with SAD (Helbig-Lang et al., 2016) and decreased significantly at follow-up across all conditions. This pattern may reflect

natural recovery or habituation rather than training-specific effects. The lack of group differences may also stem from limited generalization of CBM-I to participants' appraisals of performance, which drive PEP. While Badra et al. (2017) identified PEP as a mediator between social anxiety and IBs, their findings were correlational. Our results highlight the complexity of the IB–PEP relationship and suggest that longer, more immersive, or PEP-targeted interventions may be needed to influence this process effectively.

These findings, together with existing evidence, highlight key limitations and considerations for advancing CBM-I research. First, CBM-I effects were only observed on the closely matched ERT measure. Without additional IB measures, we cannot rule out task-specific adaptations rather than genuine cognitive shifts (see also Duken et al., 2025). If true cognitive change occurred, the lack of effects on stress reactivity suggests IB changes may not automatically reduce stress responses. Future studies should include multiple, independent IB measures to better assess the scope and downstream effects of CBM-I. Second, the CBM-I and stress induction should be closely aligned, using real-world or ecologically valid tasks that match trained cognitive themes. For example, the Trier Social Stress Test (Kirschbaum et al., 1993) reliably elicits robust physiological and psychological responses through public speaking and mental arithmetic before an evaluative audience, making it highly relevant for social anxiety studies. A third consideration is that cortisol may not be sensitive enough to capture rapid stress responses. In contrast, salivary alpha-amylase is increasingly recognized as a more immediate and sensitive indicator of autonomic nervous system activation during acute stress. Unlike cortisol, which peaks 20–30 minutes post-stressor, alpha-amylase rises within minutes, reflecting fast sympathetic arousal during stress tasks, and thus may be a more sensitive marker in this context (e.g., Nater & Rohleder, 2009). A final consideration is the range of psychophysiological markers used to assess stress reactivity. While we measured HR, HRV, and cortisol, other biomarkers such as skin conductance response, pupillometry, electromyography, and event-related potentials (ERPs), are also highly relevant (Abado et al., 2023). For example, pupillometry indexes cognitive and emotional load, and ERPs provide precise neural timing

of evaluative processes. Incorporating such markers in future CBM-I studies could better elucidate how training impacts autonomic and central nervous system responses to stress. In fact, a number of studies have already begun using such markers. For example, several clinical investigations have examined ERPs, particularly the N400 and P600 components, as potential markers of biased interpretational processing in mental disorders and symptoms such as worry and social anxiety (e.g., Feng et al., 2019, 2020; Moser et al., 2008). Overall, these studies have found associations between N400 and P600 amplitudes and biased processing. However, the strength and nature of these associations appear to depend on the type of assessment task used. Moreover, studies assessing pre–post changes in ERPs following interpretation training are scarce, and those that do exist have not consistently produced the expected results (e.g., Feng et al., 2020). To conclude, more fine-tuned sets of stress markers may be required, specifically tailored to the context of CBM-I.

In conclusion, this study provides evidence that CBM-I can successfully modify IBs in individuals with moderate social anxiety, yet modification was evident only in the negative, but not the positive, direction compared to a neutral control condition. However, these cognitive changes did not transfer to self-reported or psychophysiological stress responses (HR, HRV, cortisol) or PEP. Transfer within a subclinical sample may require more intensive training, better alignment between training and stressors, stronger stress induction, or alternative biomarkers. Future research should address these factors to clarify how CBM-I influences cognition and stress physiology.

References

Abado, E., Okon-Singer, H., & Aue, T. (2023). Neurophysiological mechanisms underlying cognitive biases to emotional information: Latest developments and new directions. *Biological Psychology*, *177*, 108486. https://doi.org/10.1016/j.biopsycho.2023.108486

Adolph, D., Zhang, X. C., Teismann, T., Wannemüller, A., & Margraf, J. (2025). Respiratory sinus arrhythmia—common and distinct mechanisms of emotional adjustment in the Depressive and Anxiety Disorders Spectrum? *Psychophysiology*, *62*(6), e70079. https://doi.org/10.1111/psyp.70079

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5.* (5th ed.). American Psychiatric Association.

Badra, M., Schulze, L., Becker, E. S., Vrijsen, J. N., Renneberg, B., & Zetsche, U. (2017). The association between ruminative thinking and negative interpretation bias in social anxiety. *Cognition and Emotion*, *31*(6), 1234–1242. https://doi.org/10.1080/02699931.2016.1193477

Baldauf, M., Thomas, A., & Strauß, B. (2014). Häufigkeit sozialphobischer Symptome und ihr Bezug zu interpersonalen Merkmalen in einer Stichprobe von Medizinstudierenden. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, *64*(02), 76–81. https://doi.org/10.1055/s-0033-1358719

Becker, E. S., Ferentzi, H., Ferrari, G., Möbius, M., Brugman, S., Custers, J., Geurtzen, N., Wouters, J., & Rinck, M. (2016). Always approach the bright side of life: A general positivity training reduces stress reactions in vulnerable individuals. *Cognitive Therapy and Research*, *40*(1), 57–71. https://doi.org/10.1007/s10608-015-9716-2

Blackwell, S. E., Woud, M. L., & MacLeod, C. (2017). A question of control? Examining the role of control conditions in experimental psychopathology using the example of cognitive bias modification research. *The Spanish Journal of Psychology*, *20*. https://doi.org/10.1017/sjp.2017.41

Brozovich, F., & Heimberg, R. G. (2008). An analysis of post-event processing in social anxiety disorder. *Clinical Psychology Review*, *28*(6), 891–903. https://doi.org/10.1016/j.cpr.2008.01.002

Chen, J., Short, M., & Kemps, E. (2020). Interpretation bias in social anxiety: A systematic review and meta-analysis. *Journal of Affective Disorders*, *276*, 1119–1130. https://doi.org/10.1016/j.jad.2020.07.121

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). The Guilford Press.

Clarke, P. J. F., Branson, S., Chen, N. T. M., van Bockstaele, B., Salemink, E., MacLeod, C., & Notebaert, L. (2017). Attention bias modification training under working memory load increases the magnitude of change in attentional bias. *Journal of Behavior Therapy and Experimental Psychiatry*, *57*, 25–31. https://doi.org/10.1016/j.jbtep.2017.02.003

Collins, A., Scott, R. B., Hirsch, C. R., Ottaviani, C., Krahé, C., & Meeten, F. (2022). A systematic review of the literature on interpretation bias and its physiological correlates. *Biological Psychology*, *173*, 108398. https://doi.org/10.1016/j.biopsycho.2022.108398

Condren, R. (2002). HPA axis response to a psychological stressor in generalised social phobia. *Psychoneuroendocrinology*, *27*(6), 693–703. https://doi.org/10.1016/S0306-4530(01)00070-1

Duken, S. B., Moriya, J., Hirsch, C., Woud, M. L., van Bockstaele, B., & Salemink, E. (2025). Reliability and validity of four cognitive interpretation bias measures in the context of social anxiety. *Behavior Research Methods*, *57*, 48. https://doi.org/10.3758/s13428-024-02576-0

Feng, Y.-C., Krahé, C., Meeten, F., Sumich, A., Mok, C. L. M., & Hirsch, C. R. (2020). Impact of imagery-enhanced interpretation training on offline and online interpretations in

worry. *Behaviour Research and Therapy*, *124*, 103497. https://doi.org/10.1016/j.brat.2019.103497

Feng, Y.-C., Krahé, C., Sumich, A., Meeten, F., Lau, J. Y. F., & Hirsch, C. R. (2019). Using event-related potential and behavioural evidence to understand interpretation bias in relation to worry. *Biological Psychology*, *148*, 107746. https://doi.org/10.1016/j.biopsycho.2019.107746

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.

Fodor, L. A., Georgescu, R., Cuijpers, P., Szamoskozi, Ş., David, D., Furukawa, T. A., & Cristea, I. A. (2020). Efficacy of cognitive bias modification interventions in anxiety and depressive disorders: A systematic review and network meta-analysis. *The Lancet Psychiatry*, *7*(6), 506–514. https://doi.org/10.1016/S2215-0366(20)30130-9

Fydrich, T. (2016). SPAI – Soziale Phobie und Angstinventar. In K. Geue, B. Strauß, & E. Brähler (Eds.), *Diagnostische Verfahren in der Psychotherapie* (3., überarbeitete und erweiterte Auflage). Hogrefe. https://doi.org/10.1026/02700-0000

Helbig-Lang, S., von Auer, M., Neubauer, K., Murray, E., & Gerlach, A. L. (2016). Post-event processing in social anxiety disorder after real-life social situations – An ambulatory assessment study. *Behaviour Research and Therapy*, *84*, 27–34. https://doi.org/10.1016/j.brat.2016.07.003

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. *Annual Review of Clinical Psychology*, *12*(1), 281–305. https://doi.org/10.1146/annurev-clinpsy-021815-093436

Hollocks, M. J., Pickles, A., Howlin, P., & Simonoff, E. (2016). Dual cognitive and biological correlates of anxiety in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *46*(10), 3295–3307. https://doi.org/10.1007/s10803-016-2878-2

Jones, E. B., & Sharpe, L. (2017). Cognitive bias modification: A review of meta-analyses. *Journal of Affective Disorders*, *223*, 175–183. https://doi.org/10.1016/j.jad.2017.07.034

Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 617–627. https://doi.org/10.1001/archpsyc.62.6.617

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*(1–2), 76–81. https://doi.org/10.1159/000119004

Klumbies, E., Braeuer, D., Hoyer, J., & Kirschbaum, C. (2014). The reaction to social stress in social phobia: Discordance between physiological and subjective parameters. *PLOS ONE*, *9*(8), e105670. https://doi.org/10.1371/journal.pone.0105670

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Liu, H., Li, X., Han, B., & Liu, X. (2017). Effects of cognitive bias modification on social anxiety: A meta-analysis. *PLOS ONE*, *12*(4), e0175107. https://doi.org/10.1371/journal.pone.0175107

MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology*, *8*(1), 189–217. https://doi.org/10.1146/annurev-clinpsy-032511-143052

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology*, *109*(4), 602–615. https://doi.org/10.1037/0021-843X.109.4.602

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, *1*(1), 167–195. https://doi.org/10.1146/annurev.clinpsy.1.102803.143916

Meeten, F. (2017). *Modification of interpretation biases in worry: An examination of cognitive and physiological responses.* [Unpublished thesis]. King's College London.

Morrison, A. S., & Heimberg, R. G. (2013). Social anxiety and social anxiety disorder. *Annual Review of Clinical Psychology*, *9*(1), 249–274. https://doi.org/10.1146/annurev-clinpsy-050212-185631

Moser, J. S., Hajcak, G., Huppert, J. D., Foa, E. B., & Simons, R. F. (2008). Interpretation bias in social anxiety as detected by event-related brain potentials. *Emotion*, *8*(5), 693–700. https://doi.org/10.1037/a0013173

Nater, U. M., & Rohleder, N. (2009). Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: Current state of research. *Psychoneuroendocrinology*, *34*(4), 486–496. https://doi.org/10.1016/j.psyneuen.2009.01.014

Nowakowski, M. E., Antony, M. M., & Koerner, N. (2015). Modifying interpretation biases: Effects on symptomatology, behavior, and physiological reactivity in social anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, *49*, 44–52. https://doi.org/10.1016/j.jbtep.2015.04.004

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rozenman, M., Gonzalez, A., Logan, C., & Goger, P. (2020). Cognitive bias modification for threat interpretations: Impact on anxiety symptoms and stress reactivity. *Depression and Anxiety*, *37*(5), 438–448. https://doi.org/10.1002/da.23018

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., & Keller, M. B. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, *54*(5), 573–583. https://doi.org/10.1016/S0006-3223(02)01866-8

Salemink, E., & van den Hout, M. (2010). Validation of the “recognition task” used in the training of interpretation biases. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(2), 140–144. https://doi.org/10.1016/j.jbtep.2009.11.006

Salemink, E., van den Hout, M., & Kindt, M. (2007). Trained interpretive bias and anxiety. *Behaviour Research and Therapy*, *45*(2), 329–340. https://doi.org/10.1016/j.brat.2006.03.011

Salemink, E., van den Hout, M., & Kindt, M. (2009). Effects of positive interpretive bias modification in highly anxious individuals. *Journal of Anxiety Disorders*, *23*(5), 676–683. https://doi.org/10.1016/j.janxdis.2009.02.006

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar De Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., & Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry*, *27*(1), 281–295. https://doi.org/10.1038/s41380-021-01161-7

Turner, S. M., Beidel, D. C., Dancu, C. V., & Stanley, M. A. (1989). An empirically derived inventory to measure social fears and anxiety: The Social Phobia and Anxiety Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *1*(1), 35–40. https://doi.org/10.1037/1040-3590.1.1.35

Turton, R., Cardi, V., Treasure, J., & Hirsch, C. R. (2018). Modifying a negative interpretation bias for ambiguous social scenarios that depict the risk of rejection in women with anorexia nervosa. *Journal of Affective Disorders*, *227*, 705–712. https://doi.org/10.1016/j.jad.2017.11.089

van Bockstaele, B., Clarke, P. J. F., Notebaert, L., MacLeod, C., & Salemink, E. (2020). Effects of cognitive load during interpretation bias modification on interpretation bias and stress reactivity. *Journal of Behavior Therapy and Experimental Psychiatry*, *68*, 101561. https://doi.org/10.1016/j.jbtep.2020.101561

World Health Organization. (2023, September 27). *Anxiety disorders*. https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders

Woud, M. L. (Ed.). (2023). *Interpretational processing biases in emotional psychopathology: From experimental investigation to clinical practice*. Springer International Publishing. https://doi.org/10.1007/978-3-031-23650-1

Woud, M. L., Blackwell, S. E., Shkreli, L., Würtz, F., Cwik, J. C., Margraf, J., Holmes, E. A., Steudte-Schmiedgen, S., Herpertz, S., & Kessler, H. (2021). The effects of modifying dysfunctional appraisals in posttraumatic stress disorder using a form of cognitive bias modification: Results of a randomized controlled trial in an inpatient setting. *Psychotherapy and Psychosomatics*, *90*(6), 386–402. https://doi.org/10.1159/000514166

Woud, M. L., Holmes, E. A., Postma, P., Dalgleish, T., & Mackintosh, B. (2012). Ameliorating intrusive memories of distressing experiences using computerized reappraisal training. *Emotion*, *12*(4), 778–784. https://doi.org/10.1037/a0024992

Woud, M. L., Postma, P., Holmes, E. A., & Mackintosh, B. (2013). Reducing analogue trauma symptoms by computerized reappraisal training – Considering a cognitive prophylaxis? *Journal of Behavior Therapy and Experimental Psychiatry*, *44*(3), 312–315. https://doi.org/10.1016/j.jbtep.2013.01.003

Woud, M. L., Zlomuzica, A., Cwik, J. C., Margraf, J., Shkreli, L., Blackwell, S. E., Gladwin, T. E., & Ehring, T. (2018). Effects of appraisal training on responses to a distressing autobiographical event. *Journal of Anxiety Disorders*, *56*, 26–34. https://doi.org/10.1016/j.janxdis.2018.03.010

Würtz, F., Kunna, M., Blackwell, S. E., Lindgraf, C., Abado, E., Amanvermez, Y., Margraf, J., Everaert, J., & Woud, M. L. (in press). Interpretation biases in anxiety – A three-level meta-analysis. *Clinical Psychological Science*, https://doi.org/10.31234/osf.io/7zkvr_v2

Yiend, J., Lee, J.-S., Tekes, S., Atkins, L., Mathews, A., Vrinten, M., Ferragamo, C., & Shergill, S. (2014). Modifying interpretation in a clinically depressed sample using 'Cognitive Bias Modification-Errors': A double blind randomised controlled trial. *Cognitive Therapy and Research*, *38*(2), 146–159. https://doi.org/10.1007/s10608-013-9571-y

Figure 1

CONSORT Flow Chart

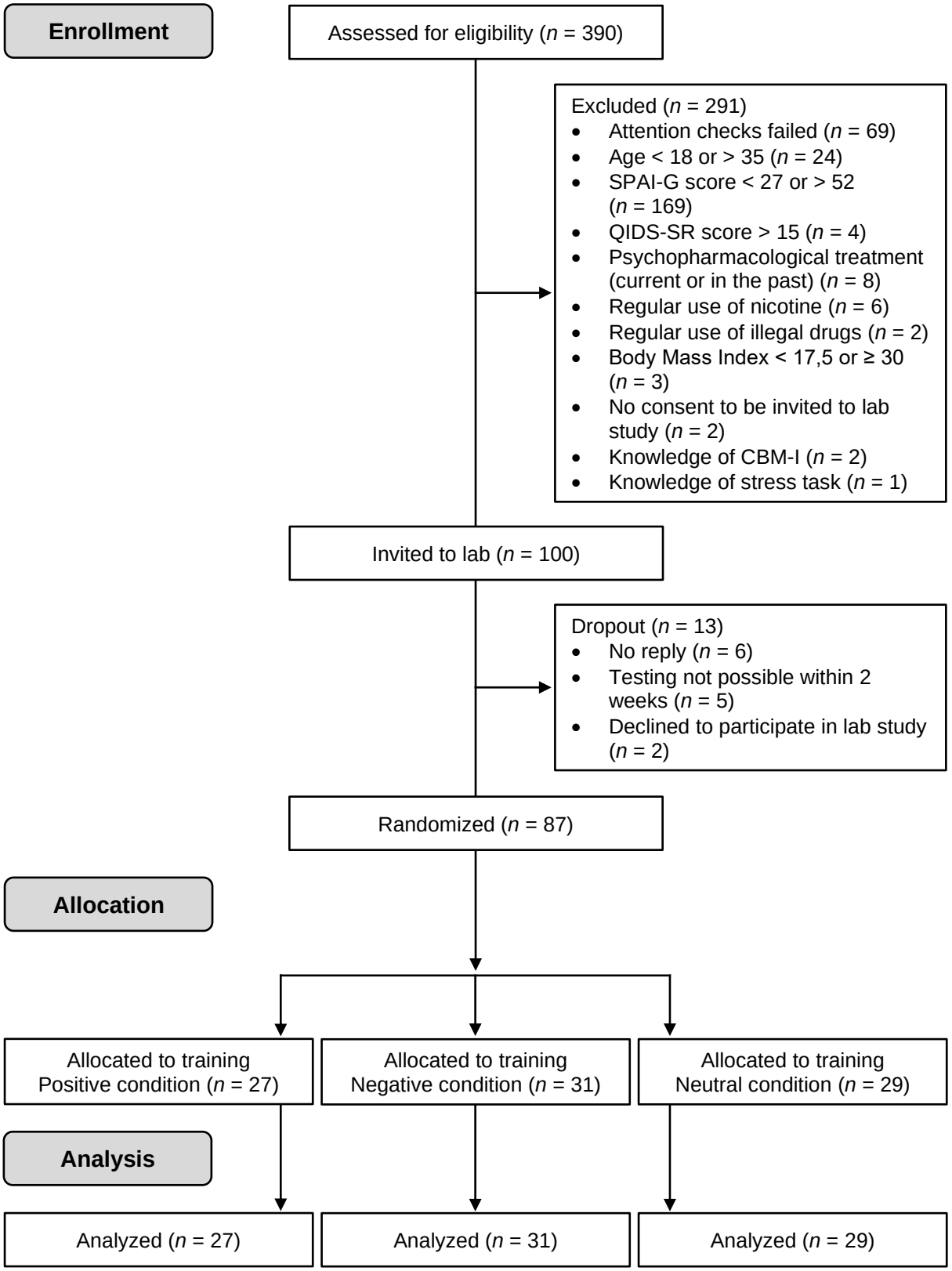


Figure 2

Encoding Recognition Task (ERT) Mean Scores at Pre- and Post-Training per Condition

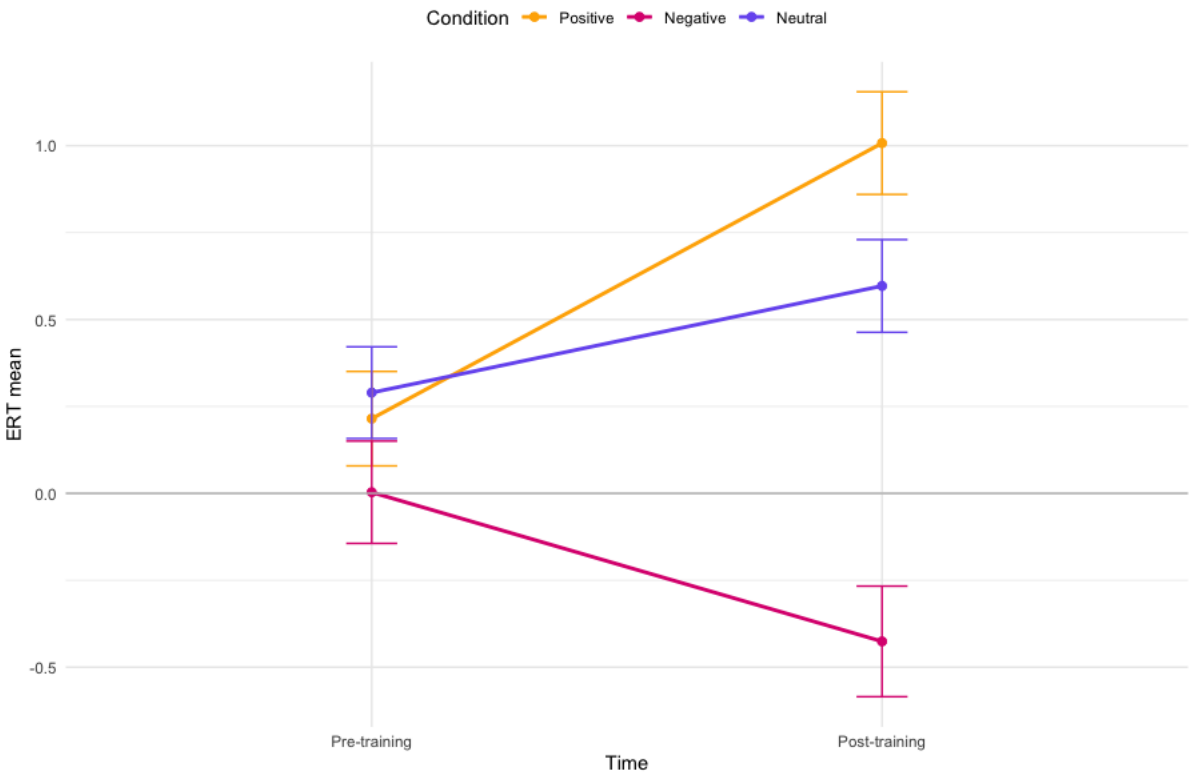
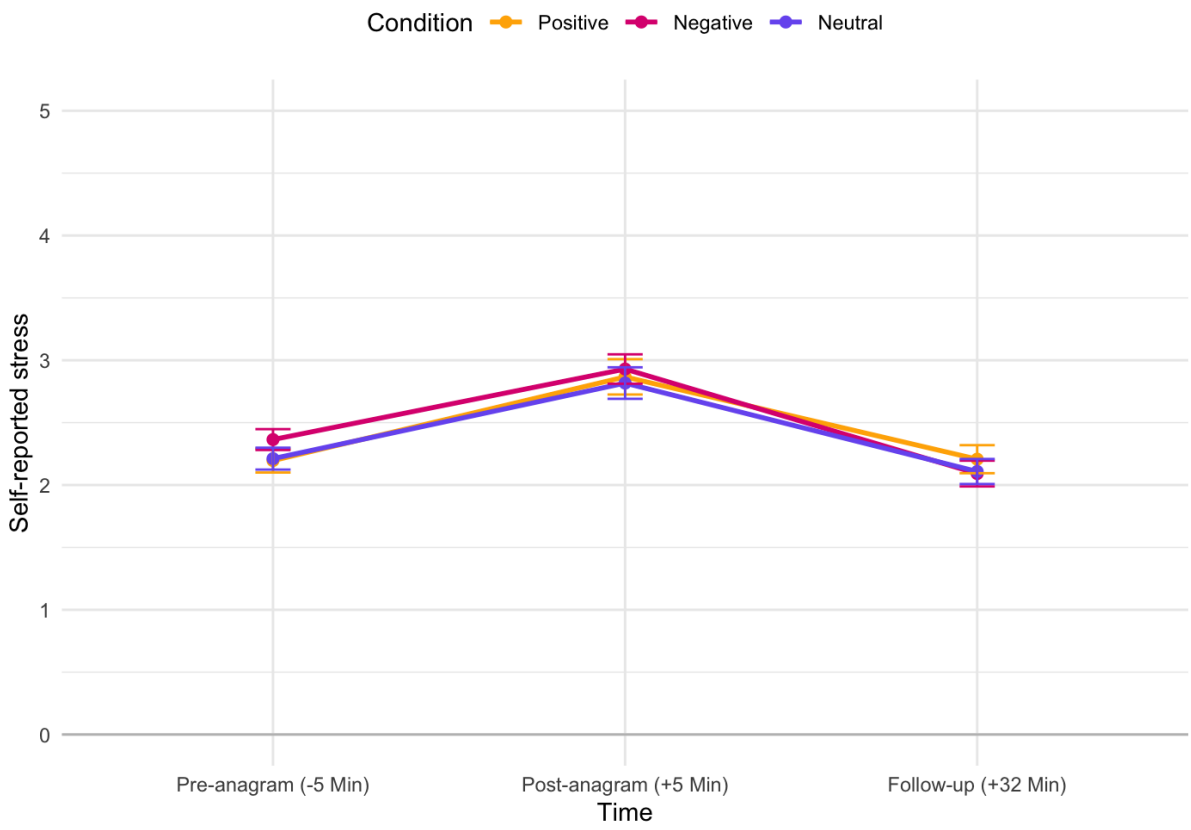


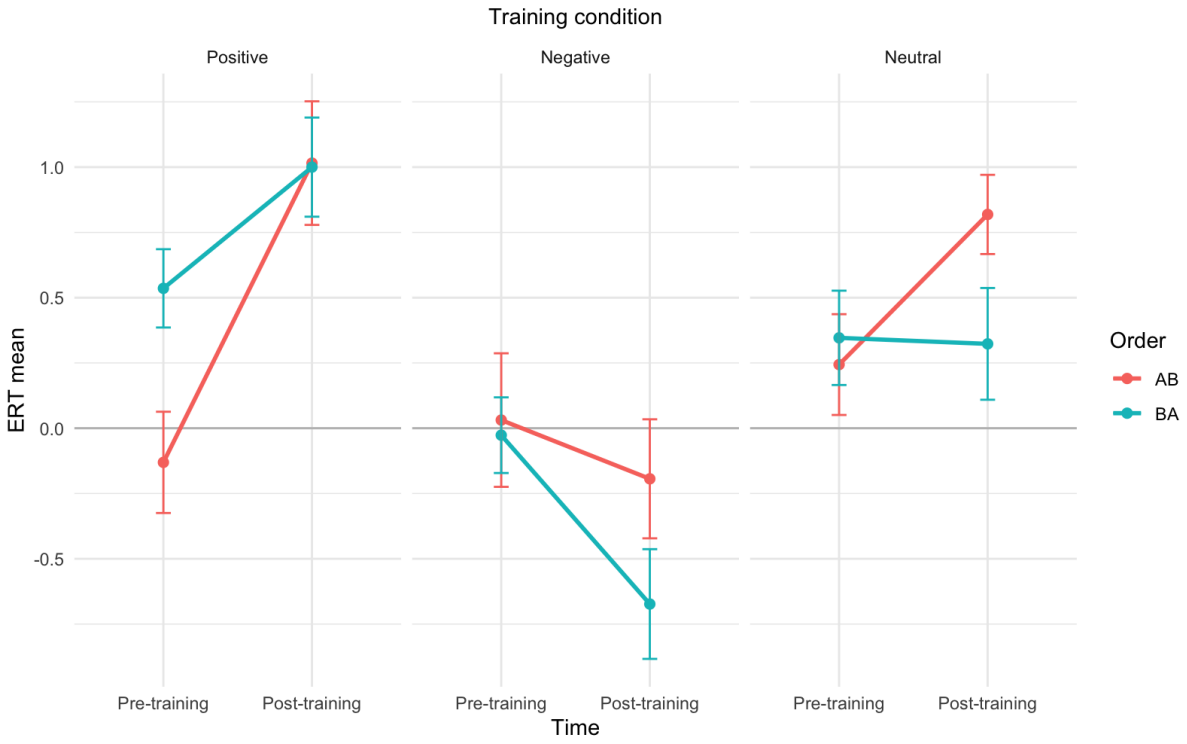
Figure 3

Self-reported Stress Mean Scores at Pre-anagram, Post-anagram and Recovery/Follow-up per Condition



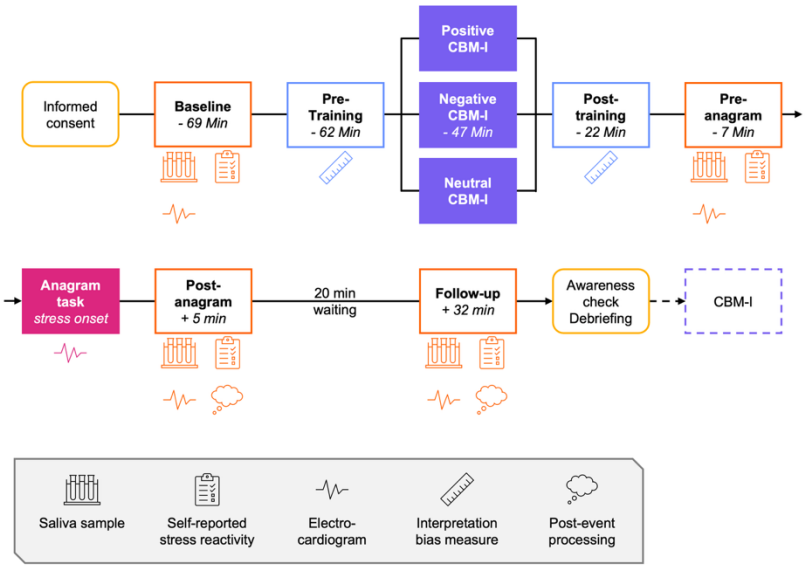
Supplementary Figure 1

Encoding Recognition Task (ERT) Mean Scores at Pre- and Post-Training per Training Condition and ERT Order



Supplementary Figure 2

Study Flow Chart



Note. CBM-I: Cognitive bias modification for interpretation. Only participants in the negative condition completed the second, then positive, CBM-I training (dotted box).

Supplementary Table 1

Descriptive Statistics and Internal Consistency of the Encoding Recognition Task

Version	Score	Pre-training		Post-training	
		M (*SD*)	α [95% CI]	*M* (*SD*)	α [95% CI]
A	Difference	0.06 (0.85)	0.75 [0.63, 0.85]	0.19 (1.03)	0.88 [0.81, 0.93]
	Positive	2.16 (0.53)	0.77 [0.65, 0.86]	2.20 (0.66)	0.86 [0.79, 0.92]
	Negative	2.10 (0.58)	0.80 [0.70, 0.88]	2.01 (0.63)	0.86 [0.70, 0.92]
B	Difference	0.28 (0.62)	0.67 [0.49, 0.80]	0.52 (0.95)	0.83 [0.74, 0.89]
	Positive	2.36 (0.53)	0.79 [0.68, 0.87]	2.55 (0.54)	0.80 [0.70, 0.88]
	Negative	2.08 (0.55)	0.80 [0.69, 0.88]	2.04 (0.54)	0.79 [0.68, 0.87]

Note. *M* = mean; *SD* = standard deviation; CI = confidence interval, calculated using the Feldt method; α = Cronbach's alpha; *n* (A pre, B post) = 45; *n* (B pre, A post) = 42. The versions did not differ significantly at Pre-training ($p = 0.181$, $d = –0.29$ [-0.72, 0.14]) or at Post-training ($p = 0.132$, $d = –0.33$ [–0.76, 0.10]).

Supplementary Table 2

Mixed ANOVA Analyzing Effects of Cognitive Bias Modification for Interpretation on Interpretation Bias

Effect	df	F	p	η^2g
Group	2, 81	12.77	< .001	.191
Order	1, 81	0.11	.743	.001
Time	1, 81	6.87	.010	.021
Group x Order	2, 81	1.70	.189	.030
Group x Time	2, 81	19.39	< .001	.107
Version x Time	1, 81	11.95	.001	.036
Group x Order x Time	2, 81	0.22	.801	.001

Note. Five participants showed identical ERT scores at pre- and post-training assessment and were thus excluded from the repeated-measures analysis due to the absence of within-subject variance.

Group: positive vs. negative vs. neutral; Order: AB vs. BA; Time: pre- vs. post-training