



Using Cognitive Bias Modification-Appraisal Training to Manipulate Appraisals about the Self and the World in Analog Trauma

Felix Würtz¹ · Julie Krans^{2,4,5,6} · Simon E. Blackwell¹ · Jan C. Cwik³ · Jürgen Margraf¹ · Marcella L. Woud¹

Accepted: 26 July 2021
© The Author(s) 2021

Abstract

Background Dysfunctional appraisals are a key mechanism in posttraumatic stress disorder (PTSD). Experimental manipulations of appraisals via Cognitive Bias Modification-Appraisal (CBM-App) training targeting cognitions related to the ‘self’ has shown to affect analog trauma symptoms. The present study aimed to conceptually replicate and extend previous findings by comparing a newly developed ‘world’ to the original ‘self’ training, and investigate the moderating role of locus of control (LOC) on intrusions.

Methods Healthy participants ($N = 173$) were exposed to distressing films as an analog trauma induction. Next, participants received positive or negative CBM-App, targeting either self- or world-relevant appraisals. Dysfunctional appraisals and LOC were assessed pre- and post-training. During the week after the laboratory session, participants recorded their intrusions and associated distress in a diary. One week later, trauma-relevant symptomatology was assessed.

Results Positive compared to negative CBM-App induced training-congruent appraisals, independent of the trained cognition (self vs. world). However, there was no effect on analog trauma symptoms and LOC did not moderate the training’s effect.

Conclusions Our results underline the validity of CBM-App as a method to experimentally manipulate appraisals. However, since we did not fully replicate previous findings further research on mechanisms associated with transfer effects is warranted.

Keywords PTSD · Cognitive Bias Modification · Appraisal · Locus of Control · Intrusions · Trauma

Introduction

Cognitive models of Posttraumatic Stress Disorder (PTSD) postulate that dysfunctional appraisals about the trauma and its sequelae (e.g. the thought that the event changed one’s life permanently to the worse, or that the world is a dangerous place) play a fundamental role in the development and maintenance of PTSD (e.g. Brewin et al., 1996; Dalgleish, 2004; Ehlers & Clark, 2000; Foa et al., 1989; cf. McNally & Woud, 2019). For example, according to the cognitive model of PTSD by Ehlers and Clark, (2000), dysfunctional appraisals may, for example, increase a feeling of current threat, which in turn triggers various symptoms, such as intrusive memories of the trauma.

In support of these models, dysfunctional trauma-related appraisals have been shown to be associated with symptoms of PTSD (Foa et al., 1999; for a meta-analysis, see Gómez de La Cuesta et al., 2019), and to directly predict PTSD symptoms longitudinally above initial symptom severity (Ehring et al., 2008; Kleim et al., 2007). Further, dysfunctional appraisals are positively associated with the use of safety

✉ Felix Würtz
felix.wuertz@rub.de

¹ Mental Health Research and Treatment Center, Faculty of Psychology, Ruhr-Universität Bochum, Massenbergrasse 9-13, 44787 Bochum, Germany

² Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands

³ Institute of Clinical Psychology and Psychotherapy, Faculty of Human Sciences, University of Cologne, Pohligstraße 1, 50969 Köln, Germany

⁴ Pro Persona Overwaal Centre for Anxiety, OCD, and PTSD, Nijmegen, the Netherlands

⁵ Department of Clinical Psychology, KU Leuven, Leuven, Belgium

⁶ Department of Clinical Psychology, Radboud University Nijmegen, Thomas van Aquinostraat 4, 6525 GD Nijmegen, the Netherlands

behaviors, which, in turn, predicts symptom severity and can maintain pathology (Beierl et al., 2019). Results from clinical studies have provided evidence that the reduction of dysfunctional appraisals precedes symptom reduction during trauma-focused cognitive behavioral therapy and cognitive processing therapy, but not vice versa (Kleim et al., 2013; Schumm et al., 2015; for a review, see Brown et al., 2019).

These findings offer robust evidence for a key role of dysfunctional appraisals in PTSD; however, these findings do not provide evidence for the potential causal role of dysfunctional appraisals. In order to investigate a potential causal risk factor, a systematic manipulation of the factor of interest is required (Kraemer et al., 1997). Techniques developed within the framework of Cognitive Bias Modification (CBM) (Koster et al., 2009; Woud & Becker, 2014) offer such an approach, via Cognitive Bias Modification – Appraisal (CBM-App) training. Previous studies have shown that such training can indeed induce training congruent appraisal styles towards (analog) trauma-relevant symptoms (Woud et al., 2012, 2013, 2018c). Specifically, participants who received a CBM-App training designed to induce a more functional appraisal style (positive training) adopted a more functional appraisal style towards a distressing film than participants who received a training designed to induce a more dysfunctional appraisal style (negative training) (Woud et al., 2012, 2013). These results showed that CBM-App successfully induced a training-congruent appraisal style. Further, those participants who received a positive training after the distressing film further reported fewer intrusions of the film in the following week (Woud et al., 2012) and participants who received a positive training prior to a distressing film reported less distress caused by the intrusions compared to those trained negatively (Woud et al., 2013). The latter effect of the CBM-App training was also found in a study in which participants re-imagined an autobiographical distressing event (Woud et al., 2018c) with positively trained participants reporting less intrusion distress compared to negatively trained participants in the week after the training. In this study, the CBM-App training was delivered after an initial reactivation of the event and the training was followed by a further reactivation. Finally, a recent randomized controlled trial compared a positive CBM-App training delivered in eight sessions over two weeks to a sham training condition as a treatment add-on for inpatients receiving trauma-focused treatment as usual (Woud et al., 2021). It was found that the positive training induced a more functional appraisal style and led to less severe PTSD symptoms directly after the training than the sham training condition.

To summarize, there is evidence from several different strands of research that dysfunctional appraisals are linked to symptom severity in PTSD, and that their reduction is associated with symptom reduction. Further, it has

been shown that a positive compared to a negative CBM-App training successfully induces a training congruent appraisal style (for a review, see Woud et al., 2017) with training-congruent effects on trauma-related symptoms in an analog and a clinical setting, which is line with core predictions of cognitive models of PTSD. These latter findings, however, represent an early stage of investigation, and systematic follow-up work is needed to both replicate and refine them.

Accordingly, our first aim was to conceptually replicate previous findings concerning the effects of the CBM-App training on both appraisal style and analog trauma symptoms (e.g. Woud et al., 2012, 2013, 2018b). We used a design similar to the study by Woud et al., (2012), therefore participants were exposed to distressing film clips and were allocated to either positive or negative CBM-App, followed by an assessment of analog trauma symptoms (e.g. via a one-week intrusion diary). Our second aim was to extend previous findings by investigating whether a CBM-App training targeting appraisals of the *world* (e.g. *The world is a dangerous place*) would have equal effects as the original training targeting appraisals of the *self* (e.g. *The trauma turned me into a bad person*). The stimuli of the CBM-App trainings used in previous studies (Woud et al., 2012, 2013, 2018b, c) were all based on the self-subscale of the Posttraumatic Cognitions Inventory (PTCI) (Foa et al., 1999), due to the strong association between self-appraisals and symptoms of PTSD (Bryant & Guthrie, 2005, 2007). However, this leaves it unclear (i) whether targeting trauma-related appraisals with other contents (e.g. derived from the other PTCI subscales ‘world’ or ‘self-blame’) would also induce a valence-congruent appraisal style, (ii) whether the respective training would then have content specific effects on the PTCI subscales, and (iii) whether this would also influence analog trauma symptoms correspondingly. To illustrate, it has been shown that appraisals regarding the world are also associated with symptoms of PTSD (Lancaster et al., 2011; van Heemstra et al., 2020; for a meta-analysis, see Gómez de La Cuesta et al., 2019), which thus offers a promising route for the development of CBM-App training versions targeting a broader spectrum of appraisals (the self and the world). In future clinical applications, these different training versions could potentially be matched to patients depending on the nature of their dysfunctional appraisals or be used in combination with each other. This is of particular interest as it has been suggested that different types of dysfunctional appraisals are associated with different dysfunctional behavioral strategies (cf. Ehlers & Clark, 2000) and thus the manipulation of specific appraisals might affect (analog) PTSD-related symptoms differently. Accordingly, in the present study, we aimed to test the effects of a CBM-App training targeting world-related appraisals. World-related appraisals were chosen instead of self-blame as there is equivocal

evidence for the association between self-blame and PTSD (Startup et al., 2007, for a recent review, see Gómez de La Cuesta et al., 2019).

Finally, our third aim was to investigate a potential factor that could be associated with the effects of CBM-App training on analog trauma-related symptoms, namely locus of control (LOC). LOC describes the extent to which individuals believe that they can control their circumstances by their actions (internal LOC), or whether their circumstances are governed by instances outside of their control (external LOC) (Rotter, 1966). In other words, the LOC describes how individuals appraise the controllability of their lives. In the context of PTSD, it has been shown that a more internal LOC is associated with resilience in the context of PTSD following a traumatic event (Chung et al., 2006; Karstoft et al., 2015), while a more external LOC is associated with more severe posttraumatic stress symptoms (Zhang et al., 2014). Given its relationship with PTSD symptoms and particularly given its definition, LOC thus overlaps conceptually with the concept of (dysfunctional) appraisals. Specifically, LOC can be regarded as a meta-concept inherent to dysfunctional appraisals, in that many dysfunctional appraisals concern the controllability of the traumatic situation itself or the associated symptoms, for example when appraising one's helplessness during the traumatic event in a negative manner (i.e. lack of control) or when appraising intrusions as a sign that one is losing control (cf. Ehlers & Clark, 2000). Based on this conceptual link, we thus investigated whether levels of LOC moderate the training's effect on intrusive memories with the idea that a more internal LOC might be associated with a lower susceptibility to effects of CBM-App training while a more external LOC be associated with greater susceptibility.

For the aims of the current study, we recruited a sample of healthy participants and exposed them to a set of distressing film clips in the laboratory (cf. James et al., 2016). Afterwards, participants received either *positive* or *negative* CBM-App training, targeting appraisals of either the *self* or the *world*. Dysfunctional appraisals were assessed pre- and post-training via the Encoding Recognition Task (ERT) (Mathews & Mackintosh, 2000). Levels of LOC as well as posttraumatic cognitions (PTCI) were also assessed pre- and post-training, and at one-week follow-up. Further, an intrusion diary was used to record intrusion frequency and distress in the week following the training. PTSD-related symptoms were measured during the follow-up assessment with the Impact of Event Scale—Revised (IES-R) (Maercker & Schützwohl, 1998; Weiss & Marmar, 1997).

Our main hypotheses were as follows: First, we expected a less dysfunctional appraisal style in the positive compared to the negative CBM-App conditions at post-training on the ERT, irrespective of the trained cognition (self / world). Second, we expected that this effect would generalize to

the PTCI at post-training and follow-up i.e. lower scores in the positive compared to the negative training conditions. However, we expected cognition-congruent effects on the PTCI subscales, thus we expected this effect to be greater on the subscale targeted in the training, i.e. a greater difference between the positive and negative condition on the self subscale for the self condition, and vice versa for the world condition. Further, we expected lower IES-R scores at follow-up as well as fewer intrusive memories and less intrusion distress on the intrusion diary for the positive compared to the negative CBM-App condition, independent of the trained cognition. Finally, we expected that LOC would moderate the training's effect on intrusions (i.e. frequency and distress) in such that the between group difference for the positive versus negative CBM-App training would be greater amongst participants with lower levels of internal LOC/higher levels of external LOC.¹ Additionally, via exploratory analyses, we investigated whether there was an association between the LOC and dysfunctional appraisals as measured via the ERT and the PTCI at baseline.

Methods

Recruitment and Screening

Participants were recruited via social media and the noticeboards at Ruhr-Universität Bochum (RUB). Interested participants contacted the study team via e-mail and were then asked to complete an initial online screening. Due to the use of the trauma film paradigm, the screening served as an eligibility check to exclude participants who could be disproportionately negatively affected by the study (e.g. due to mental health problems) or otherwise did not meet the general inclusion and exclusion criteria. Regarding the latter, the online screening assessed symptoms of depression measured via the Beck Depression Inventory-II (BDI-II) (Beck et al., 1996; German version: Hautzinger et al., 2006), and participants were also asked whether they had previously suffered from a psychological disorder or were currently suffering from one, and if so to indicate the disorder. The screening also included questions about demographic information and trait anxiety was assessed via the State and Trait Anxiety Inventory-Trait version (STAI-T) (Spielberger et al., 1983; German version: Laux et al., 1981). Participants were included if they were fluent in written and spoken German and aged between 18 and 35 years. The latter criterion was

¹ In this study LOC was analyzed on three dimensions: Internal, Powerful others-, and Chance-related, with the latter two dimensions representing an LOC laying outside of the own person, thus we took them into account as dimensions of a more external LOC (Levenson, 1973).

applied to recruit a relatively homogenous sample in order to reduce between-subject variability and by this increase statistical power. Exclusion criteria were the following: BDI-II total score of > 18 , potential suicidality (indicated by a score of > 0 on item 9 on the BDI-II), a self-reported current or past diagnosis of PTSD, color-blindness or red-green deficiency, or participation in previous studies using the same CBM-App procedure. Participants received either course credit or 15€ as an incentive for participation. The study was approved by the local ethics committee of the Psychology department at RUB (approval number: 338). All participants gave written informed consent prior to participation. We aimed to recruit enough participants to have about 40 per condition after any exclusions for example for missing diary data. This was not based on a specified power calculation to find a particular outcome, but rather based on having at least as large a sample per condition as previous studies in this line of research (e.g. Woud et al., 2012).

Material

Trauma Films

As in previous CBM-App studies (e.g. Woud, et al., 2018b), the trauma film paradigm was used as an analog trauma in the laboratory (James et al., 2016). For the present study, we used the films as used by Woud et al., (2012). The trauma films contained 11 distressing film clips (e.g. footage of the 9/11 terrorist attacks) with a duration of 1–3 min each (20 min in total), and participants were instructed to watch the films imagining themselves as bystanders actually present at the scene.

CBM-App Training

The CBM-App training was a computerized sentence-completion paradigm used in previous studies (Woud et al., 2012, 2013). Participants were presented with open-ended, ambiguous sentences targeting trauma-relevant appraisals. After each sentence, a word fragment appeared, resolving the sentence's ambiguity in a training-congruent manner. Participants were instructed to complete the word fragment by typing in its first missing letter (e.g. *My reactions to the event show that my way of processing is he_lthy / unsucc_ssful*; positive condition: *healthy*, negative condition: *unsuccessful*). There were 72 ambiguous, appraisal-related sentences and 8 neutral filler sentences, presented in 8 blocks of 10. The sentences within each block were presented in random order. Thirty-two sentences were followed by a short comprehension question regarding the last presented sentence (answers: *yes* or *no*) to ensure ongoing concentration and compliance to the task.

The training scenarios for the self training conditions were derived from the self subscale of the PTCI (as used by Woud et al., 2012, 2013, 2018b, c). For the purpose of the present study, a second training version was developed, targeting appraisals of the world, using the PTCI world subscale as a foundation to create the scenarios (e.g. *The world is full of mir_cles/haz_rds*; positive condition: *miracles*, negative condition: *hazards*). As such, there were four conditions, based on the between group factors Valence (Positive vs. Negative) and Cognition (Self vs. World): positive self-training, negative self-training, positive world-training, negative world-training. Participants were allocated to these conditions via a counterbalancing sheet, therefore, every fourth participant was allocated to the positive-self condition etc.

Manipulation Check Measures

Mood and Attentiveness Rating

To assess the effect of the trauma films and the CBM-App training on mood, state negative mood was measured using a 11-item questionnaire consisting of ten one-word items assessing negative emotions (e.g. *anxious*) and one item assessing a positive emotion (*happy*, reversed coding for sum score). For every item, participants were instructed to rate how they were feeling right now on a 11-point scale ranging from 0 (*Not at all*) to 10 (*Very much*). Reliability ranged from α [95%-CI] = 0.88 [0.86, 0.90] at baseline to α [95%-CI] = 0.91 [0.89, 0.93] at post-training. Further, after the trauma films participants were asked to rate their attentiveness to the films on a 11-point scale ranging from 0 (*not attentive at all*) to 10 (*complete attention*).

Encoding Recognition Task (ERT)

To test whether the CBM-App training successfully induced the corresponding appraisal style, an adapted version of the ERT (Mathews & Mackintosh, 2000) was administered. As in earlier CBM-App studies in the context of analog trauma, the ERT was adapted as such that stimuli were designed to assess dysfunctional appraisals related to traumatic events or its sequelae (e.g. Woud et al., 2012, 2018c). The ERT consists of two phases: First there is an encoding phase during which participants are presented with 10 novel ambiguous scenarios (e.g. *I can easily imagine how people would evaluate me if they would know the whole ...*). Every scenario is presented with an introductory title (e.g. *The whole story*) and participants are asked to keep this title in mind. As in the CBM procedure, the scenario's final word is presented as a word fragment that participants are asked to complete by typing in the first missing letter (e.g. fragment: *st_ry*,

resolved: *story*). In contrast to the CBM procedure, however, completing the word fragment does not resolve the scenarios' ambiguity.

During the recognition phase, the titles of the scenarios are presented again, and for each title four novel sentences are presented in random order. Participants are instructed to rate how similar the meaning of each sentence is to the scenario corresponding to the title using a 4-point scale ranging from 1 (*Not similar at all*) to 4 (*Very similar*). Two out of these four sentences are target sentences that resolve the ambiguity of the scenario in either a positive or a negative way (e.g. Positive: *If people knew the whole story, they would think of me as brave*, Negative: [...] *they would think of me as cowardly*). The other two sentences are foil sentences, one with a positive and one with a negative meaning (e.g. Positive: [...] *they would think of me as a precise storyteller*, Negative: [...] *they would think of me as an unprecise storyteller*). However, neither of these foil sentences resolve the scenario's ambiguity. Scoring in the current study followed that of previous studies in this line of research (Woud et al., 2012, 2018c). That is, via the similarity ratings, a relative bias score was computed by subtracting the mean similarity ratings of the negative target sentences from the mean similarity ratings of the positive target sentences. In this way positive scores (i.e. > 0) indicate a more positive than negative appraisal style, and vice versa for a negative bias.

There were two different versions of the ERT, matched with participants' training condition. That is, participants in the self condition did the ERT with scenarios related to the self while participants in the world condition did the ERT with scenarios related to the world. For all participants, the ERT was applied pre- and post-training. There were two sets of scenarios for each version (A and B), and their order was counterbalanced via a predefined counterbalancing scheme. Reliability ranged from α [95%-CI] = 0.83 [0.76, 0.90] (Version B, world ERT, pre-training) to α [95%-CI] = 0.96 [0.95, 0.98] (Version B, world ERT, post-training).

Outcome and Moderator Measures

Posttraumatic Cognitions Inventory (PTCI)

The PTCI (Foa et al., 1999; German version: Ehlers, 1999) consists of 33 statements (e.g. *The world is dangerous*) concerning appraisals of a traumatic event. It contains three subscales: Cognitions regarding the self (21 items), the world (7 items) and self-blame (5 items). For every statement, participants are instructed to rate their level of agreement on a 7-point scale ranging from 1 (*completely disagree*) to 7 (*completely agree*). For the purpose of this study, the instructions of the PTCI were adapted in such that participants were asked to answer the questionnaire in relation to the trauma films. For the baseline assessment, participants were

instructed to answer the questionnaire in relation to an own negative life event (Bryant & Guthrie, 2005, 2007; Woud et al., 2013). Reliability for the full scale ranged from α [95%-CI] = 0.95 [0.94, 0.96] at baseline to α [95%-CI] = 0.96 [0.95, 0.97] at follow-up, and for the subscales from α [95%-CI] = 0.73 [0.67, 0.80] (subscale self-blame at post-training) to α [95%-CI] = 0.95 [0.94, 0.96] (subscale self on every assessment).

Impact of Event Scale Revised (IES-R)

The IES-R (Weiss & Marmar, 1997; German version: Maercker & Schützwohl, 1998) is a 22-item self-report questionnaire assessing reactions to a specific traumatic event. The questionnaire consists of three subscales: Intrusions (7 items), avoidance (8 items) and hyperarousal (7 items). The instructions of the questionnaire were adapted to fit our study design, thus participants had to rate how they reacted to the trauma-films in the following week on a 4-point scale ranging from 1 (*Not at all*) to 4 (*Often*). Reliability for the full scale was α [95%-CI] = 0.85 [0.82, 0.88], and ranged from α [95%-CI] = 0.59 [0.50, 0.68] (hyperarousal subscale) to α [95%-CI] = 0.82 [0.78, 0.86] (intrusions subscale) for the subscales.

Intrusion Diary

Participants were instructed to record their intrusions in the week following the first laboratory session in a structured diary (Woud et al., 2018c) as well as the distress caused by each intrusion on a scale ranging from 0 (*not distressing at all*) to 100 (*extremely distressing*). Intrusions were described to the participants as sudden, involuntary memories about the trauma films that can occur in forms of mental imagery, verbal thoughts, or a combination of both (cf. Woud et al., 2012, 2013, 2018c). Participants were asked for a short description of each intrusion to ensure that they were linked to the trauma films and that they were not remembered voluntarily. Intrusions that could not be linked to the trauma films were excluded from the analysis. Intrusion distress was scored by dividing a participant's the sum of the distress ratings by the total number of intrusions reported by the participant, thus participants reporting no intrusions were not included in the intrusion distress analyses.

Levenson's Internal, Powerful Others and Chance scale (IPC)

The IPC (Levenson, 1973; German version: Krampen, 1979) consists of 24 statements about the perceived LOC (e.g. *Whether or not I get to be a leader depends mostly on my ability*). It contains three subscales, each consisting of eight items: Internal, Powerful others, and Chance. For every statement, participants are instructed to rate their level of

agreement on a 7-point Likert scale ranging from -3 (*strong disagreement*) to 3 (*strong agreement*). Reliability ranged from α [95%-CI] = 0.46 [0.34, 0.58] (subscale Powerful others at baseline) to α [95%-CI] = 0.70 [0.63, 0.76] (subscale Powerful others at follow-up).

Additional Questionnaire Measures

Further questionnaires were assessed to describe the sample at baseline. State-anxiety was assessed via the STAI-State version (STAI-S) (Spielberger et al., 1983; German version: Laux et al., 1981) and prior exposure to traumatic events was assessed via the Trauma History Checklist (THC) (Holmes & Steel, 2004). At the end of the study, a feedback questionnaire was administered to assess the participants' diary compliance and opinion about the purpose of the CBM-App training. For more information on these questionnaires, see the Supplementary material.

Procedure

After participants completed the online screening, eligible participants were invited to the laboratory. Upon arrival, participants gave written informed consent and were allocated to one of the four CBM conditions via a predefined counterbalancing sheet. After that, participants were given the baseline questionnaire battery consisting of the THC, the STAI-S, the first IPC, the first PTCI, and first mood rating. After that, the trauma film clips were presented, followed by a second mood rating and the attentiveness rating. Then, the pre-training ERT was applied, followed by the CBM-training. After that, mood was assessed again and participants completed the post-training ERT, the post-training PTCI, and IPC questionnaire, respectively. At the end of the first day in the laboratory, participants were given the intrusion diary and its use was explained.

One week later participants returned to the laboratory. At first, they handed in their intrusion diary. While the experimenter checked whether the intrusions reported in the diary were linked to the trauma films, participants completed the follow-up questionnaires, i.e. the IES-R, the follow-up PTCI, the follow-up IPC, the film-relevance and feedback questionnaire. Afterwards, each intrusion that could not be clearly linked to one of the trauma films was discussed with the participants. Finally, participants were debriefed and received their incentive. For a tabular overview of the procedure and further information regarding used materials, see Supplementary Material.

Statistical Approach

Baseline characteristics and the attentiveness to the trauma films were compared using 2×2 (Valence: Positive,

Negative; Cognition: Self, World) between group ANOVAs. A chi-squared test was conducted for the variable 'Gender'. The effect of the trauma films on mood was analyzed using a mixed linear model with a random intercept for individual participants (Between-subject factors: Valence: Positive, Negative; Cognition: Self, World; Within-subjects factor: Time: Pre-Films, Post-Films). We expected to only find a significant main effect of the within subject factor 'Time' which would be indicative of a mood worsening after trauma film exposure across all four conditions.

As a manipulation check for the effect of the CBM-App training on the ERT post-training, we conducted a $2 \times 2 \times 2$ (Valence: Positive, Negative; Cognition: Self, World; Order ERT Set: AB, BA) ANCOVA, controlling for the pre-training ERT bias score. Here, our expectation was to find a significant main effect of Valence, with means showing a more positive bias score in the positive conditions compared to the negative conditions. Effects of the CBM-App training on the PTCI self- and world subscale at post-training and follow-up were tested via mixed linear models with random intercepts for individual participants (Between subject factors: Valence: Positive, Negative; Cognition: Self, World; Within subject factor: Subscale: Self, World), with the PTCI total score at pre-training as the covariate separately for post-training and follow-up.² We expected to find a significant main effect of Valence with lower PTCI scores in the positive compared to the negative valence conditions. However, we expected to find this effect to be qualified by a significant Valence x Cognition x Subscale three-way interaction, with a greater effect of Valence on the training-congruent PTCI subscale (i.e. the self subscale for the self training and the world subscale for the world training).

Effects of the training on the IES-R and intrusion frequency and distress assessed via the diary were analyzed using 2×2 (Valence: Positive, Negative; Cognition: Self, World) ANOVAs. For each analysis, we expected to find a significant main effect of Valence but no other significant main effects or interactions, indicating lower analog symptoms in the positive than in the negative valence conditions irrespective of the trained cognition.

Further, we investigated the moderating role of the pre-training LOC in the effect of Valence on the number of intrusions and the intrusion distress reported in the diary. This was done using stepwise regression. We expected to find significant two-way interactions between Valence and the respective IPC subscale, with smaller effects of Valence for a higher internal LOC and greater effects of Valence

² Due to the different number of items in each subscale (Self: 21, World: 7), mean scores instead of sum scores were used as the dependent variable. The mean centered total sum score was used as the covariate.

Table 1 Baseline characteristics and sample description

Measure	Positive CBM-App <i>M</i> (<i>SD</i>)		Negative CBM-App <i>M</i> (<i>SD</i>)		Statistics*
	Self (<i>N</i> =44)	World (<i>N</i> =42)	Self (<i>N</i> =44)	World (<i>N</i> =43)	
Age	23.93 (3.63)	24.05 (3.94)	23.61 (3.21)	24.49 (3.74)	$F(1, 169) < .81, p > .371$
Sex: Female/male	27/17	26/16	37/7	32/11	$\chi^2(3) = 7.49, p = .058$
BDI-II	5.09 (4.64)	3.98 (4.07)	4.68 (4.90)	4.56 (3.95)	$F(1, 169) < .85, p > .359$
STAI-Trait _a	36.07 (9.37)	35.47 (8.70)	35.22 (10.05)	35.88 (8.17)	$F(1, 165) < .21, p > .655$
STAI-State	34.93 (8.77)	34.29 (10.36)	31.86 (5.65)	32.88 (6.39)	$F(1, 169) < 3.39, p > .067$
THC _b	.80 (1.15)	.93 (1.24)	1.02 (1.44)	.84 (.92)	$F(1, 169) < .76, p > .385$
Mood pre-trauma film	13.45 (13.35)	12.81 (12.78)	10.36 (7.42)	11.16 (9.30)	$F(1, 169) < 2.02, p > .157$
ERT bias score pre-training	.16 (1.03)	.45 (.77)	.08 (1.06)	.18 (.96)	$F(1, 169) < 1.80, p > .181$
PTCI pre-training	94.52 (35.08)	80.78 (33.07)	90.80 (30.87)	85.02 (36.76)	$F(1, 169) < 3.56, p > .060$
IPC—Int pre-training	35.23 (5.71)	37.26 (5.33)	35.05 (5.47)	35.26 (4.63)	$F(1, 169) < 1.94, p > .166$
IPC – Pow pre-training	19.23 (4.84)	17.29 (5.91)	16.77 (6.07)	18.86 (4.97)	Valence x Cognition: $F(1, 169) = 5.86, p = .017, \eta^2_G = .034$
IPC—Cha pre-training	19.55 (5.70)	17.95 (5.93)	19.55 (5.34)	19.05 (6.52)	$F(1, 169) < 1.37, p > .243$

*Results of 2×2 (Valence: Positive, Negative; Cognition: Self, World) ANOVAs. Only the largest *F* statistic / smallest *p* value is shown. a=During the online screening, 4 participants (2 allocated to the positive world, 2 to the negative self condition) did not answer the STAI-T. b=Due to an administration error the first 107 participants answered only the first 9 questions of the THC instead of the complete set of 13 questions. An independent groups Welch-test revealed no difference in the THC sum score between the participants who answered 9 or 13 questions, respectively, $t(171) = .65, p = .518, d = .101$. *BDI-II* Beck Depression Inventory-II, *ERT* Encoding Recognition Task, *IPC* Levenson's Internal, Powerful Others, and Chance scale, *Int* Internal, *Pow* Powerful Others, *Cha* Chance, *PTCI* Posttraumatic Cognitions Inventory, *STAI-Trait/State* State and Trait Anxiety Inventory Trait/State version, *THC* Trauma History Checklist

for a more Powerful others or Chance related LOC. Finally, in exploratory analyses, we investigated the associations between the ERT, the PTCI, and each subscale of the IPC at baseline via computing Pearson correlations with two-sided significance tests, except for the association between the ERT and the PTCI for which we tested one-sided due to our expectation for a negative correlation.

To test for the equivalence of the self- and the world training, we supplemented the frequentist analyses of our main hypotheses with the computation of Bayes Factors (BFs) which allow quantifying the evidence in favor of either the H0 (effect of a variable is not greater than zero) or the H1 (effect of a variable is greater than zero). In this study, a $BF_{10} < 1$ is indicative of the presence of an effect while $BF_{10} > 1$ is indicative of its absence. More extreme numbers indicate stronger evidence with $0.33 \leq BF_{10} \leq 3$ being considered as anecdotal evidence for the presence or absence of an effect, respectively³ (cf. Meyer et al., 2020 for interpretation conventions).

All analyses were run in RStudio version 1.1.463 (RStudio Team, 2016). ANOVAs and ANCOVAs with type III sums of squares were calculated using the package 'afex' (Singman et al., 2020). Effect sizes are expressed as

generalized eta-squared (η^2_G). Linear mixed models were calculated using the package 'nlme' (Pinheiro et al., 2020) using maximum likelihood estimation. Effect sizes for these models were calculated by dividing the coefficient of a specific effect in the model by the square root of the sum of variances of the random effects (Brysbaert & Stevens, 2018). Post-hoc pairwise comparisons with Bonferroni correction for following up significant interactions were calculated using the package 'emmeans' (Lenth, 2020). BFs were calculated using the package 'BayesFactor' (Morey & Rouder, 2018) with default (Cauchy) priors and 'participant' as a random factor for within-subject analyses. Hierarchical linear regression for moderation analyses was computed using the package 'stats' (R Core Team, 2020). For more information on the computation of BFs, contrasts and variable preparations, see Supplementary Material.

Results

Participant Characteristics

A total of $N = 174$ participants were tested. Due to an administrative error, one participant (allocated to the Positive World condition) with a BDI-II total score of 20 was included in the testing sessions, and thus excluded from further analyses. The final sample therefore consisted of $N = 173$ participants (sex: 70.52% female, age: $M = 24.02$,

³ For example, a $BF_{10} = 2$ means that the given data would be twice as likely under the H1 than under the H0, and vice versa for a $BF_{10} = .05$.

Table 2 Model based descriptive statistics of Manipulation check and outcome data

Measure	Positive CBM-App <i>EMM</i> (<i>SE</i>)		Negative CBM-App <i>EMM</i> (<i>SE</i>)	
	Self (<i>N</i> =44)	World (<i>N</i> =42)	Self (<i>N</i> =44)	World (<i>N</i> =43)
Mood post-trauma film	37.66 (2.50)	31.10 (2.56)	41.93 (2.50)	34.60 (2.53)
Attention to movie	9.41 (.12)	9.36 (.12)	9.61 (.12)	9.37 (.12)
ERT bias score post-training	1.33 (.14)	1.25 (.14)	- 0.26 (.14)	- 0.57 (.14)
PTCI Self and World subscale combined mean, post-training	2.95 (.12)	2.93 (.12)	3.33 (.12)	3.34 (.12)
PTCI Self and World subscale combined mean, follow-up	2.49 (.13)	2.58 (.13)	2.78 (.12)	2.78 (.13)
IES-R total score	34.69 (1.18)	31.12 (1.21)	31.37 (1.18)	34.03 (1.21)
Intrusion frequency	4.07 (.62)	3.24 (.64)	3.20 (.62)	3.77 (.63)
Intrusion distress _a	38.11 (3.75)	34.49 (3.91)	33.40 (3.91)	38.11 (3.80)
Number of excluded intrusions	6	11	13	3

Estimated marginal means and standard errors of the respective models are presented. a=Only participants who reported at least one intrusion (*N*=141) were included in the intrusion distress analyses. One participant allocated to the negative world condition did not return for the second day and is therefore excluded from analyses concerning the second day. For analyses in which the interaction with the covariate was included, values represent the estimated marginal mean and standard error at the mean level of the covariate. *ERT* Encoding Recognition Task, *PTCI* Posttraumatic Cognitions Inventory, *IES-R* Impact of Event Scale - Revised

SD = 3.62, range = 18–34 years). There were no baseline differences between the four groups except for the Powerful others subscale of the IPC for which post-hoc pairwise comparisons revealed no significant group differences ($t_s(169) < 2.20$, $p_s > 0.221$) (see Table 1). One participant (allocated to the Negative World condition) did not return for the second day. As such, this participant was only included in analyses regarding the first day. Three participants (two allocated to the Positive World condition, one to the Positive Self condition) completed the follow-up questionnaires at home and returned them via e-mail together with the diary, as they could not attend the second session in the laboratory due to illness or holidays. These participants were included in all analyses.

Manipulation Checks

Mood Change Pre-post Trauma Film Clips

We found the expected significant main effect of Time, $t(169) = 16.79$, $p < 0.001$, $d = 1.49$, $BF_{10} = 1.36 \times 10^{37}$, indicating a general significant increase in negative mood from pre- to post trauma films across all four conditions (Pre: $M = 11.94$, $SD = 10.95$; Post: $M = 36.39$, $SD = 20.92$). Unexpectedly, there was a significant Time x Valence interaction, $t(169) = 2.16$, $p = 0.032$, $d = 0.19$, $BF_{10} = 1.23$, indicating that mood changed differently over time for the positive versus negative conditions. Post-hoc pairwise comparisons showed that negative mood significantly increased in both the positive and negative conditions (Positive: $t(85) = 10.32$, $p < 0.001$, $d = 1.14$; Negative: $t(86) = 13.44$,

$p < 0.001$, $d = 1.48$), however, there was no significant difference between the positive and the negative conditions in mood after the trauma films, $t(171) = 1.54$, $p = 0.125$, $d = 0.19$. Further, there was a significant Time x Cognition interaction, $t(169) = 2.42$, $p = 0.017$, $d = 0.21$, $BF_{10} = 2.19$. Post-hoc pairwise comparisons showed that negative mood increased in both the self and the world conditions (Self: $t(87) = 13.70$, $p < 0.001$, $d = 1.50$; World: $t(84) = 10.08$, $p < 0.001$, $d = 1.12$). Comparing negative mood after the trauma films, results showed that participants in the self condition had more negative mood post-film than participants in the world condition, $t(171) = 2.75$, $p = 0.007$, $d = 0.33$.

In summary, the negative mood of participants increased irrespective of the condition, yet negative mood post-trauma films was higher in the self compared to the world conditions. For descriptive statistics of manipulation check and outcome data, see Table 2. For further descriptive statistics such as interactions with the covariates, see Table S1 of the Supplementary Material.

Attentiveness to the Trauma Films

The groups all showed high levels of attentiveness to the trauma films and did not differ in this aspect as indicated by non-significant main effects and a non-significant interaction ($F_s(1169) < 1.45$, $p_s > 0.230$, $BF_{10}s < 0.33$).

ERT

As expected, we found a significant main effect of Valence, $F(1, 163) = 154.14$, $p < 0.001$, $\eta^2_G = 0.424$,

$BF_{10} = 1.16 \times 10^{21}$, indicating a more positive bias score after the CBM-App training in the positive compared to the negative conditions (Positive: $EMM = 1.29$, $SE = 0.10$; Negative: $EMM = -0.41$, $SE = 0.10$). This effect was independent of the trained cognition as indicated by a non-significant main effect of Cognition, $F(1, 163) = 1.84$, $p = 0.177$, $\eta^2_G = 0.009$, $BF_{10} = 0.34$, as well as a non-significant Valence x Cognition interaction, $F(1, 163) = 0.69$, $p = 0.408$, $\eta^2_G = 0.003$, $BF_{10} = 0.24$. For analyses of foil sentences, see Supplementary Material.

Outcome variables

PTCI

At post-training, we found the expected significant main effect of Valence, $t(168) = 3.35$, $p = 0.001$, $d = 0.21$, $BF_{10} = 2.76$, indicating lower PTCI scores at post-training in the positive compared to the negative conditions (Positive: $EMM = 2.94$, $SE = 0.08$; Negative: $EMM = 3.33$, $SE = 0.08$). This effect was independent of Cognition as indicated by a non-significant main effect of Cognition, $t(168) = 0.06$, $p = 0.955$, $d < 0.01$, $BF_{10} = 0.13$, and a non-significant Valence x Cognition two-way interaction, $t(168) = 0.10$, $p = 0.920$, $d = 0.01$, $BF_{10} = 0.19$. Unexpectedly, however, the Valence x Cognition x Subscale interaction was non-significant $t(168) = 0.14$, $p = 0.887$, $d = 0.01$, $BF_{10} = 0.22$, indicating that the effect of the training Valence on the respective PTCI subscale did not differ between the self and the world training.

At follow-up, we found inconclusive evidence for a main effect of Valence $t(167) = 1.98$, $p = 0.049$, $d = 0.13$, $BF_{10} = 0.46$, and this was independent of the trained cognition as indicated by a non-significant main effect of Cognition, $t(167) = 0.36$, $p = 0.722$, $d = 0.02$, $BF_{10} = 0.14$, and a non-significant Valence x Cognition two-way interaction, $t(167) = 0.35$, $p = 0.730$, $d = 0.02$, $BF_{10} = 0.18$. Further, as in the post-training analyses, the expected Valence x Cognition x Subscale interaction was non-significant, $t(168) = 0.59$, $p = 0.595$, $d = 0.02$, $BF_{10} = 0.26$.

In summary, at post-training we found lower PTCI scores on both the self- and the world subscale in the positive compared to the negative condition. Unexpectedly, however, there was no differential effect of the self or the world training on the corresponding subscales. At follow-up, effects followed the same pattern as at post-training.

IES-R

For the IES-R, we did not find the expected main effect of Valence $F(1, 168) = 0.03$, $p = 0.864$, $\eta^2_G = 0.000$, $BF_{10} = 0.17$, indicating no overall difference in analog trauma symptoms between the positive and negative

conditions. Further, there was no significant main effect of Cognition $F(1, 168) = 0.15$, $p = 0.704$, $\eta^2_G = 0.001$, $BF_{10} = 0.16$, showing that the self- and the world conditions did not differ. Unexpectedly, there was a significant Valence x Cognition interaction $F(1, 168) = 6.85$, $p = 0.010$, $\eta^2_G = 0.039$, $BF_{10} = 4.55$. Post-hoc pairwise comparisons, however, revealed no differences between the four groups ($ts(168) < 2.13$, $ps > 0.212$, $BF_{10}s < 1.17$) with weak anecdotal evidence for a higher IES-R score in the positive self condition compared to the positive world condition, $BF_{10} = 1.06$, as well as compared to the negative self condition, $BF_{10} = 1.16$. In summary, we did not find the expected lower overall IES-R scores in the positive compared to the negative conditions.

Intrusion Diary

In total, 81.98% ($n = 141$) of the participants reported at least one intrusive memory of the trauma films in the diary.⁴ There was no difference between the four conditions in whether participants had intrusive memories or not, $\chi^2(3) = 0.87$, $p = 0.834$. Contrary to our expectation, the main effect of Valence on intrusion frequency was not significant $F(1, 169) = 0.07$, $p = 0.790$, $\eta^2_G < 0.001$, $BF_{10} = 0.17$ indicating no difference between the positive and negative conditions in the number of intrusions. Further, the main effect of Cognition as well as the Valence x Cognition two-way interaction were non-significant ($Fs < 1.25$, $ps > 0.267$, $\eta^2_G < 0.008$, $BF_{10} < 0.40$). We found a similar pattern for intrusion distress, in that neither the main effect of Valence or Cognition, nor their two-way interaction were significant ($Fs(1, 137) < 1.19$, $ps > 0.279$, $\eta^2_Gs < 0.100$, $BF_{10}s < 0.43$). In summary, we did not find the expected lower intrusion frequency and associated distress in the positive compared to the negative condition.

LOC Moderation

Contrary to our expectations we did not find any significant interactions between the LOC subscales and Valence, indicating that LOC was no moderator of the effect of Valence on the intrusion frequency. However, the powerful others LOC predicted the number of intrusions with a higher powerful others LOC being associated with a

⁴ In total $n = 651$ intrusions were reported. Of those, thirty-three (5.07%) were excluded as they were either reported as being actively remembered (e.g. when thinking of the study), not being associated with the trauma films, or participants did not describe them in the diary and did not give a description during the second laboratory session. The number of excluded intrusions did not differ between the conditions ($F(1, 167) < 1.68$, $ps > .197$, $BF_{10}s < .47$). For more information, see Table 2.

Table 3 Moderation of the effect of Valence on diary intrusion frequency and distress through LOC

Predictors	Intrusion Frequency (<i>N</i> =172)						Intrusion Distress (<i>N</i> =141)					
	<i>R</i> ²	<i>F</i>	<i>B</i>	<i>SE</i>	<i>t</i>	BF ₁₀	<i>R</i> ²	<i>F</i>	<i>B</i>	<i>SE</i>	<i>t</i>	BF ₁₀
Step 1	.01	.45				.01	.01	.41				.01
Valence (1 = positive, -1 = negative)			.08	.31	.27				.27	1.92	.14	
Cognition (1 = self, -1 = world)			.07	.31	.21				-.27	1.92	.14	
Valence x Cognition			.35	.31	1.11				2.08	1.92	1.08	
Internal	ΔR^2	<i>F</i> -change	<i>B</i>	<i>SE</i>	<i>t</i>	BF ₁₀	ΔR^2	<i>F</i> -change	<i>B</i>	<i>SE</i>	<i>t</i>	BF ₁₀
Step 2	.02	3.37				.82	<.01	.58				.24
IPC-Int			-.58	.32	1.84				1.50	1.97	.76	
Step 3	.02	3.35				1.26	<.01	<.01				.25
IPC Int x Valence			-.58	.32	1.83				<.01	1.98	<.01	
Powerful others												
Step 2	.04**	7.45**				5.87	.01	1.37				.35
IPC-Pow			.86	.31	2.73**				2.30	1.97	1.17	
Step 3	.02	2.99				.96	.01	1.18				.44
IPC-Pow x Valence			.54	.31	1.73				-2.14	1.97	1.09	
Chance												
Step 2	.02	3.24				.82	.02	3.00				.79
IPC-Cha			.57	.31	1.80				3.38	1.95	1.73	
Step 3	<.01	<.01				.22	<.01	.17				.27
IPC-Cha x Valence			<.01	.32	<.01				.81	1.96	.41	

p* < .05, *p* < .01. Scores on each IPC subscale were z-standardized before the analysis. Only participants who reported at least one intrusion were included in the intrusion distress analyses. *IPC* Levenson's Internal, Powerful Others, and Chance Scale, *Int* Internal, *Pow* Powerful Others, *Cha* Chance

higher number of intrusions in the diary. For intrusion distress, in contrast to our expectations, we did not find any significant interactions between the LOC subscales and Valence, demonstrating that the LOC was not a moderator of the effect of Valence on intrusion distress (for details and statistics of moderation analyses, see Table 3).

Exploratory Correlational Analyses

To further examine our pattern of results, we investigated the baseline associations between the ERT, the PTCI, and the LOC. We found the ERT bias score to be negatively correlated with the PTCI total score, demonstrating that a more positive appraisal style on the ERT was associated with less dysfunctional appraisals on the PTCI. Further, the PTCI was negatively correlated with the Internal LOC and positively correlated with the Powerful others and Chance related LOC, indicating that more dysfunctional appraisals on the PTCI were associated with a less internal and more external LOC. However, we did not find a significant association between the ERT bias score and the subscales of the IPC, indicating no association between appraisals as measured via the ERT and the LOC

(for correlation indices, see Table 4). For additional analyses and results, see the Supplementary Material.

Discussion

The present study had three aims: (i) conceptually replicating findings of previous CBM-App studies that CBM-App induces a training congruent appraisal style in analog

Table 4 Baseline correlations between dysfunctional appraisals and LOC

Measure	1	2	3	4	5
1. ERT bias score	–	–	–	–	–
2. PTCI total score	–.24*** _a	–	–	–	–
3. IPC – Int	.13 _b	–.24** _b	–	–	–
4. IPC – Pow	–.05 _b	.23** _b	–.18* _b	–	–
5. IPC – Cha	–.07 _b	.25** _b	–.25*** _b	.40*** _b	–

N = 173. Pearson correlations are presented. * < .05, ** *p* < .01, *** *p* < .001. *a* = one-sided, *b* = two-sided. *ERT* Encoding Recognition Task, *PTCI* Posttraumatic Cognitions Inventory, *IPC* Levenson's Internal, Powerful Others, and Chance Scale, *Int* Internal, *Pow* Powerful Others, *Cha* Chance

trauma and leads to corresponding effects on trauma-related symptoms, (ii) examining the effects of a CBM-App training targeting cognitions related to the world, and (iii) testing the moderating role of LOC regarding the training's effect on intrusions and intrusion distress. The CBM-App training successfully induced a valence-congruent appraisal style, i.e. positive training led to a more functional appraisal style than negative training on the ERT. Importantly, this effect was independent of the type of trained cognition (i.e. self vs. world). The training's effect generalized to posttraumatic cognitions assessed with the PTCI directly after the training. That is, PTCI scores in the positive training condition were lower than in the negative training condition at post-training, again for both types of cognitions. However, this effect was not maintained until follow-up, in that the evidence for a difference between the positive and the negative conditions was inconclusive at this timepoint. Nevertheless, we found these effects to be independent of the type of trained cognition. Further, contrary to our expectations we did not find that the self- and the world training had different effects on the respective PTCI subscales. We did not find any group differences on the IES-R, or in the diary regarding intrusion frequency or distress. Finally, and the LOC did not serve as a moderator on the training's effects on neither intrusion frequency nor distress.

Regarding our first aim, we were successful in conceptually replicating previous findings of the effect of CBM-App on appraisals (Woud et al., 2012, 2013, 2018b, c) in that the positive compared to a negative CBM-App training induced the corresponding appraisal style on the ERT. This effect also generalized to the PTCI directly after the training, which further supports the training's validity. However, in line with Woud et al., (2018c), we found this effect to be not maintained until follow-up as shown by inconclusive evidence on group differences at this time-point.

When broadening the scope of our interpretations, we did not replicate earlier findings (Woud et al., 2012, 2013, 2018c) on the effect of CBM-App training on analog trauma symptoms such as diary-intrusion frequency and distress or the IES-R, as we did not find any CBM-App effects on these outcome measures. There are a number of potential reasons why the present study did not find a training effect on analog trauma symptoms, and in particular intrusive memories. Importantly, our findings partly match earlier results on transfer effects of CBM-App in analog trauma. While one study found fewer intrusions on an intrusion diary in a positive compared to a negative CBM-App training when the training was presented after the trauma films (Woud et al., 2012), this effect did not replicate in further studies, both in the context of the trauma film paradigm and an autobiographical distressing event (Woud et al., 2013, 2018a, b, c). Taking our findings into account, this offers evidence for the conclusion that positive vs. negative CBM-App applied in

analog trauma does not reliably affect intrusion frequency. Moving on to further outcomes, our results also do not replicate earlier studies that found less intrusion distress after positive compared to negative CBM-App (Woud et al., 2013, 2018c) as well as lower analog symptoms assessed via a symptom questionnaire (Woud et al., 2018c). Here, our findings contribute to equivocal evidence on whether CBM-App in the context of analog trauma generalizes to intrusion distress or a symptom questionnaire. Some variability in results between studies may result lack of reliable effectiveness of the CBM-App in analog settings or from differences in samples and details of the administration leading to variation in how effectively the film or negative life event could function as an analog of a traumatic event. Generally, these findings raise questions about the circumstances under which effects of CBM-App may transfer to intrusive memories. One reason for our findings and the earlier equivocal evidence on the transfer of CBM-App could be that applying CBM-App training in a single session is not always sufficient to robustly affect intrusive memories and analog PTSD symptoms in general. This explanation is also put forward by Vermeulen et al., (2019) (for a discussion on dose response effects of CBM-App training, see Hallion & Ruscio, 2011; Jones & Sharpe, 2017). Given that we only found equivocal evidence for an effect of the CBM-App training on appraisals at follow-up, it seems plausible that the change in appraisal style induced was transient and insufficient to have an impact on analog symptoms. As such, future research may want to focus on testing whether the training's effect depends on the training's dose, optimally in combination with the stability of the induced training effect. Inconsistent (far) transfer effects are a more general and known problem in CBM research (cf. Hertel & Mathews, 2011; MacLeod et al., 2009) and identifying the mechanisms that may boost and hinder transfer will be an important issue for future research. Here, novel methods of investigating intrusions, for example via Ecological Momentary Assessment (EMA) (e.g. Rattell et al., 2018) could provide more sensitivity to detect potential training effects and their temporal stability.

Regarding our second aim, we found that the CBM-App training targeting cognitions related to the world was similarly effective to the CBM-App training targeting cognitions related to the self in inducing a valence-congruent appraisal style on the ERT. Further, both trainings generalized similarly to the PTCI in such that the positive conditions had lower PTCI scores at post-training compared to the negative conditions. Unexpectedly, no clear training effect was found at follow-up for the positive and negative conditions, irrespective of the targeted cognition. Further, in contrast to our expectations, we did not find training-specific effects on the PTCI subscales, thus there were no specific effects of the CBM-App training on the respective PTCI subscale from which the training's stimuli were derived. While this is

in line with the above-mentioned findings that both trainings generalized equally to the PTCI, it questions whether the self- and the world-training actually induced their respective appraisal style, or whether a general positive/negative appraisal style was induced irrespective of the appraisals' content.

Our third aim was to investigate a moderating factor potentially associated with the training's transfer to analog trauma symptoms. Specifically, we tested the moderating role of baseline LOC in the training's effect on intrusive memories. Contrary to our expectations, we did not find the LOC to be a moderator in this context, neither for intrusion frequency nor distress. The most parsimonious explanation is that the LOC is just not a moderating factor in the current paradigm. Whereas LOC has indeed been shown to be associated with PTSD symptoms (Karstoft et al., 2015; Zhang et al., 2014), Chung and Reed, (2017) found no association between LOC and dysfunctional appraisals following the experience of stillbirth. These findings are partly mirrored by our results: We found that the baseline powerful others LOC to predict more intrusions in the diary, yet we found no association for the internal or chance related LOC delivering mixed evidence for the role of LOC in the occurrence of intrusive memories. Further, through our correlational results at baseline we found LOC to be associated with dysfunctional appraisals measured via the PTCI but not via the ERT, questioning whether dysfunctional appraisals and LOC can be interpreted as strongly conceptually linked. In this context, our findings do not support the LOC as a crucial factor in the transfer of the effects of CBM-App to analog symptoms of PTSD after a trauma film, yet the partially exploratory nature of our analyses as well as the low reliability of our LOC measure warrants cautious interpretation. While LOC did not moderate the effect of CBM-App on intrusions in our study, given the analog nature of our study it will still be valuable to investigate LOC further in the context of actual trauma and PTSD, for example its role regarding the prevention and treatment of PTSD. Further, simply measuring LOC as in the current study would not allow causal conclusions to be drawn, and future studies focusing specifically on the role of LOC could benefit from including an experimental manipulation of LOC in order to investigate the impact of more internal or external LOC.

Our findings need to be interpreted in the context of a number of limitations. First, participants receiving the world training only received an ERT including scenarios related to the world and vice versa for the self training. This leaves it unclear whether the different training types actually were unsuccessful in inducing their content-specific appraisal style or whether this reflects a lack of specific generalization to the PTCI. Further, our conceptualization of intrusions in the diary as it was introduced to the participants included intrusive images as well as

thoughts, and a combination of both. While this is in line with prior studies applying CBM-App in an analog trauma context we cannot rule out that this broad conceptualization of intrusions led to an inclusion of other concepts that are not easy to distinguish for participants such as ruminative thoughts about the trauma films. Further, we found low reliability indices for the IPC with which LOC was assessed and the hyperarousal subscale of the IES-R, which in turn affects the reliability of our conclusions concerning the results using these measures. We further note that there were some statistically significant differences between the conditions in negative mood after watching the trauma film. While this adds another caveat to the interpretation of our results, these differences were small compared to the overall changes in mood observed, and it is difficult to see how they might have influenced our findings, if at all. However, it would be preferable if experimental conditions were balanced on all pre-manipulation measures, and in future methods such as stratifying randomization based on mood change while watching the film could be used to avoid such imbalances. Additionally, while the application of the trauma film paradigm in a healthy sample is commonly used as an experimental psychopathology model of trauma exposure (cf. James et al., 2016), our study has limited generalizability to an actual trauma context. Further, although we used the same film as previous studies (Woud et al., 2012, 2013), participants in our study recorded fewer intrusions (e.g. in the negative condition) compared to these previous studies. This might indicate that some differences between the studies (e.g. sample, details of administration) may have led to the film being a less effective analog trauma in our study and reduced our ability to find effects of the training. Also, our study did not include a neutral control condition. Our findings therefore cannot deliver evidence on whether analog trauma-related symptoms were decreased or increased by the positive or negative training, respectively.

To conclude, our CBM-App training successfully induced a valence-congruent appraisal style following analog trauma, using both a CBM-App training targeting appraisals of the self, or a newly developed training targeting cognitions of the world, providing evidence for the validity of CBM-App as an experimental manipulation of trauma-relevant appraisals as measured via the ERT and the PTCI. However, given that we found no transfer to the IES-R or intrusions, our study highlights the importance of investigating factors that are associated with an (un-)successful transfer of CBM-App effects to such analog trauma-related symptoms in future research. While we investigated the LOC as such a potential moderating factor, our study suggests that its role in the transfer appears negligible.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10608-021-10257-x>.

Authors' contributions Julie Krans, Simon E. Blackwell, Jan C. Cwik, Jürgen Margraf, and Marcella L. Woud contributed to the study design. Felix Würtz was involved in data collection. Felix Würtz, Julie Krans, Simon E. Blackwell, and Marcella L. Woud were involved in the data analysis and interpretation. Felix Würtz, Simon E. Blackwell, and Marcella L. Woud prepared the first version of the manuscript. All authors contributed to the refinement of the manuscript and approved it.

Funding Open Access funding enabled and organized by Projekt DEAL. Felix Würtz is supported by a doctoral scholarship of the Studienstiftung des deutschen Volkes. Marcella L. Woud is funded by the Deutsche Forschungsgemeinschaft (DFG) via the Emmy Noether Programme (WO 2018/3–1) and the Daimler and Benz Foundation (32–12/4). The funding bodies had no role in the design of the study, the collection, analysis, and interpretation of the data, or the preparation of the manuscript.

Data availability The anonymized datasets generated and analyzed during the current study as well as used analysis scripts are available on the Open Science Framework (OSF) via the following link: <https://osf.io/hn7bz>

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study was approved by the local ethics committee of the Psychology department at RUB (approval number: 338). All participants gave written informed consent prior to participation.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Manual for the beck depression inventory-II. *San Antonio, TX: Psychological Corporation*, 1, 82.
- Beierl, E. T., Böllinghaus, I., Clark, D. M., Glucksman, E., & Ehlers, A. (2019). Cognitive paths from trauma to posttraumatic stress disorder: A prospective study of Ehlers and Clark's model in survivors of assaults or road traffic collisions. *Psychological Medicine*. <https://doi.org/10.1017/S0033291719002253>
- Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of posttraumatic stress disorder. *Psychological Review*, 103(4), 670–686. <https://doi.org/10.1037/0033-295X.103.4.670>
- Brown, L. A., Belli, G. M., Asnaani, A., & Foa, E. B. (2019). A review of the role of negative cognitions about oneself, others, and the world in the treatment of PTSD. *Cognitive Therapy and Research*, 43(1), 143–173. <https://doi.org/10.1007/s10608-018-9938-1>
- Bryant, R. A., & Guthrie, R. M. (2005). Maladaptive appraisals as a risk factor for posttraumatic stress. *Psychological Science*, 16(10), 749–752. <https://doi.org/10.1111/j.1467-9280.2005.01608.x>
- Bryant, R. A., & Guthrie, R. M. (2007). Maladaptive self-appraisals before trauma exposure predict posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 75(5), 812–815. <https://doi.org/10.1037/0022-006X.75.5.812>
- Brybaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, 1(1), 1–20. <https://doi.org/10.5334/joc.10>
- Chung, M. C., Preveza, E., Papandreu, K., & Prevezas, N. (2006). The relationship between posttraumatic stress disorder following spinal cord injury and locus of control. *Journal of Affective Disorders*, 93(1–3), 229–232. <https://doi.org/10.1016/j.jad.2006.02.021>
- Chung, M. C., & Reed, J. (2017). Posttraumatic stress disorder following stillbirth: Trauma characteristics, locus of control. *Posttraumatic Cognitions*. *Psychiatric Quarterly*, 88(2), 307–321. <https://doi.org/10.1007/s11126-016-9446-y>
- Dalgleish, T. (2004). Cognitive approaches to posttraumatic stress disorder: The evolution of multirepresentational theorizing. *Psychological Bulletin*, 130(2), 228–260. <https://doi.org/10.1037/0033-2909.130.2.228>
- Ehlers, A. (1999). *Posttraumatische belastungsstörung*. Hogrefe Göttingen.
- Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, 38(4), 319–345. [https://doi.org/10.1016/S0005-7967\(99\)00123-0](https://doi.org/10.1016/S0005-7967(99)00123-0)
- Ehring, T., Ehlers, A., & Glucksman, E. (2008). Do cognitive models help in predicting the severity of posttraumatic stress disorder, phobia, and depression after motor vehicle accidents? A prospective longitudinal study. *Journal of Consulting and Clinical Psychology*, 76(2), 219–230. <https://doi.org/10.1037/0022-006X.76.2.219>
- Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. F., & Orsillo, S. M. (1999). The Posttraumatic Cognitions Inventory (PTCI): Development and validation. *Psychological Assessment*, 11(3), 303–314. <https://doi.org/10.1037/1040-3590.11.3.303>
- Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy*, 20(2), 155–176. [https://doi.org/10.1016/S0005-7894\(89\)80067-X](https://doi.org/10.1016/S0005-7894(89)80067-X)
- Gómez de La Cuesta, G., Schweizer, S., Diehle, J., Young, J., & Meiser-Stedman, R. (2019). The relationship between maladaptive appraisals and posttraumatic stress disorder: a meta-analysis. *European Journal of Psychotraumatology*. <https://doi.org/10.1080/20008198.2019.1620084>
- Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, 137(6), 940–958. <https://doi.org/10.1037/a0024355>
- Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck Depressions-Inventar (BDI-II)*. Harcourt Test Services.
- Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification: Past perspectives, current findings, and future applications. *Perspectives on Psychological Science*, 6(6), 521–536. <https://doi.org/10.1177/1745691611421205>
- Holmes, E. A., & Steel, C. (2004). Schizotypy: A vulnerability factor for traumatic intrusions. *Journal of Nervous and Mental Disease*, 192(1), 28–34. <https://doi.org/10.1097/01.nmd.0000105997.28152.4b>
- James, E. L., Lau-Zhu, A., Clark, I. A., Visser, R. M., Hagenars, M. A., & Holmes, E. A. (2016). The trauma film paradigm as an experimental psychopathology model of psychological trauma:

- Intrusive memories and beyond. *Clinical Psychology Review*, 47, 106–142. <https://doi.org/10.1016/j.cpr.2016.04.010>
- 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>
- Karstoft, K. I., Armour, C., Elklit, A., & Solomon, Z. (2015). The role of locus of control and coping style in predicting longitudinal PTSD-trajectories after combat exposure. *Journal of Anxiety Disorders*, 32, 89–94. <https://doi.org/10.1016/j.janxdis.2015.03.007>
- Kleim, B., Ehlers, A., & Glucksman, E. (2007). Early predictors of chronic post-traumatic stress disorder in assault survivors. *Psychological Medicine*, 37(10), 1457–1467. <https://doi.org/10.1017/S0033291707001006>
- Kleim, B., Grey, N., Wild, J., Nussbeck, F. W., Stott, R., Hackmann, A., & Ehlers, A. (2013). Cognitive change predicts symptom reduction with cognitive therapy for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 81(3), 383–393. <https://doi.org/10.1037/a0031290>
- Koster, E. H. W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology*, 118(1), 1–4. <https://doi.org/10.1037/a0014379>
- Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry*, 54(4), 337. <https://doi.org/10.1001/archpsyc.1997.01830160065009>
- Krampen, G. (1979). Differenzierungen des Konstruktes der Kontrollüberzeugung Deutsche Bearbeitung und Anwendung der IPC-Skalen. *Zeitschrift Für Experimentelle Und Angewandte Psychologie*, 26(4), 573–595.
- Lancaster, S. L., Rodriguez, B. F., & Weston, R. (2011). Path analytic examination of a cognitive model of PTSD. *Behaviour Research and Therapy*, 49(3), 194–201. <https://doi.org/10.1016/j.brat.2011.01.002>
- Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *Das State-Trait-Angstinventar (STAI) Theoretische Grundlagen und Handanweisung*. Weinheim: Beltz.
- Lenth, R. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. Retrieved from <https://cran.r-project.org/package=emmeans>.
- Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, 41(3), 397–404.
- MacLeod, C., Koster, E. H. W., & Fox, E. (2009). Whither cognitive bias modification research? commentary on the special section articles. *Journal of Abnormal Psychology*, 118(1), 89–99. <https://doi.org/10.1037/a0014878>
- Maercker, A., & Schützwohl, M. (1998). Erfassung von psychischen Belastungsfolgen : Die Impact of Event Skala- revidierte Version (IES-R). *Diagnostica*, 44(3), 130–141. <https://doi.org/10.1037/t55092-000>
- 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>
- McNally, R. J., & Woud, M. L. (2019). Innovations in the study of appraisals and PTSD: A commentary. *Cognitive Therapy and Research*, 43(1), 295–302. <https://doi.org/10.1007/s10608-018-09995-2>
- Meyer, T., Brewin, C. R., King, J. A., Nijmeijer, D., Woud, M. L., & Becker, E. S. (2020). Arresting visuospatial stimulation is insufficient to disrupt analogue traumatic intrusions. *PLoS ONE*, 15(2), 1–20. <https://doi.org/10.1371/journal.pone.0228416>
- Morey, R. D., & Rouder, J. N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12–4.2. Retrieved from <https://cran.r-project.org/package=BayesFactor>.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2020). nlme: Linear and Nonlinear Mixed Effects Models. Retrieved from <https://cran.r-project.org/package=nlme>.
- R Core Team. (2020). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>.
- Rattel, J. A., Grünberger, L. M., Reichenberger, J., Liedlgruber, M., Miedl, S. F., Blechert, J., & Wilhelm, F. H. (2018). Frequency of intrusions and appraisal of related distress after analogue trauma: a comparative ecological momentary assessment methods study. *Cognitive Therapy and Research*. <https://doi.org/10.1007/s10608-018-9941-6>
- Rotter, J. B. (1966). General expectations for internal versus external control of reinforcement. *Psychological Monographs General and Applied*, 80(1), 1–28. <https://doi.org/10.1037/h0092976>
- RStudio Team. (2016). RStudio: Integrated Development Environment for R. Boston, MA: RStudio, Inc. Retrieved from <http://www.rstudio.io.com/>.
- Schumm, J. A., Dickstein, B. D., Walter, K. H., Owens, G. P., & Chard, K. M. (2015). Changes in posttraumatic cognitions predict changes in posttraumatic stress disorder symptoms during cognitive processing therapy. *Journal of Consulting and Clinical Psychology*, 83(6), 1161–1166. <https://doi.org/10.1037/ccp0000040>
- Singman, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). afex: Analysis of Factorial Experiments. Retrieved from <https://cran.r-project.org/package=afex>
- Spielberger, C. D., Gorsuch, R., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory*. Consulting Psychologists Press.
- Startup, M., Makgekgenene, L., & Webster, R. (2007). The role of self-blame for trauma as assessed by the Posttraumatic Cognitions Inventory (PTCI): A self-protective cognition? *Behaviour Research and Therapy*, 45(2), 395–403. <https://doi.org/10.1016/j.brat.2006.02.003>
- van Heemstra, H. E., Scholte, W. F., Ehring, T., & Boelen, P. A. (2020). Contextualizing cognitions: the relation between negative post-traumatic cognitions and post-traumatic stress among palestinian refugees. *International Journal of Cognitive Therapy*, 13(2020), 159–172. <https://doi.org/10.1007/s41811-020-00066-7>
- Vermeulen, M., Brown, A. D., Raes, F., & Krans, J. (2019). Decreasing event centrality in undergraduates using cognitive bias modification of appraisals. *Cognitive Therapy and Research*, 43(1), 214–225. <https://doi.org/10.1007/s10608-018-9936-3>
- Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale - Revised. In J. P. Wilson & T. M. Keane (Eds.), *The Impact of Events Scale—Revised* (pp. 399–411). Guilford Press.
- Woud, M. L., & Becker, E. S. (2014). Editorial for the special issue on cognitive bias modification techniques: an introduction to a time traveller’s tale. *Cognitive Therapy and Research*, 38(2), 83–88. <https://doi.org/10.1007/s10608-014-9605-0>
- Woud, M. L., Blackwell, S. E., Shkreli, L., Würtz, F., Cwik, J. C., Margraf, J., & 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*. <https://doi.org/10.1159/000514166>
- Woud, M. L., Blackwell, S. E., Steudtler-Schmiedgen, S., Browning, M., Holmes, E. A., Harmer, C. J., & Reinecke, A. (2018a). Investigating d-cycloserine as a potential pharmacological enhancer of an emotional bias learning procedure. *Journal of Psychopharmacology*, 32(5), 569–577. <https://doi.org/10.1177/0269881118754679>
- Woud, M. L., Cwik, J. C., Blackwell, S. E., Kleim, B., Holmes, E. A., Adolph, D., & Margraf, J. (2018b). Does napping enhance the effects of Cognitive Bias Modification-Appraisal training? *An Experimental Study*. *PLOS ONE*, 13(2), e0192837. <https://doi.org/10.1371/journal.pone.0192837>

- 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., Verwoerd, J., & Krans, J. (2017). Modification of cognitive biases related to posttraumatic stress: A systematic review and research agenda. *Clinical Psychology Review, 54*, 81–95. <https://doi.org/10.1016/j.cpr.2017.04.003>
- Woud, M. L., Zlomuzica, A., Cwik, J. C., Margraf, J., Shkreli, L., Blackwell, S. E., & Ehring, T. (2018c). 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>
- Zhang, W., Liu, H., Jiang, X., Wu, D., & Tian, Y. (2014). A longitudinal study of posttraumatic stress disorder symptoms and its relationship with coping skill and locus of control in adolescents after an earthquake in China. *PLoS ONE, 9*(2), 1–7. <https://doi.org/10.1371/journal.pone.0088263>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.