International Journal of Cognitive Therapy (2024) 17:700–724
https://doi.org/10.1007/s41811-024-00218-z

An Anger-based Approach-Avoidance Modification Training Targeting Dysfunctional Beliefs in Adults with Elevated Stress – Results from a Randomized Controlled Pilot Study

Marie Keinert[1] · Lena Schindler-Gmelch[1] · Bjoern M. Eskofier[2] · Matthias Berking[1]

Accepted: 9 June 2024 / Published online: 2 July 2024
© The Author(s) 2024

Abstract

We examined the feasibility, acceptability, and clinical potential of a smartphone-based Approach-Avoidance Modification Training (AAMT) that required users to deliberately display anger and positive emotions (AAMT-AP) to move (dys-)functional stress-related beliefs away from and towards themselves, respectively. The per protocol sample consisted of $N = 30$ participants randomized to an AAMT-AP condition, a swipe control condition, or an inactive control condition. Feasibility and acceptability of the intervention and study design were assessed directly and changes in clinical outcomes (perceived stress (primary outcome), dysfunctional beliefs, symptoms of depression, and emotion regulation skills) one week after intervention completion. The results indicated that AAMT-AP was feasible and acceptable to participants. Moreover, notable effect sizes for clinical outcomes in comparison with both control conditions provided preliminary evidence for its therapeutic potential. Future research should evaluate the efficacy of AAMT-AP in sufficiently powered randomized controlled trials.

Keywords Approach-avoidance modification training · Stress · Dysfunctional beliefs · Smartphone-based intervention · Anger · Emotions

✉ Marie Keinert
marie.keinert@fau.de

[1] Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

[2] Machine Learning and Data Analytics Lab, Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

Background

Chronic stress levels have been rising steadily in the past decades on a global scale (UK Health & Safety Executive, 2020), a trend exacerbated by the COVID-19 pandemic (American Psychological Association, 2020, 2021). Chronic stress burdens individuals by causing physical and mental illness (e.g., Cohen et al., 2007; Hammen, 2005; O'Connor et al., 2021), and impacts communities with additional direct and indirect costs (Hassard et al., 2018). The immense impact of chronic stress emphasizes the importance of creating easily accessible and effective interventions that can reduce stress and prevent the development of stress-related pathologies.

According to Lazarus and Folkman's seminal transactional stress model (Lazarus & Folkman, 1984), stress emerges when an individual perceives a demanding situation to threaten significant personal goals, while available coping resources are deemed insufficient to meet those situational demands. Personal beliefs influence subjective appraisals and determine a person's cognitions, emotions, and coping efforts with respect to both the situation and its subsequent outcome (Lazarus & Folkman, 1984). Accordingly, since dysfunctional beliefs contribute to the emergence of stress (Kuroda, 2016; Otto et al., 1997), they should be considered a prime target for therapeutic interventions. In traditional cognitive behavioral therapy, sophisticated techniques of cognitive restructuring are used to question the general validity of dysfunctional beliefs, and to introduce patients to more functional (i.e., adaptive) perspectives (Clark, 2013). Complex cognitive restructuring methods, however, are difficult to integrate into readily accessible, cost-effective, web-based interventions. Moreover, cognitive restructuring targets the so-called propositional processing of dysfunctional beliefs (Gawronski & Bodenhausen, 2006), and requires higher-order cognitive capacities that have been shown to be impaired when under chronic stress (Allen et al., 2017; Girotti et al., 2018). To address these issues, a more feasible alternative could be to focus instead on associative, rather than propositional processing of dysfunctional, stress-related beliefs (Gawronski & Bodenhausen, 2006).

One promising intervention for such an endeavor could be Approach-Avoidance Modification Trainings (AAMTs). AAMTs are designed to require individuals to repeatedly draw disorder-specific functional (i.e., adaptive) stimuli toward themselves (approach) and/or drive dysfunctional (i.e., maladaptive) stimuli away from themselves (avoidance) by either maneuvering a computer gaming joystick or by executing a swiping-movement on a smartphone screen (Cervera-Torres et al., 2021; Kollei et al., 2017; Lukas & Berking, 2018; Lukas et al., 2019; Rinck & Becker, 2007). Based on the premise that actions performed in AAMTs are associated with a certain valence, the evaluation of the action and its valence are then inferred to the original stimuli that triggered the reactions (Van Dessel et al., 2019). For example, if a person drives or pushes away an undesirable stimulus, this avoidance-gesture is evaluated as negative, and this negative valence is also transferred to the respective stimulus. Drawing on consistency theory (Festinger, 1957) it can be argued that motioning a negatively appraised stimulus toward

oneself or motioning a positively appraised stimulus away from oneself creates cognitive dissonance. Since the behavioral reactions are prescribed and, thus, individuals cannot change their behavior in the training, they are likely to reduce cognitive dissonance by reevaluating the stimuli in accordance with their behavioral response (i.e., reevaluate stimuli that are approached to be positive and stimuli that are avoided to be negative).

Empirically, results on the efficacy of AAMTs in the mental health context are mixed (for review, see Loijen et al., 2020). Particularly in the field of alcohol use disorder, studies on joystick-AAMTs have shown good efficacy in changing approach biases towards alcohol and in decreasing relapse rates (Eberl et al., 2013; Rinck et al., 2018). In other areas of mental health, however, the evidence is less clear. For example, in social anxiety disorder and bulimic eating disorder, some studies found positive effects on symptom reduction or modified approach-avoidance biases (Brockmeyer et al., 2015; Taylor & Amir, 2012), whereas others did not (Asnaani et al., 2014; Brockmeyer et al., 2019). Studies on swipe-based AAMTs in subclinical samples were more conclusive: They found beneficial effects on symptoms of procrastination ($d = 0.75$, Lukas & Berking, 2018) body dissatisfaction ($d = 0.62$, Kollei et al., 2017), and alexithymia ($d = 1.14$, Lukas et al., 2019). However, in those studies, the AAMT intervention was complemented by a counseling session; hence, the effects could not clearly be attributed exclusively to AAMT.

To our knowledge, only three studies to date utilized AAMTs in the context of stress (Becker et al., 2016 reporting on two separate studies; Ferrari et al., 2018). In all these studies, participants were instructed to approach and/or avoid positive and negative emotional pictures. The effects of the training on both approach-avoidance tendencies and on self-reported stress were assessed before and after a behavioral stress task administered after the AAMT. The initial investigation of Becker and colleagues found no effects on stress (Becker et al., 2016). In contrast, their second study, (in which a sad mood state was induced before the intervention), revealed lower self-reported tension in dysphoric individuals (but not non-dysphoric ones) after completing the behavioral stress task in the AAMT condition compared to participants in the control group ($d = 1.15$). This notwithstanding, no effects were found prior to completing the behavioral stress task (Becker et al., 2016). In the third study, conducted by Ferrari and colleagues, none of the training conditions affected self-reported or behavioral indicators of stress (Ferrari et al., 2018).

The limited literature suggests a significant need to improve and prove the efficacy of AAMTs ability to reduce stress. Based on the inferential account proposed by Van Dessel and colleagues (2019), it is conceivable that this might be achieved by specifically enhancing the valence of approach-avoidance responses. Assuming that emotions embody inherent valence far stronger than the valence associated with simple wrist movements with a joystick or via screen swipe (Smith & Ellsworth, 1985), one could argue that utilizing the deliberate expression of negative emotions to move dysfunctional stimuli away from oneself and deliberate expression of positive emotions to draw functional stimuli toward oneself leads to a significantly stronger manipulation of the stimuli's valence than is possible with traditional AAMTs. Furthermore, consistency theory (Festinger, 1957) would have us predict that intentionally responding to a shared/rejected belief with a positive/negative emotion will

generate far greater dissonance than responding to such beliefs with simple wrist movements. Thus, emotion focused AAMTs should generate more pressure to share functional beliefs and to reject dysfunctional beliefs in order to reduce this dissonance. Moreover, emotions are seemingly inextricably associated with strong approach and avoidance tendencies (Elliot et al., 2013). Associating functionally adaptive beliefs with approach tendencies and dysfunctional beliefs with avoidance tendencies also likely increases the salience of the former and reduces the salience of the latter. Finally, emotions are associated with specific, biologically determined behavioral tendencies (Lazarus, 1991). An emotion particularly relevant to the stress context is anger, which can arise as a result of stress (Novaco, 1979). Anger triggers fighting against something or someone that threatens to hinder one's goal attainment (Berking & Whitley, 2014; Novaco, 2010, 2016). Analogously, anger might be used to combat dysfunctional beliefs. It is noteworthy that anger not only incites the necessary mindset to confront threatening, dysfunctional beliefs, but arguably also the heightened state of arousal conducive to promoting cognitive change and of the perception of enhanced personal control over negative events (Remington et al., 2000; Lerner et al., 2015). Thus, it can be argued that anger is a particularly promising candidate when it comes to overcoming dysfunctional stress-related beliefs.

Following a similar line of thinking, positive emotions such as joy, pride, or love, could be very effective approach responses in AAMTs. These affirming emotions carry strong positive valence, they are experienced in the context of positively appraised stimuli, and they are inherently associated with approach motivation (Elliot et al., 2013; Smith & Ellsworth, 1985). Thus, compared to simple arbitrary wrist movements, positive emotions should have a much greater capacity to increase positive valence for functional stimuli, to maximize cognitive dissonance (in response to unshared functional beliefs), and to cognitively link functionally adaptive stimuli with approach motivation.

To explore the potential of an AAMT that utilizes both anger and positive emotions (AAMT-AP), we developed a smartphone intervention, in which participants approached stress-reducing, functionally adaptive beliefs, and avoided stress-enhancing, dysfunctional beliefs by deliberately displaying positive emotions or anger, respectively. As there currently exist no interventions comparable to this innovative AAMT-AP, the main objectives of this pilot study were to collect preliminary feasibility and acceptability data for the intervention and to examine the merits of our study design. Our secondary aim was to explore the clinical potential of the intervention in individuals suffering from elevated stress.

Methods

Study Design, Participants, and Procedures

In order to evaluate the feasibility and acceptability of AAMT-AP and to explore its clinical potential, we conducted a randomized, controlled pilot study at the Department of Clinical Psychology and Psychotherapy of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). The study was preregistered in the German Clinical

Trials Register (Deutsches Register Klinischer Studien; DRKS00023007) and conducted in accordance with the CONSORT recommendations (Schulz et al., 2010; see Supplementary Materials Table S1 for the CONSORT checklist). All study procedures were approved by the German Psychological Societies ethics committee (BerkingMatthias2020-09-10AM) and in accordance with the Declaration of Helsinki's research guidelines. We also evaluated AAMTs utilizing emotions other than anger (e.g., sadness, disgust, and anxiety). Findings from these components of the study will be reported elsewhere. For a detailed description of the study protocol, see Keinert et al. (2023).

Recruitment commenced in July 2020 and was completed in May 2022. Participants were recruited via social media, several university e-mail lists, and by posting flyers on public notice boards at highly frequented places in Erlangen. Interested individuals were directed to an online screening questionnaire administered via the platform Unipark (https://www.unipark.com). After providing written informed consent, prospective participants were screened for the following inclusion criteria: (1) a score of 19 or above on the Perceived Stress Scale-10 (PSS-10; Klein et al., 2016), (2) minimum age 18 years, and (3) informed consent. Exclusion criteria were: (1) presence of psychotic or bipolar disorders, (2) physically impaired emotional expression (e.g., facial paralysis), (3) smoking in excess of ten cigarettes daily (owing to potential confounding influences on cortisol samples), and (4) current suicidal ideation. Individuals who met all the inclusion and none of the exclusion criteria were provided written information about study procedures via e-mail. They were randomly assigned to study groups according to a randomization list with an allocation ratio of 1:1. We performed block randomization with a block size of eight using random numbers in Microsoft Excel. The randomization list was generated by a postdoctoral researcher not otherwise involved in the study. The allocation based on the randomization list was performed by a research assistant. The subsample reported on in the current paper consisted of $n = 43$ individuals allocated to either the AAMT-AP condition, to the swipe control condition, or to the inactive control condition. Because of a coding error in the experimental software, seven participants not meeting the PSS-10 cut-off at screening were inadvertently enrolled in the study. The PSS-10 score of these participants was between 5 and 10 points below the required cut-off ($M = 7.4$, $SD = 2.07$). Six further participants had not completed the training and/or assessments as intended. These errors were remedied by excluding all 13 individuals from the per-protocol analyses of clinical outcomes. Recruitment continued until the necessary sample size for the per-protocol analyses of $n = 10$ individuals per group was attained with fully eligible participants. The sample size was determined following Bell and colleague's (2018) recommendations for pilot studies. For the feasibility analyses, we did include the data from the otherwise excluded participants, except for the three participants who did not complete the feasibility measures directly after the training (see Fig. 1). Thus, the sample used for the feasibility evaluation consisted of $n = 40$ individuals, the sample for the per-protocol analyses of clinical outcomes consisted of $n = 30$ individuals (i.e., 10 participants in the AAMT-AP, 10 in the swipe control, and 10 in the inactive control condition; see Fig. 1), and the sample for the complementary intention-to-treat (ITT)-analyses consisted of n = 36 individuals. A majority (80%) of the primary, per

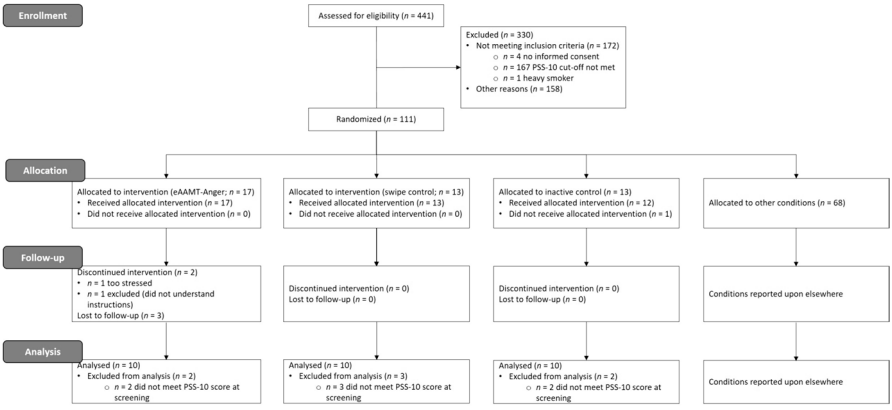


Fig. 1 Study Flow [Note. PSS-10 = Perceived Stress Scale-10; AAMT-AP = Approach-Avoidance Modification Training with anger and positive emotions]

protocol sample was female. The mean age of this sample was 24.2 years (SD = 4.4; range: 18–33 years). Further sociodemographic and baseline clinical characteristics are summarized in Table 1. Sociodemographic and baseline clinical characteristics of the feasibility and ITT-samples are reported in the supplementary materials (Table S2 and S3).

As indicated in Fig. 1, measures were assessed directly before the training (pre) and directly after the training (between). A one-week follow-up (post) survey was sent to participants by e-mail. Control group participants received all questionnaires via e-mail. All questionnaires were assessed using the web-based assessment tool Unipark (www.unipark.com). Psychology students received course credit for participation. Other participants had the option through May 2021 to win 500 € in a lottery or to receive compensation in the amount of 20 € from June 2021.

Intervention and Control Conditions

Participants in the AAMT-AP and the swipe control condition completed 20–30 min of the app-based training on four consecutive days using a study-issued smartphone. Both groups were first introduced to the AAMT concept and provided condition-specific instructions on how to execute the training, both (1) by the experimenter and (2) via the smartphone application itself. AAMT stimuli consisted of 78 stress-related belief statements with corresponding images (see Figure S1 in the supplementary materials for examples). The pool of 30 stress-enhancing beliefs (e.g., "If I fail at my work, then I am a failure as a person") and 48 stress-reducing beliefs (e.g., "I am allowed to make mistakes") was derived from the Dysfunctional Attitude Scale (DAS; Rojas et al., 2015) by an expert group of clinicians and stress researchers. The training sessions were composed of 60 trials each, including all 30 stress-enhancing beliefs with corresponding photos, and a random sample of 30 stress-reducing beliefs with photos; stimuli presentation was in random order at a 1:1 ratio. The mean duration of one trial in the AAMT-AP condition was 10.03 s.

Table 1 Sociodemographic and Clinical Baseline-Characteristics of the Per Protocol Sample

	AAMT-AP $n=10$	Swipe control $n=10$	Inactive control $n=10$	Statistics	p
Female gender, *n* (%)	7 (70)	9 (90)	8 (80)	Fisher's exact test	0.85
Age, *M (SD)*	22.4 (3.17)	24.7 (4.92)	25.5 (4.72)	$F(2,27)=1.37$	0.27
Age range	19–29	18–32	19–33		
Education, *n* (%)				Fisher's exact test	0.70
≤ 10 years	1 (10)	0 (0)	0 (0)		
> 10 years	7 (70)	6 (60)	6 (6)		
University degree	2 (20)	4 (40)	4 (40)		
Occupation, *n* (%)				Fisher's exact test	1
Student	9 (90)	8 (80)	8 (80)		
employed	1 (10)	2 (20)	2 (20)		
Relationship status, *n* (%)				Fisher's exact test	1
single	10 (100)	9 (90)	9 (90)		
married	0 (0)	1 (1)	1 (1)		
Psychotherapy, *n* (%)					
Former	1 (10)	0 (0)	1 (10)	Fisher's exact test	1
Current	1 (10)	0 (0)	0 (0)	Fisher's exact test	1
PSS-10, *M (SD)*	20 (5.68)	18.7 (7.53)	19.6 (5.68)	$F(2,27)=0.11$	0.90
DAS-A, *M (SD)*	28.3 (10.17)	28.1 (12.23)	33 (9.07)	$F(2,27)=0.69$	0.51
DAS-H, *M (SD)*	19.3 (6.83)	16.9 (5.92)	21.6 (9.25)	$F(2,27)=1.00$	0.38
CES-D, *M (SD)*	19.1 (5.3)	15.6 (8.37)	21 (9.96)	$F(2,27)=1.41$	0.33
ERSQ, *M (SD)*	2.26 (0.56)	2.57 (0.59)	2.18 (0.43)	$F(2,27)=1.55$	0.23

Note. AAMT-AP = Approach-Avoidance Modification Training with anger and positive emotions; PSS-10 = Perceived Stress Scale-10; DAS-A = Dysfunctional Attitude Scale-Approval; DAS-H = Dysfunctional Attitude Scale-Helpfulness; CES-D = Center for Epidemiological Studies Depression Scale; ERSQ = Emotion Regulation Skills Questionnaire

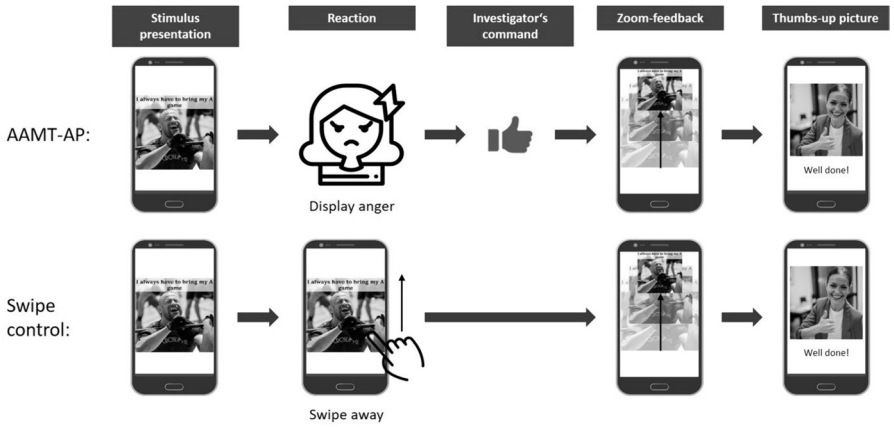


Fig. 2 Exemplary Trial Flow of the AAMT-AP and the Swipe Control Conditions [Note. AAMT-AP = Approach-Avoidance Modification Training with anger and positive emotions]

(*SD* = 5.15) and in the swipe control condition 3.53 s. (*SD* = 1.62). After participants performed each respective approach- or avoidance-response correctly, the app provided an immediate zoom-feedback (i.e., zoom-in for approach, zoom-out for avoidance) and a thumbs-up picture (see Figure S2 in the supplementary materials for example). When participants executed an incorrect response, the app provided a respective feedback. Figure 2 illustrates an exemplary trial for both the AAMT-AP and the swipe control conditions.

In the AAMT-AP condition, participants were instructed to either draw the presented stimuli toward themselves or to drive the stimuli away by deliberately displaying positive emotions or anger, respectively. To maximize intervention appeal, we introduced some variation by using multiple positive emotions. Specifically, each training session consisted of three blocks, with the positive emotion to be displayed by participants alternating after each block. The positive emotions were always performed in the same order (Day 1: joy, relaxation, and love; Day 2: excitement, tranquility, and gratitude; Day 3: happiness, resolve, and contentedness; Day 4: courage, confidence, and pride). Short instructive video clips, integrated into the app, served to demonstrate how to display the respective emotions with (1) a facial expression, (2) a verbal statement, and (3) a body gesture consistent with the emotion. Anger, for example, could be displayed by an angry facial expression, by slamming one's fist on the table, and saying, "Such a stupid attitude, I will not listen to its advice!". In an adjacent area behind a partition, an experimenter remotely used a tablet to control the smartphone-app. The two devices were connected via Bluetooth- or Wi-Fi-connection established using Google's Nearby API (https://developers.google.com/nearby/). To determine whether participants displayed the correct emotion, the training session was recorded with a video camera and observed by the experimenter via live stream behind said partition. After participants' emotion display, the experimenter commanded the smartphone-app to provide the respective zoom-feedback.

This was followed by an experimenter rating of the emotion display and participants' rating of the subjective intensity of the emotion.

In the swipe control condition, participants did not display emotions, but approached or avoided the presented stimuli by performing finger swipe-movements "towards" themselves (i.e., downwards) or "away from" themselves (i.e., upwards), respectively. In this condition, the app was not controlled by the experimenter, but operated automatically.

Participants in the inactive control condition did not engage in any training. They received links to all study questionnaires (pre, between, and post) via e-mail.

Measures

Feasibility and Acceptability

To evaluate the *feasibility of the intervention*, we recorded technical problems, adherence, usability, acceptability, and adverse events. The experimenter recorded technical difficulties that arose during sessions; it was decided (a priori) to consider the intervention feasible if at least 95% of the sessions transpired without interruptions of more than five minutes related to technical problems. With regards to adherence, we considered the intervention feasible if at least 50% of participants completed the intervention as instructed. This cut-off was established in collaboration with an expert for smartphone-based interventions. With regards to usability, at T2 participants were asked to rate how understandable the app-instructions were on a 6-point Likert scale (0 = *not understandable* to 5 = *very understandable*) and how easy it was for them to display the different emotions at the beginning and at the end of the training (0 = *very difficult* to 5 = *very easy*). Acceptability of the intervention was operationalized by two items administered at T2 which asked participants to rate how helpful the training was in modifying their stress-related beliefs and whether they would recommend the training to a friend on a 6-point Likert scale (0 = *not at all* to 5 = *very much*). Moreover, participants were asked to evaluate the training via an open-ended question. Adverse events, i.e., severe mood decompensation during the training, were assessed and recorded by the experimenter. A further possible harm of the training might be an increase in the experience of anger caused by its repeated display and reinforcement during the training. Therefore, we assessed the experience of anger and other affective states during the past week at T1-T3 with an adapted version of the Positive and Negative Affect Schedule (Krohne et al., 1996) on a five-point Likert scale (0 = *not at all* to 4 = *very much*).

With regards to the *feasibility of the study design*, we collected data on the proportion of eligible participants agreeing to participate in the study. We additionally analyzed the number of completed follow-up assessments and considered the design feasible if at least 75% of all assessments were completed.

Clinical Outcomes

The primary clinical outcome was change in perceived stress from pre to post as assessed with the PSS-10. The PSS-10 is a widely used self-report measure consisting of ten items measuring perceived personal stress during the past week. The items are rated on a 5-point Likert scale (0 = *never* to 4 = *very often*). The questionnaire has shown good validity and reliability in several studies (Cronbach alpha = 0.84; Beutel et al., 2016; Klein et al., 2016; Reis et al., 2019).

Secondary outcomes included the approval of stress-related beliefs and the perceived helpfulness of stress-related beliefs. Beliefs were derived from the short version of the DAS based on determinations by an expert group supervised by the senior author regarding the alignment of individual items for clinically relevant stress problems. One of the ten items represented a functional attitude and nine items represented dysfunctional attitudes (see Table S4 of the supplementary materials for the entire measure). The DAS is a self-report instrument originally developed to assess dysfunctional attitudes in patients diagnosed with depression. In our adapted version, participants rate the items twice on a 7-point Likert scale (approval [DAS-A]: 1 = *fully disagree* to 7 = *fully agree* and helpfulness [DAS-H]: 1 = *not at all helpful* to 7 = *very helpful*). The singular functional attitude item was reversely coded to compute a total score.

Additionally, we assessed symptoms of depression with the German version of the Center for Epidemiological Studies Depression Scale (CES-D; Hautzinger et al., 2012). The CES-D is a commonly used self-report measure assessing the frequency of symptoms of depression during the past week with 20 items on a 4-point Likert scale (0 = *some of the time* to 3 = *most of the time*). We used the German version of the measure, which has demonstrated good validity and reliability in many studies (Cronbach alpha = 0.89—0.92; Hautzinger & Bailer, 1993; Jahn et al., 2018; Roch et al., 2016; Stein et al., 2014).

We deemed it critical to assess emotion regulation skills of participants using the Emotion Regulation Skills Questionnaire (ERSQ; German version: Berking & Znoj, 2008). The ERSQ is a self-report measure consisting of 27 items that assesses the successful application of the following emotion regulation skills in the past week: *awareness, sensations, clarity, understanding, modification, acceptance, tolerance, readiness to confront distressing situations,* and *compassionate self-support*. Each item is rated on a 5-point Likert-scale (0 = *not at all* to 4 = *almost always*). An overall indicator of general emotion regulation skills is computed as the mean of all items. The ERSQ has demonstrated good validity and reliability consistently in research studies (Cronbach alpha = 0.89—0.97; Berking et al., 2008, 2011, 2013; Berking & Znoj, 2008; Wirtz et al., 2014).

Additional outcomes measured, but not analyzed in the current study included participants' state stress level (self-constructed). Moreover, in the active conditions, we also assessed the current stress level before each training with paper–pencil questionnaires, and recorded performance measures (i.e., ratings of the emotion display in face, body, and voice by the experimenter, as well as a self-rating of the felt intensity of the emotion by the participant) with self-constructed items in the app after each stimulus. In a subsample of participants, the system usability

scale (Brooke, 1996) was administered after the training. In addition, we collected saliva samples and physiological data with a radar system (Will et al., 2018) at pre and between assessments to reinforce the evaluation of the intervention with the help of biological markers of stress (e.g., cortisol, heart rate variability). We intend to publish findings on these measures in forthcoming publications, which will additionally source data from our related research endeavors so as to ensure sufficient statistical power.

Statistical Analyses

Feasibility and acceptability outcomes were analyzed with the help of descriptive statistics. The post-ratings of the understandability of the app-instructions, the helpfulness of the training, and whether participants would recommend the training to a friend were compared by computing effect sizes. Following recommendations for small sample sizes proposed by Lakens (2013), we computed *Hedges g* and (a priori) set $g = 0.2/0.5/0.8$ as small/moderate/large effect (Cohen, 1988). Moreover, we conducted exploratory independent *t*-tests to compare the ratings of the AAMT-AP versus the swipe control group. To explore the efficacy of the intervention on clinical outcomes, we focused on computing effect sizes to gain preliminary insights. We determined that even small effects (according to commonly accepted conventions) on the primary outcome would justify conducting a larger randomized controlled trial. Due to the partial overlap of the assessment periods between pre measures and between measures for some self-report outcomes, the change between pre and post was chosen as central outcome. Here, we also computed *Hedges g* as effect sizes for the difference in change between groups with $g = 0.2/0.5/0.8$ as small/moderate/large effect (Cohen, 1988). For the effect sizes, 95% bootstrap confidence intervals were computed. Additionally, we ran exploratory two-way mixed analyses of variance (ANOVAs). In these analyses, we included treatment as between-subjects factor (AAMT-AP vs. inactive control group or AAMT-AP vs. swipe control group) and time (pre vs. post) as within-subjects factor. The interaction effect between treatment and time reflects differences across treatment conditions in terms of changes on the dependent variable from pre- to post-treatment. To maximize power, we first compared the AAMT-AP condition to the inactive control condition. Only in the case of a significant interaction, did we also compare the AAMT-AP condition to the swipe control condition. This procedure was based on the premise that not receiving any treatment at all will in all cases result in smaller effects than receiving the swipe-based intervention (which has been shown to be superior to inactive control conditions in the literature, e.g., Kollei et al., 2017; Lukas & Berking, 2018; Lukas et al., 2019). We conducted one-sided tests for the comparison between AAMT-AP and inactive control condition since we expected a greater decrease in perceived stress, dysfunctional attitudes and symptoms of depression and a greater increase in emotion regulation skills in the AAMT-AP group. For the comparison between AAMT-AP and swipe control condition, however, we could not make clear predictions concerning the direction of the differences due to the novelty of the AAMT-AP intervention. Therefore, we decided to conduct two-sided tests. Group differences

in sociodemographic variables and baseline scores of the primary and secondary outcomes were analyzed with one-way ANOVA for continuous variables and Fisher's exact test for categorical data. We used per-protocol analyses as our primary approach as we were interested in potential effects occurring in individuals really utilizing the intervention. To explore the potential of average efficacy in individuals who had all been provided with the opportunity to use the intervention (regardless of whether they actually chose to do so), we also performed ITT analyses, including all participants who met the inclusion criteria and were allocated to a study condition. In the ITT analyses, we used multiple imputation with $m = 20$ imputations. The modelling was based on the sum scores of all available questionnaire data. All analyses were conducted with R Studio Version 4.1.2 (R Core Team, 2017). For computing effect sizes, we used the rstatix-package (Kassambra, 2021) and for the ANOVA the ez-package (Lawrence, 2016).

Results

Feasibility and Acceptability

The sample for feasibility and acceptability analyses consisted of $n = 40$ individuals (i.e., 15 participants in the AAMT-AP, 13 in the swipe control condition, and 12 in the inactive control condition).

Feasibility and Acceptability of the Intervention

With regards to technical feasibility, it is of note that connectivity problems between smartphone and experimenter tablet occurred in 2.94% of the AAMT-AP sessions and in 3.85% of sessions in the swipe control condition, which each led to a delay of the training of about 20 min. With regards to adherence, it is of note that in the AAMT-AP condition, one participant had to be excluded from the study because of difficulties in understanding and following the instructions, despite multiple attempts at explanation. A second participant dropped out of the study during the intervention, reporting to be too stressed out to properly comply with the intervention protocol. Thus, the adherence to the intervention was 93.75% in the AAMT-AP condition. Adherence to the intervention in the swipe control condition was 100%. Regarding usability, the ratings of the understandability of the instructions in the AAMT-AP condition ranged from 1 to 5 with a mean of 4.33 ($SD = 1.05$). In the swipe control condition, the average understandability rating was 5 (SD = 0). The effect size was large with $g = -0.88$ and the exploratory t-test (Welch) showed that this difference was significant ($t(14) = -2.47$, $p = 0.03$). With regards to the ease of the required display of emotions, the mean rating of displaying anger was 2.87 ($SD = 1.51$, range = 0–5) at the beginning of the training and 4.73 ($SD = 0.46$, range = 4–5) at the end of the training, with a mean change of 1.87 ($SD = 1.55$). For the display of positive emotions, the mean rating was 3.54 ($SD = 0.59$, range = 1–5) at the beginning of the training and 3.91 ($SD = 0.71$, range = 1–5) at the end of the training, with a mean change of 0.37 ($SD = 0.37$). With regards to acceptability, the mean rating of

the perceived helpfulness of the training in invalidating stress-related beliefs in the AAMT-AP was 3.4 ($SD = 0.74$, range = 3–5) and 3.15 ($SD = 1.72$, range = 0–5) in the swipe control condition. The effect size for this difference was negligible ($g = 0.18$) and Welch's t-test did not show significant differences in the ratings between groups ($t(15.76) = 0.48$, $p = 0.64$). The question of whether participants would recommend the training to a friend received a mean score of 3.27 ($SD = 1.33$, range = 1–5) in the AAMT-AP condition and a mean score of 3.23 ($SD = 1.88$, range = 0–5) in the swipe control condition. Again, the effect size was negligible with $g = 0.02$ and the exploratory independent t-test was not significant ($t(26) = 0.06$, $p = 0.95$). Replying to the open-ended question, one participant in the AAMT-AP described the display and experience of different emotions during the course of the training to be exhausting. In the swipe control condition, three participants felt the training was monotonous. With regards to adverse events, no decompensation of study participants was reported by the experimenters. Change in the experience of anger from T1 to T3 was $M = -0.17$ ($SD = 1.27$) in the AAMT-AP, $M = 0$ ($SD = 0.74$) in the swipe control, and $M = 0.08$ ($SD = 1.31$) in the inactive control condition. The effect size for the between group comparisons of AAMT-AP and swipe control condition ($g = 0.16$ [-0.57; 1.39]) as well as AAMT-AP and inactive control condition ($g = 0.19$ [-0.58; 1.26;]) were negligible.

Feasibility of the Study Design

The flow of participants through the study is displayed in Fig. 1. Recruitment took place between July 2020 and May 2022. During this period, 441 individuals were screened for eligibility. Study criteria were met by 269 individuals, of whom 111 (41.26%) participated in the study. The large majority (88.1%) of this final sample completed all assessments (70.59% in the AAMT-AP condition, 100% in the swipe control condition, and 92.31% in the inactive control group). Participants who did not complete the assessments stopped responding to emails, therefore we were not able to explore the reasons.

Clinical Outcomes

Preliminary Analyses

The groups did not differ in sociodemographic characteristics and baseline measures of study outcomes (see Table 1).

Per-protocol Analyses

The per-protocol sample consisted of $n = 30$ individuals (i.e., $n = 10$ per group). Descriptive statistics of the primary and secondary outcomes are displayed in Table 2 and effect sizes in Table 3. With regards to perceived stress as the primary outcome, the effect size for the change in PSS-10 from pre to post in the AAMT-AP compared to the inactive control condition was $g = 0.77$ [0.05; 1.93;]. This

Table 2 Means and Standard Deviations of Primary and Secondary Outcomes (Per-Protocol Sample)

	AAMT-AP ($n = 10$)				Swipe control ($n = 10$)				Inactive control ($n = 10$)			
	Pre		Post		Pre		Post		Pre		Post	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
PSS-10	20	5.68	13.5	3.89	18.7	7.53	16.2	7.04	19.6	5.68	18.3	6.46
DAS-A	28.3	10.17	16	4.37	28.1	12.23	15.6	7.62	33	9.07	27	10.3
DAS-H	19.3	6.83	13.4	3.78	16.9	5.92	12.7	4.24	21.6	9.25	22	10.4
CES-D	19.1	5.30	13.2	4.96	15.6	8.37	11	8.19	21	9.96	18.9	12.0
ERSQ	2.26	0.56	2.73	0.54	2.57	0.59	2.96	0.51	2.18	0.43	2.49	0.41

Note. AAMT-AP = Approach-Avoidance Modification Training with anger and positive emotions; PSS-10 = Perceived Stress Scale-10; DAS-A = Dysfunctional Attitude Scale-Approval; DAS-H = Dysfunctional Attitude Scale-Helpfulness; CES-D = Center for Epidemiological Studies Depression Scale; ERSQ = Emotion Regulation Skills Questionnaire

Table 3 Effect Sizes for Between Group Differences in the Change in Outcomes from Pre to Post

Measure	Heges *g* (CI)	
	Inactive Control	Swipe control
PSS-10		
AAMT-AP	0.773 [-0.05; 1.93]	0.718 [-0.12; 1.74]
Swipe control	0.182 [-0.58; 1.34]	
DAS-A		
AAMT-AP	0.685 [-0.08; 1.62]	-0.02 [-1.03; 0.88]
Swipe control	0.762 [-0.06; 1.79]	
DAS-H		
AAMT-AP	1.04 [0.25; 2.08]	0.332 [-0.63; 1.27]
Swipe control	0.859 [0.05; 1.96]	
CES-D		
AAMT-AP	0.391 [-0.51; 1.34]	0.250 [-0.74; 1.13]
Swipe control	0.266 [-0.82; 1.15]	
ERSQ		
AAMT-AP	0.312 [-0.5;1.56]	0.177 [-0.81;1.04]
Swipe control	0.143 [-0.76;1.46]	

Note. PSS-10 = Perceived Stress Scale-10; AAMT-AP = Approach-Avoidance Modification Training with anger and positive emotions; DAS-A = Dysfunctional Attitude Scale-Approval; DAS-H = Dysfunctional Attitude Scale-Helpfulness; CES-D = Center for Epidemiological Studies Depression Scale; ERSQ = Emotion Regulation Skills Questionnaire

corresponds to a medium to large effect size. The effect size in the comparison of AAMT-AP with the swipe control condition was $g = 0.72$ [-0.12; 1.74], which also corresponds to a medium to large effect. Figure 3 illustrates the results for the change in the PSS-10. The exploratory ANOVA showed significant differences in the reduction of perceived stress between the AAMT-AP and the inactive control

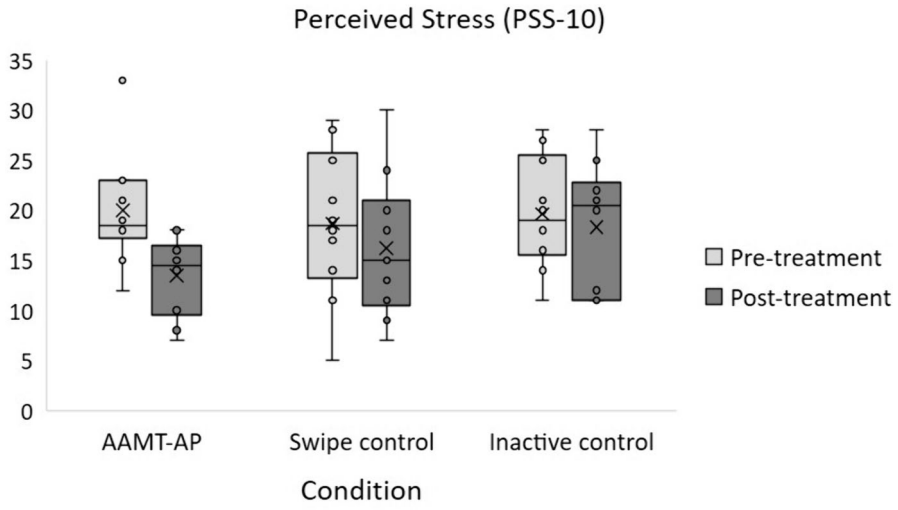


Fig. 3 Boxplot for the PSS-10 at Pre- and Post-treatment in the Three Study Conditions [Note. PSS-10 = Perceived Stress Scale-10; AAMT-AP = Approach-Avoidance Modification Training with anger and positive emotions]

condition (interaction effect: $F(1,18) = 3.26$, $p = 0.04$). The exploratory ANOVA for the comparison of the AAMT-AP and the swipe control group, however, was not significant ($F(1,18) = 2.81$, $p = 0.11$).

With regards to secondary outcomes, for the change in *approval of dysfunctional attitudes* assessed with the DAS-A, effect sizes were moderate to large in the AAMT-AP ($g = 0.69$ [-0.08;1.62]) and swipe control conditions ($g = 0.76$ [-0.06;1.79]) compared to the inactive control condition, respectively. Concerning the change in the perceived *helpfulness of dysfunctional attitudes* assessed with the DAS-H, effect sizes were large in the AAMT-AP ($g = 1.04$ [0.25;2.08]) and the swipe control condition ($g = 0.86$ [0.05;1.96]) compared to the inactive control condition, respectively. Effect sizes for the reduction of *symptoms of depression* were small in the AAMT-AP ($g = 0.39$ [-0.51;1.34]) and swipe control condition ($g = 0.27$ [-0.82;1.15]) compared to the inactive control condition, respectively. With regards to changes in *emotion regulation skills*, there was an increase in the ERSQ total score with small effects in the AAMT-AP compared to the inactive control condition ($g = 0.31$ [-0.5;1.56]). The effect size in the swipe control compared to the inactive control condition, however, was negligible ($g = 0.14$ [-0.76;1.46]).

The requirements for conducting the exploratory ANOVA were violated for the DAS-A (normality), DAS-H (normality and homogeneity of variance), and CES-D (homogeneity of variance). Therefore, following recommendations by Field and Wilcox (2017) we conducted a robust mixed ANOVA with the WRS2-package (Mair & Wilcox, 2020) for those outcomes. None of the exploratory ANOVAs showed significant differences in change between the AAMT-AP and the inactive control condition (DAS-A: $F(2,9.23) = 1.51$, $p = 0.14$; DAS-H: $F(2,8.85) = 0.99$, $p = 0.20$; CES-D: $F(2,8.86) = 0.84$, $p = 0.23$; ERSQ: $F(1,18) = 0.53$, $p = 0.24$).

ITT Analyses

The ITT sample consisted of $n = 36$ individuals (i.e., 15 participants in the AAMT-AP, 10 in the swipe control condition, and 11 in the inactive control condition). The ITT analyses yielded similar results as the exploratory per protocol analyses. With regard to the PSS-10, the interaction between group and time was significant for the comparison of the AAMT-AP and the inactive control condition ($F(1,1399.04) = 2.61$, $p = 0.05$, $\eta^2 = 0.05$) whereas it was not significant for the comparison of the AAMT-AP and the swipe control condition ($F(1,12704.62) = 1.61$, $p = 0.10$, $\eta^2 = 0.03$). For the other outcomes, the interaction between group (AAMT-AP vs. inactive control) and time (T1 vs. T3) was not significant (DAS-A: $F(1,1863.01) = 1.23$, $p = 0.13$, $\eta^2 = 0.02$; DAS-H: $F(1,1109.74) = 1.54$, $p = 0.11$, $\eta^2 = 0.04$; CES-D: $F(1,1295.94) = 0.62$, $p = 0.22$, $\eta^2 = 0.02$; ERSQ: $F(1,1579.05) = 0.30$, $p = 0.29$, $\eta^2 = 0.01$).

Discussion

This randomized controlled pilot study aimed to evaluate the feasibility and acceptability and explore the clinical potential of a 4-day, smartphone-based AAMT intervention that utilizes anger and positive emotions to move AAMT-stimuli (i.e., stress-related [dys-]functional beliefs) away from/toward oneself.

With regards to the feasibility of the intervention, the occurrence of technical problems remained below the benchmark of 95% for both the AAMT-AP and the swipe control condition. Moreover, the app instructions were rated as understandable, with a slight and significant superiority of the swipe control condition, which might be explained by the more complex task in the AAMT-AP condition. The moderate rating of the ease of emotion display in the AAMT-AP condition does suggest that an emotion-based AAMT is feasible. The intervention adherence was comparatively high in both the AAMT-AP condition (93.75%) and the swipe control condition (100%). Future studies should estimate more reliably with larger samples the percentage of individuals who could be too stressed to engage in the AAMT-AP safely and effectively. Concerning acceptability, the mean ratings of the training's helpfulness and its recommendation to a friend was high in the AAMT-AP and the swipe control conditions, and did not differ between the two. This finding suggests that the AAMT-AP intervention was well accepted by participants. The open evaluation of the training showed that the swipe control condition was experienced to be monotonous by some, which has also been found in a previous study investigating a similar intervention (Beard et al., 2012). This finding underlines the advantages of an engaging intervention with more variation like AAMT-AP. One participant, however, described the repeated use of emotions during AAMT-AP to be exhausting. Therefore, future studies should investigate the effects of training periods with different lengths in order to identify a minimum effective dose. It is also thinkable that the intervention duration may be individually increased or decreased with increasing time of practice with the app, depending on the individual progress made. With regards to adverse events, our results suggest that AAMT-AP did not cause

decompensation or increase in the experience of anger and can thus be considered safe. Results on the feasibility of the study design were mixed. The proportion of eligible individuals who truly participated in the study was 41.26%. Moreover, considering the total sample size, the recruitment period of 23 months was relatively long. In fact, the study period coincided with the Covid-19 pandemic, which hindered recruitment due to unpredictably changing legal parameters and strict quarantine rules. For example, people were not allowed to participate if they had any symptoms related to Covid-19, which became especially problematic during the flu season. Moreover, we were only able to spread the study information online since face-to-face classes and events were prohibited. Presumably, recruitment rate would be significantly higher in post-Covid times. Adherence to study procedures was very good with 88.1% completed assessments in total.

Concerning the clinical potential of AAMT-AP, it is of note that a moderate-to-large superiority of the AAMT-AP condition over the inactive control condition was found with regards to reductions in perceived stress as the primary clinical outcome. Despite the small sample size, this effect was also significant in the exploratory statistical comparisons. Thus, the findings indicate that AAMT-AP has clinical potential. Moreover, when compared to the swipe control condition, the AAMT-AP was also superior with a moderate to large effect ($g = 0.72$). Furthermore, effect sizes found in this study suggest a superiority to the majority of web-based stress interventions systematically assessed in the meta-analysis of Heber et al., (2017). Finally, the finding that effect sizes on perceived stress were small in the more traditional swipe-AAMT matches the findings for joystick-based AAMTs in similar studies (Becker et al., 2016; Ferrari et al., 2018). Here, only one study found effects on stress reactivity in a subsample of dysphoric individuals (Becker et al., 2016). However, these effects emerged only in a subgroup and only after a sad mood induction before the training, whereas in the study without mood induction there were no effects. Hence, the effects might be attributable to the sad mood state participants were subjected to. Thus, the majority of findings on previous joystick-based AAMTs indicates a lack of efficacy against ameliorating stress. These findings are consistent with the hypothesis that integrating emotional responses into the AAMT-procedure enhances their efficacy. When explained within the theoretical framework of the inferential account (Van Dessel et al., 2019), the greater valence of the emotions deliberately displayed as compared to swipe- or joystick-movements performed might account for these observed effects. However, emotions are complex constructs, and it is alternatively or additionally possible that the effects might be caused by adding other components (e.g., associated appraisals or behavioral tendencies) to the associative network of the respective belief.

In line with the assumption that AAMT-AP is superior to previous joystick-based AAMTs, an alternative explanation might be the stimuli used. We, in contrast to those studies, used stimuli implicated in the stress process (i.e., stress-related beliefs with matching emotional pictures). Hence, our AAMT might target stress more directly and therefore might be more effective in reducing it. Moreover, in the aforementioned studies increased stress level was not an inclusion criterion, which increases the likelihood of floor effects. However, our results suggest that the swipe-version of the AAMT also is not effective although the present study only included

individuals with elevated stress burden. Thus, it is plausible that the larger effect size found in the present study is mainly attributable to the participants' deliberate manifestation of emotions in vocalization, facial expression, and gesture.

With regards to the secondary outcomes, effect sizes found in our study indicate that AAMT-AP can reduce the approval of and perceived helpfulness of dysfunctional beliefs with moderate-to-large effect sizes. Interestingly, effect sizes were similar for the AAMT-AP and the swipe control condition. This is an important finding considering that dysfunctional beliefs are a main contributing factor to the emergence of stress (Kuroda, 2016) and the development of psychopathology in general (Beck & Haigh, 2014). As already mentioned above, one main difference between our study and other AAMT interventions in the context of stress and related conditions such as depression is the use of (dys-)functional beliefs as stimuli. So far, most studies used stimulus-driven approaches with emotional pictures (e.g., Becker et al., 2016, 2019; Ferrari et al., 2018; Sweet et al., 2021; Vrijsen et al., 2018). To our knowledge, only two studies used additional verbal statements as stimulus material (Lukas et al., 2021 in depression and Lukas & Berking, 2018 in procrastination), also with promising effects on symptomatology. However, the effect on the appraisal of (dys-)functional beliefs has never been investigated. Our results suggest the potential of AAMT interventions to change the appraisal of (dys-)functional beliefs and open the path to a new way of designing and implementing AAMT interventions.

Considering that effect sizes of reductions in perceived stress, but not in dysfunctional beliefs differed between the intervention groups, the underlying mechanism of stress reduction in our training is not sufficiently explained by a mere change in the appraisal of (dys-) functional beliefs. Instead, the findings suggest a specific effect resulting from participants' deliberate display of emotions. By mobilizing psychological and physiological resources (Berking & Whitley, 2014; Novaco, 2016), anger might, beyond the mere modification of stress-related beliefs, facilitate active coping with stress and help participants regain a sense of control over stressful situations (Lerner et al., 2015). This might account for the bigger effects of the AAMT-AP compared to the swipe control condition for perceived stress. The exact mechanisms of change, however, need to be investigated with the help of systematic mediation analyses in larger samples.

In terms of symptoms of depression, it is not surprising that effect sizes were small, since depression was not the main target of our intervention. Moreover, floor effects might have played a role, as we did not include a clinical sample. However, the fact that effects were found, although small, is promising and aligned with previous studies showing beneficial effects of AAMT-interventions on symptoms of depression (Becker et al., 2019; Lukas et al., 2021; Vrijsen et al., 2018). Thus, emotion-based AAMTs should be developed in the context of depression, and future studies should investigate their feasibility and efficacy in clinical samples with a confirmed diagnosis of depression.

When engaging in AAMT-AP, participants train to deliberately display different emotions, and so they are consequently confronted with the characteristic features of those emotions as they enact them. Hence, the AAMT-AP might be considered a kind of implicit emotion regulation training. However, only small effects emerged

for emotion regulation, but, notably, only in the AAMT-AP condition, corroborating its superiority to the less emotionally engaging swipe control condition. Of note, training in the AAMT-AP condition focused on the display of anger and therefore might have selectively enhanced anger regulation skills, which might not generalize to other emotions. Since the ERSQ measures general emotion regulation skills, this might explain the small effects. Moreover, we only investigated effects of the intervention on the total ERSQ-score. However, it is possible that by merely displaying specific emotions, AAMT-AP participants train certain strategies related to the recognition of different emotions (e.g., awareness and clarity), whereas other strategies related to dealing with emotional states (e.g., modification, tolerance) might not be affected. This could explain the small effect size, and future studies should investigate AAMT-AP effects on specific emotion regulation strategies.

There are several limitations of the current study to be considered. First, as mentioned above, this is a pilot study with a small sample size that did not have sufficient power for reliable confirmatory analyses. Second, the sample was very homogenous with predominantly young female student participants. Therefore, results cannot be generalized to the general population. Future studies should evaluate the AAMT-AP in a larger sample of the general population of individuals suffering from elevated stress, including middle-aged and older adults. Third, participants in the inactive control condition only completed online questionnaires at home. Hence, it was not possible to control for unspecific effects of participating in an intervention like those resulting from interacting with the research team. To control for such effects, future studies should include placebo control conditions, which require participants to come to the lab and interact with the research team. As such, in the present study we did not control for effects of the mere confrontation with stress-related beliefs. This confrontation alone could stimulate participants to recognize and question their implicit beliefs. Thus, future studies should include a sham-condition in which participants are presented with the same stimulus material as participants in the AAMT-AP condition. Fourth, our intervention was conducted as a Wizard-of-Oz-paradigm (Green & Wei-Haas, 1985) with an experimenter observing participants' performance—since the smartphone app was not equipped to automatically assess emotional responses. The presence of another person, however, might influence participants' performance and serve as a stressor (Guerin, 1986), thereby reducing the efficacy of the intervention or artificially boost participants' performance (Uziel, 2007), thereby increasing efficacy. Once the implementation of automated assessment of emotional responses is feasible, we will develop an autonomous version of the app. Fifth, we used the PSS as both selection criterion and outcome measure. This harbors the risk that stress reduction effects found are the results of a regression to the mean and not the actual effects of the intervention. Future studies should use different measures for screening and outcome assessment. Finally, we cannot draw conclusions on long-term effects of our interventions, as we only included a one-week post-assessment. Thus, future studies should include long-term follow-up assessments, and might also involve more numerous or extended training sessions.

These limitations aside, results of this pilot study are promising for several reasons: (1) Effective interventions against stress are needed to prevent the development of severe, stress-related physical and mental conditions; (2) When automatic

emotion recognition becomes feasible, the AAMT-AP will operate more autonomously with less possibility for experimenter bias effects. As such, it will hold all advantages of smartphone-based interventions, including high accessibility and availability, cost-effectiveness, and easy dissemination (Ben-Zeev et al., 2015; Ly et al., 2012; Marshall et al., 2020); (3) Dysfunctional beliefs are implicated in the development and maintenance of psychopathology (Beck & Haigh, 2014). If AAMT interventions are confirmed to successfully modify such beliefs, they could be used as an adjunct to psychotherapy for a broad range of clinical applications. Further steps that need to be taken before the dissemination of the intervention include determining the optimal setting and number of training sessions (e.g., Eberl et al., 2014), as well as suited target groups. To overcome problems of credibility and lack of motivation due to repetitiveness reported in similar trainings (Beard et al., 2012), a comprehensible and convincing treatment rationale should be provided. Moreover, gamification elements could be integrated to increase participants' engagement.

In sum, findings of the present study are consistent with the assumed potential of AAMT-AP to provide accessible, easy-to-disseminate interventions that foster successful coping with elevated stress and, potentially, other clinically relevant conditions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s41811-024-00218-z.

Acknowledgements We thank all participants taking part in our study and the undergraduate students for supporting the conductance of the study. Moreover, we are grateful to Dr. Stefan Gradl for developing and maintaining the AAMT-App. We thank Giulia A. Chillemi for providing editorial support and Hannah Streit and Lydia Rupp for proofreading the manuscript.

Author Contributions **M. Keinert:** conceptualization, data curation, formal analysis, investigation, methodology, writing – original draft. **L. Schindler-Gmelch:** writing – review & editing. **B. M. Eskofier:** software, writing – review & editing. **M. Berking:** conceptualization, formal analysis, funding acquisition, methodology, project administration, resources, supervision, writing – review & editing.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Bavarian Ministry of Science and Arts [as part of the Bavarian Research Association on Healthy Use of Digital Technologies and Media (ForDigitHealth)] and the German Research Foundation [SFB 1483 – Project-ID 442419336, EmpkinS]. The funding sources were not involved in the study design, the collection, analysis, and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

Data Availability The data set and R code is available in the open science framework (osf.io/cmqd5).

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

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

Allen, A. P., Curran, E. A., Duggan, Á., Cryan, J. F., Chorcoráin, A. N., Dinan, T. G., Molloy, D. W., Kearney, P. M., & Clarke, G. (2017). A systematic review of the psychobiological burden of informal caregiving for patients with dementia: Focus on cognitive and biological markers of chronic stress. *Neuroscience & Biobehavioral Reviews, 73*, 123–164. https://doi.org/10.1016/j.neubiorev.2016.12.006

American Psychological Association. (2020). *Stress in America™ 2020: A national mental health crisis*. American Psychological Association. https://www.apa.org/news/press/releases/stress/2020/sia-mental-health-crisis.pdf

American Psychological Association. (2021). Stress in America™: One year later, a new wave of pandemic health concerns. *American Psychological Association*. https://doi.org/10.1037/e502832021-001

Asnaani, A., Rinck, M., Becker, E., & Hofmann, S. (2014). The effects of approach-avoidance modification on social anxiety disorder: A pilot study. *Cognitive Therapy & Research, 38*(2), 226–238. https://doi.org/10.1007/s10608-013-9580-x

Beard, C., Weisberg, R. B., & Primack, J. (2012). Socially anxious primary care patients' attitudes toward Cognitive Bias Modification (CBM): A qualitative study. *Behavioural and Cognitive Psychotherapy, 40*(5), 618–633. https://doi.org/10.1017/S1352465811000671

Beck, A. T., & Haigh, E. A. P. (2014). Advances in cognitive theory and therapy: The Generic Cognitive Model. *Annual Review of Clinical Psychology, 10*(1), 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734

Becker, E. S., Barth, A., Smits, J. A. J., Beisel, S., Lindenmeyer, J., & Rinck, M. (2019). Positivity-approach training for depressive symptoms: A randomized controlled trial. *Journal of Affective Disorders, 245*, 297–304. https://doi.org/10.1016/j.jad.2018.11.042

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*, 57–71. https://doi.org/10.1007/s10608-015-9716-2

Bell, M. L., Whitehead, A. L., & Julious, S. A. (2018). Guidance for using pilot studies to inform the design of intervention trials with continuous outcomes. *Clinical Epidemiology, 10*, 153–157. https://doi.org/10.2147/CLEP.S146397

Ben-Zeev, D., Schueller, S. M., Begale, M., Duffecy, J., Kane, J. M., & Mohr, D. C. (2015). Strategies for mHealth Research: Lessons from 3 Mobile Intervention Studies. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(2), 157–167. https://doi.org/10.1007/s10488-014-0556-2

Berking, M., Ebert, D., Cuijpers, P., & Hofmann, S. G. (2013). Emotion regulation skills training enhances the efficacy of inpatient cognitive behavioral therapy for major depressive disorder: A randomized controlled trial. *Psychotherapy and Psychosomatics, 82*(4), 234–245. https://doi.org/10.1159/000348448

Berking, M., Margraf, M., Ebert, D., Wupperman, P., Hofmann, S. G., & Junghanns, K. (2011). Deficits in emotion-regulation skills predict alcohol use during and after cognitive behavioral therapy for alcohol dependence. *Journal of Consulting and Clinical Psychology, 79*(3), 307–318. https://doi.org/10.1037/a0023421

Berking, M., Orth, U., Wupperman, P., Meier, L. L., & Caspar, F. (2008). Prospective effects of emotion-regulation skills on emotional adjustment. *Journal of Counseling Psychology, 55*(4), 485–494. https://doi.org/10.1037/a0013589

Berking, M., & Whitley, B. (2014). *Affect regulation training* (1st ed.). Springer New York. https://doi.org/10.1007/978-1-4939-1022-9

Berking, M., & Znoj, H. (2008). Entwicklung und Validierung eines Fragebogens zur standardisierten Selbsteinschätzung emotionaler Kompetenzen (SEK-27). *Zeitschrift Fur Psychiatrie Psychologie Und Psychotherapie, 56*(2), 141–153. https://doi.org/10.1024/1661-4747.56.2.141

Beutel, M. E., Klein, E. M., Aufenanger, S., Brähler, E., Dreier, M., Müller, K. W., Quiring, O., Reinecke, L., Schmutzer, G., Stark, B., & Wölfling, K. (2016). Procrastination, distress and life satisfaction across the age range – A German representative community study. *PLoS ONE, 11*(2), e0148054. https://doi.org/10.1371/journal.pone.0148054

Brockmeyer, T., Friederich, H.-C., Küppers, C., Chowdhury, S., Harms, L., Simmonds, J., Gordon, G., Potterton, R., & Schmidt, U. (2019). Approach bias modification training in bulimia nervosa and binge-eating disorder: A pilot randomized controlled trial. *The International Journal of Eating Disorders, 52*(5), 520–529. https://doi.org/10.1002/eat.23024

Brockmeyer, T., Hahn, C., Reetz, C., Schmidt, U., & Friederich, H.-C. (2015). Approach bias modification in food craving—A proof-of-concept study. *European Eating Disorders Review, 23*(5), 352–360. https://doi.org/10.1002/erv.2382

Brooke, J. (1996). SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability Evaluation in Industry* (pp. 189–194). Taylor & Francis Ltd.

Cervera-Torres, S., Ruiz Fernández, S., Lachmair, M., Riekert, M., & Gerjets, P. (2021). Altering emotions near the hand: Approach–avoidance swipe interactions modulate the perceived valence of emotional pictures. *Emotion, 21*(1), 220–225. https://doi.org/10.1037/emo0000651

Clark, D. A. (2013). Cognitive restructuring. In S. G. Hofmann (Ed.), *The Wiley handbook of cognitive behavioral therapy* (Vol. 1, pp. 23–44). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118528563.wbcbt02

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Routledge. https://doi.org/10.4324/9780203771587

Cohen, S., Janicki-Deverts, D., & Miller, G. E. (2007). Psychological stress and disease. *JAMA, 298*(14), 1685–1687. https://doi.org/10.1001/jama.298.14.1685

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: Do clinical effects replicate and for whom does it work best? *Developmental Cognitive Neuroscience, 4*, 38–51. https://doi.org/10.1016/j.dcn.2012.11.002

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2014). Implementation of approach bias re-training in alcoholism—How many sessions are needed? *Alcoholism: Clinical and Experimental Research*, *38*(2), 587–594. https://doi.org/10.1111/acer.12281

Elliot, A. J., Eder, A. B., & Harmon-Jones, E. (2013). Approach-Avoidance Motivation and Emotion: Convergence and Divergence. *Emotion Review, 5*(3), 308–311. https://doi.org/10.1177/1754073913477517

Ferrari, G. R. A., Möbius, M., Becker, E. S., Spijker, J., & Rinck, M. (2018). Working mechanisms of a general positivity approach-avoidance training: Effects on action tendencies as well as on subjective and physiological stress responses. *Journal of Behavior Therapy and Experimental Psychiatry, 59*, 134–141. https://doi.org/10.1016/j.jbtep.2018.01.005

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy, 98*, 19–38. https://doi.org/10.1016/j.brat.2017.05.013

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Girotti, M., Adler, S. M., Bulin, S. E., Fucich, E. A., Paredes, D., & Morilak, D. A. (2018). Prefrontal cortex executive processes affected by stress in health and disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 85*, 161–179. https://doi.org/10.1016/j.pnpbp.2017.07.004

Green, P., & Wei-Haas, L. (1985). The rapid development of user interfaces: Experience with the Wizard of OZ Method. *Proceedings of the Human Factors Society Annual Meeting, 29*(5), 470–474. https://doi.org/10.1177/154193128502900515

Guerin, B. (1986). Mere presence effects in humans: A review. *Journal of Experimental Social Psychology, 22*(1), 38–77. https://doi.org/10.1016/0022-1031(86)90040-5

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1*, 293–319. https://doi.org/10.1146/annurev.clinpsy.1.102803.143938

Hassard, J., Teoh, K. R. H., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of Occupational Health Psychology, 23*(1), 1–17. https://doi.org/10.1037/ocp0000069

Hautzinger, M., & Bailer, M. (1993). *ADS. Allgemeine Depressions Skala* [General Depression Scale]. Beltz.

Hautzinger, M., Bailer, M., Hofmeister, D., & Keller, F. (2012). *Allgemeine Depressionsskala—2., überarbeitete und neu normierte Auflage* [General Depression Scale - 2nd revised and newly standardized edition]. Hogrefe.

Heber, E., Ebert, D. D., Lehr, D., Cuijpers, P., Berking, M., Nobis, S., & Riper, H. (2017). The Benefit of Web- and Computer-Based Interventions for Stress: A Systematic Review and Meta-Analysis. *Journal of Medical Internet Research, 19*(2), e5774. https://doi.org/10.2196/jmir.5774

Jahn, R., Baumgartner, J. S., van den Nest, M., Friedrich, F., Alexandrowicz, R. W., & Wancata, J. (2018). Kriteriumsvalidität der deutschsprachigen Version der CES-D in der Allgemeinbevölkerung. *Psychiatrische Praxis, 45*(8), 434–442. https://doi.org/10.1055/a-0584-9803

Kassambara, A. (2021). *rstatix: Pipe-friendly framework for basic statistical tests* (R package version 0.7.0) [Computer software]. CRAN. https://CRAN.R-project.org/package=rstatix

Keinert, M., Eskofier, B. M., Schuller, B. W., Böhme, S., & Berking, M. (2023). Evaluating the feasibility and exploring the efficacy of an emotion-based approach-avoidance modification training (eAAMT) in the context of perceived stress in an adult sample—Protocol of a parallel randomized controlled pilot study. *Pilot and Feasibility Studies, 9*(1), 155. https://doi.org/10.1186/s40814-023-01386-z

Klein, E. M., Brähler, E., Dreier, M., Reinecke, L., Müller, K. W., Schmutzer, G., Wölfling, K., & Beutel, M. E. (2016). The German version of the Perceived Stress Scale – psychometric characteristics in a representative German community sample. *BMC Psychiatry, 16*, 159. https://doi.org/10.1186/s12888-016-0875-9

Kollei, I., Lukas, C. A., Loeber, S., & Berking, M. (2017). An app-based blended intervention to reduce body dissatisfaction: A randomized controlled pilot study. *Journal of Consulting and Clinical Psychology, 85*(11), 1104–1108. https://doi.org/10.1037/ccp0000246

Krohne, H., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). *Diagnostica, 42*(2), 139–156. https://doi.org/10.1037/t49650-000

Kuroda, Y. (2016). Dysfunctional attitudes lead to depressive symptoms by generating subjective stress. *The Journal of Psychology, 150*(3), 358–370. https://doi.org/10.1080/00223980.2015.1054335

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lawrence, M. A. (2016). *ez: Easy analysis and visualization of factorial experiments* (R package version 4.4–0) [Computer software]. CRAN. https://CRAN.R-project.org/package=ez

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Publishing Company.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology, 66*(1), 799–823. https://doi.org/10.1146/annurev-psych-010213-115043

Loijen, A., Vrijsen, J. N., Egger, J. I. M., Becker, E. S., & Rinck, M. (2020). Biased approach-avoidance tendencies in psychopathology: A systematic review of their assessment and modification. *Clinical Psychology Review, 77*, 101825. https://doi.org/10.1016/j.cpr.2020.101825

Lukas, C. A., & Berking, M. (2018). Reducing procrastination using a smartphone-based treatment program: A randomized controlled pilot study. *Internet Interventions, 12*, 83–90. https://doi.org/10.1016/j.invent.2017.07.002

Lukas, C. A., Eskofier, B., & Berking, M. (2021). A gamified smartphone-based intervention for depression: Randomized controlled pilot trial. *JMIR Mental Health, 8*(7), e16643. https://doi.org/10.2196/16643

Lukas, C. A., Trevisi Fuentes, H., & Berking, M. (2019). Smartphone-based emotion recognition skills training for alexithymia—A randomized controlled pilot study. *Internet Interventions, 17*, 100250. https://doi.org/10.1016/j.invent.2019.100250

Ly, K. H., Dahl, J., Carlbring, P., & Andersson, G. (2012). Development and initial evaluation of a smartphone application based on acceptance and commitment therapy. *Springerplus, 1*(1), 11. https://doi.org/10.1186/2193-1801-1-11

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods, 52*(2), 464–488. https://doi.org/10.3758/s13428-019-01246-w

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Smartphone psychology: New approaches towards safe and efficacious mobile mental health apps. *Professional Psychology: Research and Practice, 51*(3), 214–222. https://doi.org/10.1037/pro0000278

Novaco, R. (2010). Anger and Psychopathology. In M. Potegal, G. Stemmler, & C. Spielberger (Eds.), *International handbook of anger. Constituent and concomitant biological, psychological, and social processes* (pp. 465–497). Springer. https://doi.org/10.1007/978-0-387-89676-2_27

Novaco, R. (2016). Anger. In G. Fink (Ed.), *Stress: Concepts, cognition, emotion, and behavior: Handbook of stress series volume 1* (pp. 285–292). Elsevier. https://doi.org/10.1016/B978-0-12-800951-2.00035-2

Novaco, R. W. (1979). The cognitive regulation of anger and stress. In P. C. Kendall & S. D. Hollon (Eds.), *Cognitive-behavioral interventions: Theory, Research, and Procedures* (pp. 241–285). Academic Press.

O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and health: A review of psychobiological processes. *Annual Review of Psychology, 72*, 663–688. https://doi.org/10.1146/annurev-psych-062520-122331

Otto, M. W., Fava, M., Penava, S. J., Bless, E., Muller, R. T., & Rosenbaum, J. F. (1997). Life event, mood, and cognitive predictors of perceived stress before and after treatment for major depression. *Cognitive Therapy and Research, 21*(4), 409–420. https://doi.org/10.1023/A:1021980209878

R Core Team (2017). *R: A language and environment for statistical computing* (Version 4.1.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Reis, D., Lehr, D., Heber, E., & Ebert, D. D. (2019). The German version of the Perceived Stress Scale (PSS-10): Evaluation of dimensionality, validity, and measurement invariance with exploratory and confirmatory bifactor modeling. *Assessment, 26*(7), 1246–1259. https://doi.org/10.1177/1073191117715731

Remington, N. A., Fabrigar, L. R., & Visser, P. S. (2000). Re-examining the Circumplex model of affect. *Journal of Personality and Social Psychology*, *79*(2), 286–300. https://doi.org/10.1037//0022-3514.79.2.286

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 105–120. https://doi.org/10.1016/j.jbtep.2006.10.001

Rinck, M., Wiers, R. W., Becker, E. S., & Lindenmeyer, J. (2018). Relapse prevention in abstinent alcoholics by cognitive bias modification: Clinical effects of combining approach bias modification and attention bias modification. *Journal of Consulting and Clinical Psychology, 86*(12), 1005–1016. https://doi.org/10.1037/ccp0000321

Roch, S., Fydrich, T., Küch, D., Meyer, J., Rabe, K., Besch, D., Worringen, U., & Hampel, P. (2016). Erfassung von Depressivität und Ängstlichkeit in der stationären verhaltensmedizinisch orthopädischen Rehabilitation – eine Fragebogenvalidierung mithilfe des SKID. *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin, 26*(3), 130–136. https://doi.org/10.1055/s-0042-104066

Rojas, R., Geissner, E., & Hautzinger, M. (2015). DAS-18 Form A und Form B. Entwicklung und psychometrische Überprüfung von zwei vergleichbaren Kurzversionen der Skala Dysfunktionaler Einstellungen (DAS). *Diagnostica*, *61*(4), 173–183. https://doi.org/10.1026/0012-1924/a000123

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48*(4), 813–838. https://doi.org/10.1037/0022-3514.48.4.813

Stein, J., Luppa, M., Mahnke, J., Weyerer, S., Schomerus, G., & Riedel-Heller, S. G. (2014). Screening for depression by telephone using the German version of the Center for Epidemiological Studies Depression Scale (CES-D). *Psychiatrische Praxis, 41*(3), 135–141. https://doi.org/10.1055/s-0033-1343176

Sweet, A. M., Pearlstein, S. L., Paulus, M. P., Stein, M. B., & Taylor, C. T. (2021). Computer-delivered behavioural activation and approach-avoidance training in major depression: Proof of concept and initial outcomes. *British Journal of Clinical Psychology, 60*(3), 357–374. https://doi.org/10.1111/bjc.12287

Taylor, C. T., & Amir, N. (2012). Modifying automatic approach action tendencies in individuals with elevated social anxiety symptoms. *Behaviour Research and Therapy, 50*(9), 529–536. https://doi.org/10.1016/j.brat.2012.05.004

UK Health and Safety Executive. (2020). *Work-related stress, anxiety or depression statistics in Great Britain, 2020*. Health and Safety Executive. https://www.tuc.org.uk/sites/default/files/stress_0.pdf

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality, 41*(3), 579–601. https://doi.org/10.1016/j.jrp.2006.06.008

Van Dessel, P., Hughes, S., & De Houwer, J. (2019). How do actions influence attitudes? An inferential account of the impact of action performance on stimulus evaluation. *Personality and Social Psychology Review, 23*(3), 267–284. https://doi.org/10.1177/1088868318795730

Vrijsen, J. N., Fischer, V. S., Müller, B. W., Scherbaum, N., Becker, E. S., Rinck, M., & Tendolkar, I. (2018). Cognitive bias modification as an add-on treatment in clinical depression: Results from a placebo-controlled, single-blinded randomized control trial. *Journal of Affective Disorders, 238*, 342–350. https://doi.org/10.1016/j.jad.2018.06.025

Will, C., Shi, K., Schellenberger, S., Steigleder, T., Michler, F., Fuchs, J., Weigel, R., Ostgathe, C., & Koelpin, A. (2018). Radar-based heart sound detection. *Scientific Reports*, *8*, https://doi.org/10.1038/s41598-018-29984-5

Wirtz, C. M., Hofmann, S. G., Riper, H., & Berking, M. (2014). Emotion regulation predicts anxiety over a five-year Interval: A cross-lagged panel analysis. *Depression and Anxiety, 31*(1), 87–95. https://doi.org/10.1002/da.22198

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