Behaviour Research and Therapy 184 (2025) 104667

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Prosody-focused feedback enhances the efficacy of anti-depressive self-statements in depressed individuals – A randomized controlled trial

Jonathan F. Bauer [a,*], Lena Schindler-Gmelch [a], Maurice Gerczuk [b,c], Björn Schuller [b,c,d], Matthias Berking [a]

[a] Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-Universität Erlangen Nürnberg, Nägelsbachstraße 25a, 91052, Erlangen, Germany
[b] Chair of Health Informatics, MRI, Technical University Munich, Ismaninger Straße 22, 81675 München, Germany
[c] Munich Center for Machine Learning, Arcisstraße 21, 80333 München, Germany
[d] GLAM - Group on Language, Audio, & Music, Imperial College London, London, SW7 2AZ, UK

ARTICLE INFO

Keywords:
Major depressive disorder
Cognitive restructuring
Prosody
Voice
Interacting cognitive subsystems
Embodiment

ABSTRACT

This study was aimed to evaluate whether the efficacy of invoking anti-depressive self-statements to cope with depressed mood can be enhanced for depressed individuals by systematically guiding them to amplify the expression of conviction in their voice. Accordingly, we recruited $N = 144$ participants (48 clinically depressed individuals, 48 sub-clinically depressed individuals, and 48 non-depressed individuals). Participants were randomly assigned to an experimental or control condition. Across study conditions, participants completed a mood induction procedure, then read aloud scripted anti-depressive self-statements designed to reduce depressed mood. Participants in the experimental condition received instructions to heighten the prosodic expression of conviction in their voice; participants in the control condition received no prosodic expression instructions. Results showed that depressed participants achieved a more pronounced decrease of depressed mood in the experimental condition than in the control condition. Further, the results indicated no effects in sub-clinically depressed and non-depressed individuals. Finally, heightened conviction expressed by participants in the experimental condition was associated with lower depressed mood and diminished depressive symptom severity. Overall, our findings suggest that fostering the prosodic expression of conviction in depressed persons' voices, while they vocalize anti-depressive self-statements, represents a promising method for augmenting the efficacy of cognitive interventions for depression.

Major depressive disorder (MDD) is a prominent cause of disability worldwide (James et al., 2018). The debilitating disease is characterized by depressed mood, loss of interest, cognitive impairments, changes in psychomotor function, fatigue, suicidal ideation, feelings of worthlessness, and disturbances in sleep and appetite (APA, 2013). MDD is associated with a reduced quality of life and increased risk of comorbidities and mortality (Kessler & Bromet, 2013). Cognitive behavioral therapy (CBT) has been shown to be an effective treatment of MDD (Cuijpers et al., 2008, 2013). However, various studies have found significant rates of insufficient treatment response or non-response, relapse, and disease recurrence (Beshai et al., 2011; Casacalenda et al., 2002; Härter et al., 2004; Taylor et al., 2010; Vittengl et al., 2007). Thus, further optimization of CBT efficacy is needed.

CBT for depression is deeply rooted in Beck's cognitive model, which posits dysfunctional beliefs to be the principal cause of the disorder (Beck & Haigh, 2014; Cristea et al., 2015). Such dysfunctional beliefs commonly refer to the self (e.g., "I am a total failure"), the world (e.g., "No one likes me"), and the anticipated future (e.g., "My future is hopeless"; Beck & Haigh, 2014), and are experienced as automatic thoughts (e.g., negative self-verbalizations; Hollon & Kendall, 1980). Based on this theory, a primary goal of CBT in the treatment of depression is to identify dysfunctional beliefs and reshape them into adaptive ones (e.g., "Making mistakes does not devalue my achievements") by employing cognitive restructuring techniques (Beck & Haigh, 2014). In addition, more recent approaches stemming from the so-called third wave of CBT emphasize the importance of self-acceptance and self-compassion for functional, salutogenic beliefs and resilience (Ehret et al., 2015; Gilbert & Procter, 2006).

* Corresponding author.
E-mail addresses: jonathan.f.bauer@fau.de (J.F. Bauer), lena.gmelch@fau.de (L. Schindler-Gmelch), maurice.gerczuk@tum.de (M. Gerczuk), schuller@tum.de (B. Schuller), matthias.berking@fau.de (M. Berking).

https://doi.org/10.1016/j.brat.2024.104667
Received 21 December 2023; Received in revised form 5 August 2024; Accepted 6 December 2024
Available online 7 December 2024

Extending Beck's model, the theory of Interacting Cognitive Subsystems (ICS) suggests that interactions between activated cognitive schemas and sensory information play a critical role in maintaining depressed affect (Teasdale & Barnard, 1993). ICS theory proposes several cognitive subsystems, among them, the *acoustic* and *visual subsystem* process auditory and visual perception, the *body state subsystem* processes interoceptive sensations, and the *propositional subsystem* processes semantic concepts and relationships between them that are unrelated to emotion (e.g., factual knowledge). The *implicational subsystem* scans the information provided by all other subsystems and triggers affective responses if affect-specific patterns are detected. The ICS theory suggests that depressed mood is triggered whenever the *implicational subsystem* detects a pattern that characterizes the present situation as aversive, uncontrollable, and stable over time. The response component of such a depressogenic schema includes changes in propositional information processing (i.e., "negative thinking") and changes in body state (e.g., slumped posture, frowning, low energy in muscles needed for voice production). If these psychological and physical changes are perceived by the *implicational subsystem* as evidence that the present situation is indeed aversive, uncontrollable, and stable over time, a positive feedback loop is established, by which symptoms of depression cue the ongoing activation of depressogenic schemas and, thereby, the perpetuation of the symptoms. Thus, the goal of psychotherapeutic treatment according to the ICS theory is to disrupt this maladaptive feedback loop by introducing new information into the system that cues schemas incompatible with depression (e.g., hope, self-compassion, confidence). This can be achieved by using CBT techniques that initially challenge depressogenic beliefs (e.g., reappraisal) and then support the development of more functional, salutogenic beliefs (e.g., positive reorientation; Clark, 2013).

Drawing on ICS theory, we hypothesized that the degree of conviction used in one's voice when uttering anti-depressive self-statements moderates the efficacy of such statements for depressed individuals striving to cope with depressed mood. Our hypothesis assumes that a convincing voice differs from a non-convincing one in terms of the bodily sensations associated with underlying muscle activity. Remarkably, human speech production requires the complex motor control of more than 100 distinct muscles (Simonyan & Horwitz, 2011), the orchestration of which is continually monitored by the body-state sensory system. As such, speech is subject to classical conditioning; specific sensory patterns become inextricably associated with specific affective states (Skinner, 1965). Additionally, the level of conviction in a person's voice is estimated through speech prosody (i.e., rhythmic and intonational aspects of speech), and the perceived conviction likely influences the credibility of the information being conveyed (Goupil et al., 2021; Jiang & Pell, 2017). Thus, it can be assumed that depressed individuals are less likely to benefit from anti-depressive statements if they *hear themselves* vocalizing these statements in an *un*convincing manner (i.e., low intensity, minor pitch variation, slow speech rate, and rising intonation at the end of statements; Goupil et al., 2021; Jiang & Pell, 2017). In sum, from both a somatosensory and auditory perspective, depressed individuals risk invalidating the content and thereby diminishing the impact of their own anti-depressive self-statements because of how they insufficiently intonate them.

Numerous empirical studies substantiate that depression is associated with gait, posture and facial expression patterns (Adolph et al., 2021; Michalak et al., 2009; Pampouchidou et al., 2019), and that experimentally induced changes in those parameters affect both the processing of emotions and higher-order cognitive processes (Michalak et al., 2014, 2015; Price & Harmon-Jones, 2015). Conversely, inducing depression-associated body states (e.g., slumped posture, sluggish gait, sad facial expression, or contraction of the corrugator supercilii muscle to achieve a furrowed brow) has been shown to trigger and perpetuate affective and cognitive processes associated with depression (e.g., negative affective memory bias, biased attention toward and accelerated processing of negative affective stimuli; Davey et al., 2013; Michalak et al., 2014; Michalak et al., 2015; Schnall & Laird, 2007). Therefore, extant research supports the notion that body states are promising targets in the treatment of depression. Interestingly, none of the cited studies could find effects of deliberately manipulating body states on self-reported affect. That said, this lack of evidence aligns with ICS theory, insofar as it posits that shifts in affective schemas are typically cued by the interplay of several factors as opposed to just a single one (Teasdale & Barnard, 1993, p. 188).

Regarding voice, research indicates that several aspects of speech are associated with depression (for review see Cummins et al., 2015). For example, depressive symptom severity was found to be associated with reduced speech rate and more pause time (Cannizzaro et al., 2004; Mundt et al., 2012), lower pitch and less pitch variation (Hönig, Batliner, Nöth, Schnieder, & Krajewski, 2014; Mundt et al., 2007; Quatieri & Malyska, 2012), as well as greater harshness and breathiness (Hönig, Batliner, Nöth, Schnieder, & Krajewski, 2014; Quatieri & Malyska, 2012). Furthermore, improvements in depressive symptomatology after treatment have been shown to correspond with a normalization of prosody (Alpert et al., 2001; Mundt et al., 2007, 2012). Another study investigating vocal indicators of change processes during psychotherapy sessions found that patients' use of a "determined voice" was associated with being stubbornly stuck in the psychotherapeutic progress, whereas use of a "questioning voice" directed toward the therapist was a vocal indicator of curative progress (Tomicic et al., 2015). The authors concluded that patients typically used a determined voice in therapy when holding steadfast to their preexisting dysfunctional beliefs, thus hampering the development of more functional ones. In contrast, being more open-minded and inquisitive towards the therapist and allowing for questioning or challenging of their preexisting dysfunctional beliefs appeared to make modification of such beliefs more likely. In yet another study, Aucouturier and colleagues developed an application enabling real-time manipulation of participants' voices that made them sound happier, sadder, or more anxious. Hearing their own altered voices led to changes in participants' self-reported affect and skin conductance levels congruent with the respective emotion (Aucouturier et al., 2016). Taken together, these studies provide evidence that systematically shaping how persons intonate anti-depressive self-statements could enhance the efficacy of cognitive interventions for depression. However, no study to date has evaluated the specific effects of experimentally manipulating depressed persons' prosody when uttering anti-depressive self-statements, targeting (somato-)sensory and cognitive processing simultaneously.

To fill this gap in the literature, we tested the hypothesis that depressed individuals experience a greater reduction of depressed mood when they are explicitly prompted to vocalize *invalidating* responses to *depressogenic* statements and *validating* responses to *salutogenic* statements in a *convincing* voice, compared to depressed individuals vocalizing the same set of responses *without* receiving any instructions on prosody. Additionally, we explored whether potential effects of such an intervention would be specific to individuals meeting criteria for MDD (but not for sub-clinically depressed or non-depressed individuals) and whether the degree of expressed conviction when vocalizing anti-depressive self-statements would show associations with depressed mood and depressive symptom severity.

1. Method

1.1. Participants and procedures

A power analysis using G*Power (Faul et al., 2007) based on the effect size from a similarly designed study testing strategies for reducing depressed mood (Diedrich et al., 2016) indicated that a sample size of $N = 40$ ensures power of $\beta = 0.80$ for a between-group comparison with two factors having critical alpha set at 5%. It should be noted that this power calculation only accounts for a single intervention period, whereas we analyzed mood assessments over time with a multilevel

model. We disregarded this feature of the study in the power analysis because the number of repetitions is relatively small and unlikely to significantly affect power and because arbitrary assumptions on the intraclass correlation may lead to unreliable power estimates (De Jong et al., 2010). Surpassing our power analysis requirement, we recruited $n = 48$ participants meeting criteria for MDD, $n = 48$ participants suffering from elevated, yet subclinical (SC) depressive symptom severity (PHQ-9 > 4), and $n = 48$ never-depressed (ND) participants (PHQ-9 < 5, no self-reported prior history of MDD (clinical interview described below), resulting in a total sample of $N = 144$. All participants were at least 18 years old. Exclusion criteria included a current diagnosis of bipolar, psychotic, or substance-related disorder (except for nicotine) within the past six months, and any exposure to psychotherapeutic treatment during the past six months. Participants received up to 150€ for study participation, depending on the total number of completed assessments. Upon completion of the study, participants from the MDD sample were offered CBT at the outpatient clinic for psychotherapy of Friedrich-Alexander-Universität Erlangen-Nürnberg, where the study had been conducted. Overall, participants in the final sample had a mean age of 32.72 years (ranging from 20 to 63, $SD = 11.02$). Most participants (67%) were female, and 58% of participants in the MDD sample had at least one comorbid psychiatric diagnosis. Table 1 provides an overview of sociodemographic and comorbidity data.

Participants were recruited from the waitlist of the university's outpatient psychotherapeutic clinic by advertising on the treatment center's official website, by posting on relevant social media platforms, and by circulating flyers among local psychiatrists and directly to prospective participants in public places. All potential participants completed an initial online screening questionnaire followed by an in-person diagnostic session, during which final eligibility was assessed (see assessment section for details). All invited participants were matched for age and gender across the three samples. If potential participants met all inclusion and no exclusion criteria, the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960) was administered to them to determine baseline severity of depressive symptoms.[1] They were subsequently allocated to the experimental or control condition based on two Microsoft Excel randomization lists with block sizes of four, separated across gender. Upon completion of the experimental session, 95.8% of participants in the MDD sample started CBT at the university's outpatient clinic for psychotherapy. HRSD follow-up interviews were conducted by phone. The average number of post-experiment CBT sessions was 2.12 ($SD = 1.60$) at the 1-month follow-up interview, 6.23 ($SD = 2.43$) sessions at the 3-month follow-up interview, and 13.10 ($SD = 5.06$) sessions at the 6-month follow-up interview.[2] A flow chart of study procedures is depicted in Fig. 1. All study procedures were approved by the ethics committee of the Friedrich-Alexander-Universität Erlangen--Nürnberg. The trial was pre-registered in the German clinical trials registry under https://drks.de/search/en/trial/DRKS00023670 during data collection and prior to data analyses.

Table 1
Sociodemographic data, depressive symptom severity, and comorbid disorders.

	MDD sample	SC sample	ND sample	X^2	F	p
Age	32.98 (11.10)	32.22 (10.63)	32.85 (11.55)		0.05	0.955
Sex, female, *n* (%)	32 (66.7)	32 (66.7)	32 (66.7)	0.00		1.000
Highest education, *n* (%)				4.24		0.374
No school degree	1 (2.1)	–	–			
High school degree	20 (41.7)	24 (50.0)	24 (50.0)			
University degree	27 (56.3)	24 (50.0)	24 (50.0)			
HRSD (Mean, SD)	15.29 (5.81)	4.33 (3.93)	1.44 (1.75)		147.26	<0.001
Previous MDE, *n* (%)	18 (37.5)	12 (25.0)	–		45.98	<0.001
Current comorbid disorders	26 (54.2)	2 (4.2)	–	55.69		<0.001
Anxiety disorders	18 (37.5)	–	–			
Obsessive-compulsive disorder	8 (16.7)	–	–			
Attention deficit hyperactivity disorder	6 (12.5)	–	–			
Eating disorder	–	1 (2.1)	–			
Adjustment disorder	–	1 (2.1)	–			

Note. MDD sample = participants meeting criteria for major depressive disorder; SC sample = participants with subclinical depressive symptoms; ND sample = non-depressed participants; HRSD = Hamilton Rating Scale for Depression; MDE = major depressive episode. Due to technical difficulties, one participant's data was incomplete.

1.2. Experimental manipulation of intonation

Data for this study was collected at the university's outpatient clinic for psychotherapy between September 2020 and April 2023. General instructions for the experiment were presented to participants on a smartphone (Motorola G8 Lite) facilitated by an app programmed specifically for this purpose. The experimental context was designed to resemble a psychotherapeutic session, focusing first on the *invalidation of depressogenic self-statements* (Phase 1), followed by the *validation of anti-depressive self-statements* (Phase 2). This sequence was intended to correspond to the typical sequential procedures of clinical practitioners and to thus develop an ecologically valid intervention. To prevent floor effects and synchronize mood states between participants, we induced depressed mood in Phase 1. To prevent depressed mood from persisting beyond the duration of the experiment (if – contrary to our expectations – the intervention would not succeed in reducing depressed mood), we induced positive mood in Phase 2. All scripted self-statements are documented in the Supplemental Materials.

The experiment was comprised of two Phase 1 blocks and two Phase 2 blocks. Each block of Phase 1 started with a negative mood induction, in which participants listened to an excerpt from *Adagio in G minor* by Tomaso Giovanni Albinoni, played at half the original speed, while they read aloud five statements designed to induce depressed mood (e.g., "My future is absolutely hopeless."; see Velten, 1968; Diedrich et al., 2014; Diedrich et al., 2016 for more details on rationale and efficacy of this procedure). After this negative mood induction, the same items of each respective block were again presented to participants in sequence. This time, for each depressogenic statement, participants were asked to select one of three possible scripted coping responses (i.e., anti-depressive self-statements) to invalidate the depressogenic statement (e.g., "No! I don't see it that way! I don't give up that quickly!"). Participants were instructed to select the scripted anti-depressive self-statement that they considered most likely to effectively neutralize any exacerbating effects of the previous depressogenic statement on their mood and to read the selected statement aloud three times. For the second block of Phase 1, this procedure was repeated with a new set of scripted statements.

[1] At this point, participants had also provided a speech sample at site and completed an ecological momentary assessment designed to develop a machine learning model detecting depression from speech. As these assessments/data were not used for the present study, we do not report them in detail.

[2] Number of CBT sessions at the time of follow-up interviews was not available for four participants.

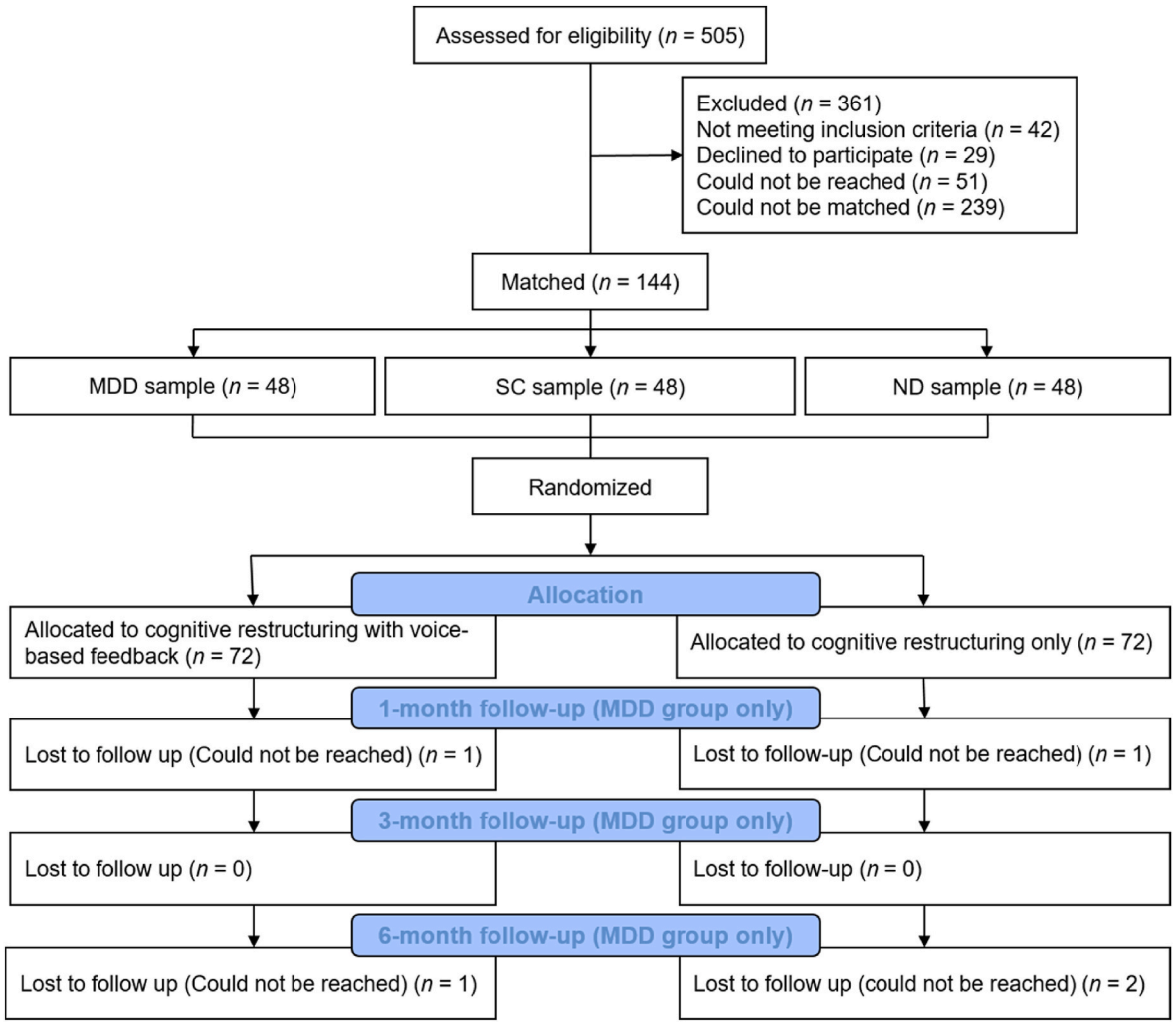

Fig. 1. Flowchart of participants.

Next, the two blocks of Phase 2 were each first preceded by *positive* mood induction consisting of reading aloud five salutogenic statements (e.g., "I deserve to be happy and to be loved."). Analogous to Phase 1, mood induction was followed by the repeated sequential presentation of each of the salutogenic statements from the prior mood induction. For each salutogenic statement, participants were asked to select one of three possible scripted coping responses to be used to validate the statement (e.g., "Yes, I too am loved!"). Participants were instructed to select the validating statement that they considered most likely to further enhance any positive effect of the salutogenic statement on their mood, and to read the statement aloud three times.

In the experimental condition, participants received additional instructions on how to intone the scripted coping response statements. As such, they were instructed to modulate their voice in a way that expressed as much conviction as possible (focusing on aspects of prosody such as loudness, emphasis, and intonation). During the intervention, the experimenter provided feedback on how participants could further maximize the expression of conviction in their voice. This feedback was individualized for each participant and again focused on loudness (e.g., "Try to speak up when reading the statement"), emphasis (e.g., "Try to emphasize words that are particularly relevant for you in this statement"), and intonation (e.g., "Try to lower your pitch at the end of the statement"). In the control condition, participants received identical instructions about the intervention procedures and occasional encouraging feedback, but no specific instructions on how to better intonate their coping statements (e.g., "Well done! Please read it another time.").

In both study conditions and for all coping statements, experimenters rated the participant's voice with regards to the level of expressed conviction using a visual analogue scale (VAS; range 0–10; rated via tablet). In addition, before and after each mood induction as well as after each experimental block, participants rated their depressed mood on a VAS (range 0–10). After the experiment, participants were debriefed by the experimenter and asked how they were currently feeling. In case of significant mood deterioration during the experiment, an experienced psychotherapist was available to provide crisis intervention. This safety measure only had to be applied for one patient.

1.3. Measures

The clinical status of participants was assessed with the German version of the Structured Clinical Interview for DSM-5 (SCID; First et al., 2016; German version: Beesdo-Baum et al., 2019). All interviewers were thoroughly trained in administering the SCID and were supervised by a senior clinical diagnostician/therapist. To assess interrater reliability, a blinded and experienced rater re-rated 10% of the videotaped diagnostic interviews. A Cohen's κ of 0.84 indicated excellent agreement on the presence versus absence of an MDD diagnosis.

The assessment of depressed mood was done on an 11-point VAS ranging from 0 to 10. Participants were asked to indicate their current depressed mood on a horizontal line with eleven points presented on a smartphone screen. VAS ratings have been previously determined to provide valid assessments of depressed mood in experimental studies on

depressed mood regulation (Diedrich et al., 2014, 2016; Ehret et al., 2018).

To assess depressive symptom severity, we used a semi-structured 17-item version of the HRSD (Miller et al., 1985). The HRSD is a widely used semi-structured clinical interview, in which each item is rated on a scale from 0 to 4 or 0 to 2, depending on the item, with the total score range being 0 to 52. While including the same items and score ranges as the original version (Hamilton, 1960), the semi-structured version (Miller et al., 1985) includes additional prompts to gain relevant information about each item. Specifically, structured versions of the HRSD provide a reliable and valid assessment of depressive symptom severity (Carrozzino et al., 2020). We examined interrater reliability by having blinded and experienced raters re-rate 10% of the videotaped interviews. Pearson's correlation coefficient for mean scores was 0.99, indicating excellent interrater reliability. To evaluate the inter-rater reliability of the original, audio-recorded conviction ratings that were conducted by the experimenter during the experiment, 10% of those recordings were randomly selected to be re-rated by two blinded raters. Between the original rating and the two re-ratings, we found intraclass correlation coefficients of 0.89 for the entire experiment, 0.87 for Phase 1, and 0.86 for Phase 2, indicating good inter-rater reliability for this measure.

1.4. Statistical analyses

For initial manipulation checks, we tested the effect of the mood induction procedures on depressed mood with paired t-tests and the effect of study condition on expressed conviction with an analysis of variance (ANOVA). To test the effect of study condition and sample on depressed mood after reading anti-depressive self-statements, we employed a multilevel model. To account for the hierarchical structure, we nested observations within each participant by adding random intercepts for participants. Further, we controlled for depressed mood before each block. Since we did not have any directed hypotheses on differential effects for the two intervention phases, we compared a model that does not include the two phases with a model that includes random intercepts for the two phases and decided to interpret the model that achieved the better fit. Two additional multilevel models were employed to explore effects of study condition and sample on depressed mood in Phase 1 and Phase 2. To examine the extent to which expressed conviction in depressed individuals is associated with depression, we computed Pearson's product-moment correlations between expressed conviction, depressed mood, and depressive symptom severity. Correlations were computed separately for the two study conditions and the total sample. Expressed conviction describes conviction ratings averaged over the entire experiment. Finally, we calculated effect sizes for the entire experiment and for both phases. Effect sizes of 0.01, 0.06, and 0.14 for η^2 and 0.2, 0.5 and 0.8 for Cohen's d were a priori defined as reflecting small, medium, and large effects, respectively (Cohen, 1988). We used two-sided tests across analyses. For calculating multilevel models, we used the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2014). The significance of the fixed effects was assessed using the lmerTest package (Kuznetsova et al., 2017), and for examining interaction effects, we used the reghelper package (Hughes & Beiner, 2021). We further used the clubSandwich package (Pustejovsky & Tipton, 2018) to obtain robust standard errors adjusting for heteroscedasticity. All other analyses were computed with IBM SPSS Statistics, Version 29. Materials and analysis code for this study are available by emailing the corresponding author.

2. Results

2.1. Preliminary analyses

2.1.1. Randomization checks

Sociodemographic data are depicted in Table 1. Randomization checks yielded no significant differences between study conditions with regards to sociodemographic and clinical characteristics (all $ps \geq 0.404$). Similarly, there were no significant differences between study conditions with regards to baseline depressed mood in the MDD sample ($t(46) = 1.03$, $p = .310$, $d = 0.29$), in the SC sample ($t(46) = 0.38$, $p = .707$, $d = 0.11$), and in the ND sample ($t(46) < 0.00$, $p = 1.000$, $d = 0.00$).

2.1.2. Manipulation checks

Effects of mood inductions were calculated with paired t-tests. Results are shown in Table 2 and suggest that only negative mood induction procedures had a moderate effect on depressed mood, whereas the positive mood induction did not affect depressed mood.

Descriptive data for conviction ratings by the experimenter over the course of the procedure are summarized in Fig. 2. The Figure illustrates how participants' expressed conviction increased during the three vocalized repetitions of each anti-depressive self-statement, as well as over the entire experiment. Additionally, it illustrates how the increase of conviction over the three repetitions and over the entire experiment differs across study conditions. An ANOVA comparing experimenter-rated expressed conviction between study conditions and samples revealed a significant main effect of study condition ($F(1,138) = 142.73$, $p < .001$, partial $\eta^2 = 0.51$) and a significant interaction between sample and condition ($F(2,138) = 8.27$, $p < .001$, partial $\eta^2 = 0.11$). Bonferroni-corrected post-hoc tests revealed that expressed conviction was higher in the experimental condition across all samples (all $ps < 0.001$) and that expressed conviction was significantly lower in the MDD sample compared to both the SC sample ($p = .008$) and the ND sample ($p = .003$). When testing differences in re-ratings of expressed conviction between study conditions, we found significant main effects of study condition for re-rater 1 ($F(1,136) = 127.03$, $p < .001$, partial $\eta^2 = 0.49$) and for re-rater 2 ($F(1,136) = 101.76$, $p < .001$, partial $\eta^2 = 0.43$), revealing higher expressed conviction for participants in the experimental condition.

2.1.3. Assumptions of multilevel models

None of the multilevel models in this study met the assumption of homoscedasticity and residuals were not normally distributed. We adjusted for heteroscedasticity by using robust standard errors for all models. We applied log and square root data transformations but omitted transformed data for the sake of brevity, since the residuals of the transformed data also showed a non-normal distribution.

2.2. Primary outcome: effects of study condition and sample on depressed mood in the entire experiment

Mean scores of depressed mood over the course of the experiment are shown in Table 3 and depicted in Fig. 3. Effect sizes for the entire experiment and each individual phase are shown in Table 4.

A chi-square test showed a significantly better fit for the model including random slopes for the experimental phases (AIC = 1296.6, BIC = 1344.5) compared to the model without random slopes for the phases (AIC = 1350.2, BIC = 1389.4; $\chi^2(2) = 57.63$, $p < .001$). Therefore, we report the results of the former. Since both sample and study condition are categorical variables, effects are reported in reference to the ND sample and the control condition. The model revealed a significant main effect of MDD sample (estimate = 1.81, $SD = 0.36$, $t(50.5) = 5.04$, $p < .001$), no significant main effect of SC sample (estimate = 0.24, $SD = 0.14$, $t(46.3) = 5.04$, $p = .082$), no significant main effect of experimental condition (estimate = −0.01, $SD = 0.04$, $t(46.0) = -0.14$, $p = .889$), and a significant interaction between MDD sample and experimental condition (estimate = −0.72, $SD = 0.34$, $t(91.6) = -2.13$, $p = .036$) on depressed mood after each reading of anti-depressive statements, controlled for depressed mood before each reading. Simple slopes analysis showed a significant effect of study condition on depressed mood in the MDD sample (estimate = −0.75, $SD = 0.23$, t

Table 2
Manipulation checks testing the effect of mood induction procedures.

Block	MDD sample			SC sample			ND sample		
	t	p	d	t	p	d	t	p	d
Negative mood induction									
Block 1	4.14	<0.001	0.60	6.51	<0.001	0.94	5.37	<0.001	0.77
Block 2	2.72	0.009	0.39	4.06	<0.001	0.59	4.09	<0.001	0.59
Positive mood induction									
Block 3	0.89	0.377	0.13	1.23	0.224	0.18	1.27	0.209	0.18
Block 4	0.22	0.830	0.03	1.95	0.057	0.28	1.43	0.159	0.21

Note. MDD sample: participants meeting criteria for major depressive disorder; SC sample: participants with subclinical depressive symptoms; ND sample: non-depressed participants.

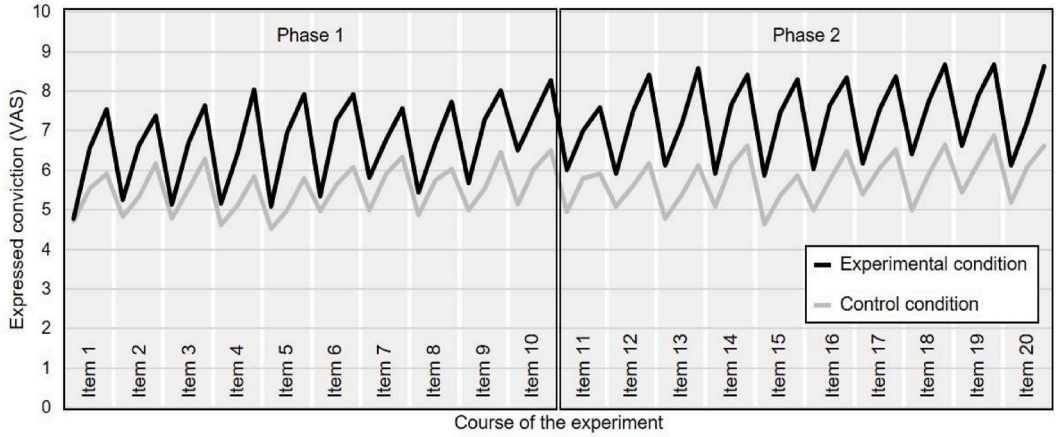

Fig. 2. Ratings of expressed conviction over the course of the experiment in the MDD sample
Note. Each item was read three times and each reading was rated by the experimenter on a Visual Analogue Scale (VAS) from 0 to 10.

Table 3
Means and standard deviations of depressed mood.

	MDD sample		SC sample		ND sample	
	EC	CC	EC	CC	EC	CC
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
t0	3.08 (1.93)	3.75 (2.52)	0.63 (1.10)	0.75 (1.19)	0.08 (0.41)	0.08 (0.28)
t1[a]	3.96 (2.22)	4.29 (2.60)	1.75 (1.85)	1.92 (1.74)	1.17 (1.49)	1.33 (1.55)
t2[b]	3.00 (1.89)	3.71 (2.29)	0.71 (1.55)	0.83 (1.27)	0.25 (0.53)	0.29 (0.55)
t3[a]	3.50 (1.82)	4.13 (2.11)	1.46 (1.74)	2.00 (2.23)	1.29 (2.22)	1.17 (1.55)
t4[b]	2.92 (1.67)	3.63 (2.45)	0.83 (1.69)	0.92 (1.64)	0.17 (0.38)	0.33 (0.87)
t5	3.13 (1.80)	3.65 (2.25)	0.67 (1.31)	0.79 (1.32)	0.08 (0.28)	0.21 (0.66)
t6[c]	3.08 (1.82)	3.52 (2.37)	0.50 (1.02)	0.71 (1.04)	0.08 (0.28)	0.04 (0.20)
t7[d]	2.38 (1.74)	3.43 (2.29)	0.50 (0.83)	0.63 (1.21)	0.08 (0.28)	0.13 (0.34)
t8[c]	2.42 (1.82)	3.35 (2.10)	0.38 (0.77)	0.50 (0.98)	0.08 (0.28)	0.04 (0.20)
t9[d]	2.08 (1.56)	3.09 (2.35)	0.33 (0.76)	0.46 (0.93)	0.08 (0.28)	0.00 (0.00)

Note. MDD sample: participants meeting criteria for major depressive disorder; SC sample: participants with subclinical depressive symptoms; ND sample: non-depressed participants; EC: experimental condition; CC: control condition.
[a] preceded by negative mood induction.
[b] preceded by invalidation of depressogenic statements.
[c] predeced by positive mood induction.
[d] preceded by validation of salutogenic statemtents.

(85.7) = −3.29, p = .001), but not in the SC sample (estimate = −0.05, SD = 0.23, t(84.0) = −0.22, p = .824) or the ND sample (estimate = 0.01, SD = 0.23, t(84.0) = −0.02, p = .981). Thus, the findings provide preliminary evidence for the specificity of the effects exclusively for individuals meeting full criteria for MDD.

2.3. Effects of study condition and sample on depressed mood in the phases of the experiment

For Phase 1 (invalidation of depressogenic statements), the model revealed a significant main effect of MDD sample (estimate = 2.23, SD = 0.41, t(47.9) = 5.47, p < .001), no significant main effect of SC sample (estimate = 0.29, SD = 0.24, t(46.2) = 1.24, p = .222), no significant main effect of experimental condition (estimate = −0.10, SD = 0.14, t(46.0) = −0.69, p = .494), and no significant interaction between MDD sample and experimental condition (estimate = −0.43, SD = 0.42, t(91.7) = −1.04, p = .302) and between SC sample and experimental condition (estimate = 0.13, SD = 0.35, t(91.90) = 0.36, p = .719).

For Phase 2 (validation of salutogenic statements), the model revealed no significant main effect of MDD sample (estimate = 0.12, SD = 0.12, t(36.8) = 1.01, p = .321), no significant main effect of SC sample (estimate = −0.03, SD = 0.06, t(47.1) = −0.56, p = .581), no significant main effect of experimental condition (estimate = −0.02, SD = 0.02, t(46.0) = −0.79, p = .436), a significant interaction between MDD sample and experimental condition (estimate = −0.39, SD = 0.15, t(87.0) = −2.65, p = .010), and no significant interaction between SC sample and experimental condition (estimate = 0.04, SD = 0.10, t(92.1) = 0.45, p = .653). Simple slopes analysis showed a significant effect of study condition on depressed mood in the MDD sample (estimate = −0.41, SD = 0.10, t(130.0) = −3.93, p < .001), but not in the SC sample (estimate = 0.03, SD = 0.10, t(129.7) = 0.26, p = .799) or the ND sample (estimate = −0.02, SD = 0.10, t(129.7) = −0.17, p = .869).

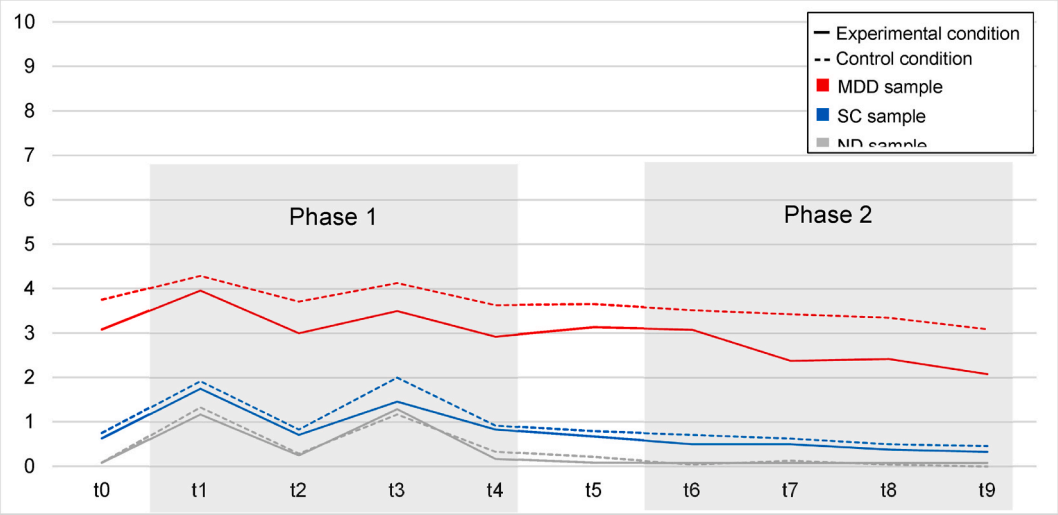

Fig. 3. Depressed mood over the course of the experiment
Note. Depressed mood was self-rated on a Visual Analogue Scale (VAS) from 0 to 10. Negative mood inductions took place between t0 and t1 and t2 and t3; positive mood inductions between t5 and t6 and between t7 and t8. Anti-depressive self-statements were read between t1 and t2, between t3 and t4, between t6 and t7, and between t8 and t9.

Table 4
Depressed mood ratings in the MDD sample pre- and post-intervention for the entire intervention, Phase 1, and Phase 2.

	Experimental condition		Control condition	
	ΔM (SD)	*d* [95%-CI]	ΔM (SD)	*d* [95%-CI]
MDD sample				
Entire Intervention	1.88 (1.65)	1.14 [0.61, 1.65]	1.09 (1.86)	0.59 [0.14, 1.02]
Phase 1	1.04 (0.96)	1.09 [0.58, 1.59]	0.67 (1.31)	0.51 [0.79, 0.93]
Phase 2	1.00 (1.02)	0.98 [0.48, 1.46]	0.44 (0.84)	0.52 [0.07, 0.95]
SC sample				
Entire Intervention	1.42 (1.44)	0.98 [0.49, 1.47]	1.46 (1.35)	1.08 [0.57, 1.58]
Phase 1	0.92 (1.14)	0.81 [0.34, 1.26]	1.00 (1.06)	0.94 [0.45, 1.42]
Phase 2	0.17 (0.57)	0.30 [–0.12, 0.70]	0.25 (0.53)	0.47 [0.04, 0.89]
ND sample				
Entire Intervention	1.08 (1.50)	0.72 [0.27, 1.17]	1.33 (1.55)	0.86 [0.38, 1.32]
Phase 1	1.00 (1.38)	0.72 [0.27, 1.17]	1.00 (1.18)	0.85 [0.37, 1.31]
Phase 2	–	–	0.04 (0.20)	0.20 [–0.20, 0.61]

Note. ΔM = mean differences between pre- and post-depressed mood; *d* refers to Cohen's *d* and indicates estimated effect sizes; CI = confidence intervals; MDD sample = participants meeting criteria for major depressive disorder; SC sample = participants with subclinical depressive symptoms; ND sample = non-depressed participants.

2.4. Exploratory analyses of associations between conviction and symptoms of depression

Descriptive statistics of expressed conviction are shown in Table 5. Across study conditions, depressed mood showed a significant association with expressed conviction ($r = -0.185$; $p = .027$). Within study conditions, expressed conviction was significantly associated with depressed mood in the experimental condition ($r = -0.526$; $p < .001$), but not in the control condition ($r = 0.067$; $p = .575$). Regarding overall depressive symptom severity, we found no significant associations between HRSD and expressed conviction across study conditions ($r = 0.149$; $p = .074$). Within study conditions, there was a significant association of HRSD with expressed conviction of participants only in the experimental condition ($r = -0.462$; $p < .001$), but not in the control condition ($r = -0.015$; $p = .901$). The significant associations could all be confirmed with re-ratings of expressed conviction (see Supplemental Table 9).

Table 5
Ratings of expressed conviction by the experimenter (original) and two re-raters.

Conviction ratings	MDD sample		SC sample		ND sample	
	EC	CC	EC	CC	EC	CC
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Original	7.01 (1.36)	5.63 (1.72)	8.18 (0.92)	4.87 (1.54)	8.31 (1.14)	5.06 (1.11)
Re-rater 1	6.94 (1.72)	4.33 (2.26)	7.89 (1.23)	4.43 (1.76)	7.87 (1.13)	4.60 (1.55)
Re-rater 2	7.36 (1.16)	4.64 (1.65)	6.88 (1.24)	4.46 (1.19)	6.43 (1.32)	4.89 (1.31)

Note. MDD sample: participants meeting criteria for major depressive disorder; SC sample: participants with subclinical depressive symptoms; ND sample: non-depressed participants; EC: experimental condition; CC: control condition.

3. Discussion

The primary goal of this study was to test the hypothesis that particularly depressed individuals would benefit to a greater extent from uttering anti-depressive self-statements if they were systematically guided to express a maximum of conviction when vocalizing these statements. To test this hypothesis, we conducted an experiment with $n = 48$ participants meeting criteria for MDD, $n = 48$ participants with subclinical depressive symptoms, and $n = 48$ never-depressed participants. As hypothesized, clinically depressed individuals who received instructions on both content and prosody of anti-depressive self-statements reported a significantly greater reduction of experimentally induced depressed mood than did similarly depressed individuals who received the same content-related instructions but no instructions on prosodic execution. Our results further suggest that the mood-regulating effect of anti-depressive self-statements, which showed a moderate effect size comparable to common emotion regulation strategies (e.g., Diedrich et al., 2014), can be increased to a large effect by adding

prosody modulation instructions. For sub-clinically depressed participants and never-depressed participants, adding prosody modulation instructions did not increase the mood-regulating effect of anti-depressive self-statements.

This finding is consistent with a large body of empirical findings indicating that systematically manipulating sensory information can have significant effects on cognitive processes associated with depression (e.g., Michalak et al., 2014; Price & Harmon-Jones, 2015). The present study extends beyond previous research by experimentally manipulating prosody insofar as depressed individuals were instructed to vocalize anti-depressive self-statements. At first glance, findings from the present study appear to contradict those of a previous one that also investigated possible correlations between prosody and treatment processes in patients (Tomicic et al., 2015). Specifically, whereas Tomicic and colleagues (2015) found that the greater use of an affirmative voice in therapy negatively predicted change processes, we found that systematically enhancing vocally expressed conviction was associated with a greater reduction of depressed mood. Presumably, the naturalistic speech samples that Tomicic and colleagues derived from treatment sessions may have been convoluted with vocalizations of *dysfunctional* self-statements and beliefs. Logically, in the case of dysfunctional self-statements, any affirmative prosody is likely to produce negative effects on treatment. Conversely, in the present study we focused exclusively on the prosody for anti-depressive self-statements. In this case, a greater expression of conviction is assumed to validate functional beliefs and, therefore, help overcome depression.

From a theoretical perspective, findings from the present study are in keeping with theories claiming that sensory perceptions can moderate the affective consequences of cognitive information processing. As reviewed in the introduction, ICS theory (Teasdale & Barnard, 1993) posits that affective schemas are activated if emotion-specific patterns are detected in the incoming multimodal stream of information. Since the rational assessment of the situation conducted in the *propositional subsystem* is a major source of information for the *implicational subsystem*, anti-depressive thoughts (or their vocalization) have the potential to reduce the likelihood of activating schemas in the *implicational subsystem* that are incompatible with the depressogenic schema (e.g., hope, self-compassion, confidence). However, if a patient vocalizes these thoughts in an unconvincing way, the *implicational subsystem* will receive information from the *body state subsystem* and the *acoustic subsystem* that is incompatible with the anti-depressive schema. The *implicational subsystem* has likely learned that the rational interpretation of the situation might be false (and simply driven by the desire to avoid painful emotions or to please one's therapist) and that information from the *body state* and *acoustic subsystem* provide important information on the validity of one's rational interpretation. Therefore, invalidating information emanating from the sensory subsystems can interfere with the activation of the anti-depressive schema and, hence, perpetuate the activation of the depressogenic schema. Thus, our finding that systematically amplifying the expression of conviction when vocalizing anti-depressive self-statements correlated with a greater decrease of depressed mood is aligned with ICS core assumptions.

Moreover, it is of note that previous studies manipulating the somatosensory input (namely posture, gait, or facial expressions) found effects on cognitive processes associated with depression, but not on self-rated negative mood (e.g., Michalak et al., 2014; Schnall & Laird, 2007). This apparent contradiction to the present study, arguably, may have resulted from our induction of validating somatosensory and auditory input while participants were engaged in anti-depressive cognitions, whereas this was not the case in previous studies. Our findings are consistent with the (ICS-based) hypothesis that the combination of anti-depressive cognitive activity and validating somatosensory and auditory information is more effective than either cognitive activity or (somato-)sensory input alone when it comes to effectively reducing depressed mood.

The finding that differences between experimental conditions were only found in the MDD sample, but not in the SC and the ND sample is consistent with the proposition of ICS theory that depressed individuals are entangled in a so-called "depressive interlock." This term refers to a vicious cycle, occurring when the activation of a depressogenic schema in the *implicational subsystem* leads to changes in body and cognitions, which are then fed back to the *implicational subsystem* where they lead to the re-activation of the depressogenic schema and so forth (Teasdale & Barnard, 1993, pp. 168–171). Thus, in individuals with subthreshold depressive symptoms, addressing only one modality (e.g., cognitions) may suffice to reduce depressed mood significantly, since they are not held captive in a perpetual state of depressive interlock. However, individuals who have developed clinically relevant depressive symptoms may need anti-depressive input from more than one modality to effectively disrupt the reciprocal feedback loop between depressogenic cognitions and (somato-)sensory perceptions, such as was the case in our experimental condition that systematically coupled cognitive and prosodic information.

Apart from theoretical implications, our findings have very specific significance for clinical practice. They provide important proof that systematically encouraging depressed patients to maximize their expression of conviction when vocalizing anti-depressive self-statements helps to enhance the anti-depressive effects of such interventions. Likely, many experienced practitioners already acknowledge the importance of prosody (and other ways of using the body to validate cognitions) when developing anti-depressive self-statements with patients. However, to the best of our knowledge CBT manuals on the treatment of depression usually focus extensively on the content of depressogenic and anti-depressive self-statements, and only rarely on prosody (or other somatosensory ways of validating anti-depressive self-statements). Consequently, there is a risk that less experienced therapists underutilize prosody when employing anti-depressive self-statements as a technique. Therefore, the results of the present study should be incorporated in future CBT manuals to more fully exploit the potential of cognitive interventions against depression. Similarly, in an ad-hoc review of digital and book-based CBT self-help interventions, we found ample references to the content of anti-depressive self-statements, but none to the importance of prosody. Although this might only be one of many reasons why self-help interventions are less effective than traditional, guided CBT (Cuijpers et al., 2019) particularly in patients with more severe depressive symptoms (Karyotaki et al., 2021), it is conceivable that without a therapist intuitively shaping the prosody of anti-depressive self-statements, these interventions cannot fully harness the curative potential of these statements. Therefore, future self-help interventions should guide users to also focus on prosody and not only on the content of anti-depressive statements.

A major limitation of the study derives from the fact that we did not balance the order in which participants invalidated depressogenic cognitions (Phase 1) and validated salutogenic self-statements (Phase 2), whereby our rationale was twofold. First, in the psychotherapeutic treatment of depression, therapists typically start with the identification and invalidation of depressogenic automatic thoughts/beliefs and only subsequently proceed with the development and utilization of salutogenic self-statements (Clark, 2013). Thus, by adhering to this organic order we, arguably, enhanced the ecological validity of the study. Secondly, experimentally inducing depressed mood in depressed individuals in scientific studies is ethically questionable, since residual depressed mood may prevail after the experiment (Frost & Green, 1982). Therefore, we decided to end the experiment with a phase including positive mood induction followed by the vocalized validation of salutogenic statements, which was considered conducive to minimizing the level of depressed mood participants might suffer after completing the experiment. However, the disadvantage of refraining from randomizing the orders of experimental phases is that we cannot disentangle the effects of content and time. Instead, the difference between the phases may either indicate that it took more time until the instructions regarding prosody lead to a greater decrease of depressed mood, or that

prosody-related instructions exclusively enhanced the efficacy of validating salutogenic statements only, but not the efficacy of invalidating depressogenic statements. Based on our observations during the experiment (and in keeping with ICS theoretical premises), we are inclined to believe that a certain number of repetitions is needed until manipulations of expressed conviction take effect. Nevertheless, future studies should balance the two phases to clarify this question empirically. In addition to this issue, we found that while the negative mood induction procedure increased depressed mood, the positive mood induction procedure did not. For depressed participants, this may be explained by a dampening effect, which led to a lack of mood improvement (Bean et al., 2022), whereas for non-depressed participants, a floor effect may have curbed or prevented mood improvement. However, our decision to include a positive mood induction component represented a precaution to assure our participants' well-being at the close of the experiment; we do not expect that it had any effect on the results or their interpretation. Nevertheless, we recommend that future studies randomize the order of negative and positive mood induction procedures prior to participants' invalidation of depressogenic self-statements and validation of salutogenic self-statements. Another limitation is that expressed conviction was exclusively, subjectively assessed by the experimenter. Thus, we are unable to examine to what extent the expression of conviction must be *genuinely* experienced (rather than merely shown) to enhance the effect of anti-depressive statements on depressed mood. The decision to forgo self-reports on conviction was based on the findings from prior feasibility trials indicating that such assessments would significantly interfere with the flow of the intervention. Thus, future studies will need to meet the challenge of complementing observer-based assessment of conviction with participant self-reports without interfering with the intervention. Finally, our multilevel models did not meet the assumption of normally distributed residuals. Most plausibly, this may be explained by a skewed distribution of depressed mood ratings and differences in variance with lower ratings and variances in the ND and SC samples compared to the MDD sample. However, there is data suggesting multilevel models are largely robust against non-normally distributed residuals (Schielzeth et al., 2020). Further, the effect sizes shown in Table 4 provide additional proof of the efficacy of our intervention for depressed patients.

Directions for future research include the application of prosody-based interventions to other mental disorders, and the use of sensory modalities other than voice when trying to enhance anti-depressive self-statements. For example, in a currently ongoing study, we are investigating whether instructing depressed individuals to invalidate depressogenic self-statements with facial expressions of rejection (e.g., frowning) and validate salutogenic statements with facial expressions of approval (e.g., slight nod) leads to a greater reduction of experimentally induced depressed affect than the vocalization of the same self-statements without any manipulation of facial expressions (https://www.empkins.de/research/sub-projects/d02/; for a similar study see Keinert et al., 2023). Ideally, future studies would compare the effects of anti-depressive self-statements alone with the effects of such statements augmented by validating prosody, by validating facial expressions, by validating body movements, and by all of these ways of embodied augmentation combined.

CRediT authorship contribution statement

Jonathan F. Bauer: Writing – original draft, Methodology, Formal analysis, Data curation. **Lena Schindler-Gmelch:** Writing – review & editing, Formal analysis. **Maurice Gerczuk:** Writing – review & editing, Software, Methodology. **Björn Schuller:** Writing – review & editing, Project administration, Funding acquisition, Conceptualization. **Matthias Berking:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Ethical approval

The study was approved by the ethics board of the Friedrich-Alexander-Universität Erlangen-Nürnberg (18-73-B).

Submission declaration

The study's hypotheses were preregistered at the German Clinical Trials Register (Deutsches Register Klinischer Studien, ID: DRKS00023670) on January 17, 2021 https://drks.de/search/en/trial/DRKS00023670).

Funding

This study was supported by grants from the German Research Foundation to Matthias Berking (BE4510/11-1), to Björn Schuller (SCHU2508/8-1), and to Jarek Krajewski (KR3698/9-1).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Lisa Meisel, Lena Grüttner, Cem Akin, and Marleen Bogdanov for their contributions to this research project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2024.104667.

Data availability

Data will be made available on request.

References

Adolph, D., Tschacher, W., Niemeyer, H., & Michalak, J. (2021). Gait patterns and mood in everyday life: A comparison between depressed patients and non-depressed controls. *Cognitive Therapy and Research, 45*(6), 1128–1140. https://doi.org/10.1007/s10608-021-10215-7

Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders, 66*(1), 59–69. https://doi.org/10.1016/S0165-0327(00)00335-9

APA. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Aucouturier, J.-J., Johansson, P., Hall, L., Segnini, R., Mercadié, L., & Watanabe, K. (2016). Covert digital manipulation of vocal emotion alter speakers' emotional states in a congruent direction. *Proceedings of the National Academy of Sciences of the United States of America, 113*(4), 948–953. https://doi.org/10.1073/pnas.1506552113

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823 https://doi.org/10.18637/jss.v067.i01.

Bean, C. A., Summers, C. B., & Ciesla, J. A. (2022). Dampening of positive affect and depression: A meta-analysis of cross-sectional and longitudinal relationships. *Behaviour Research and Therapy, 156*, Article 104153. https://doi.org/10.1016/j.brat.2022.104153

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

Beesdo-Baum, K., Zaudig, M., & Wittchen, H.-U. (2019). *SCID-5-CV: Strukturiertes Klinisches Interview für DSM-5-Störungen - Klinische Version. Hogrefe*.

Beshai, S., Dobson, K. S., Bockting, C. L. H., & Quigley, L. (2011). Relapse and recurrence prevention in depression: Current research and future prospects. *Clinical Psychology Review, 31*(8), 1349–1360. https://doi.org/10.1016/j.cpr.2011.09.003

Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., & Snyder, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and Cognition, 56*(1), 30–35. https://doi.org/10.1016/j.bandc.2004.05.003

Carrozzino, D., Patierno, C., Fava, G. A., & Guidi, J. (2020). The Hamilton rating scales for depression: A critical review of clinimetric properties of different versions.

Psychotherapy and Psychosomatics, 89(3), 133–150. https://doi.org/10.1159/000506879
Casacalenda, N., Perry, J. C., & Looper, K. (2002). Remission in major depressive disorder: A comparison of pharmacotherapy, psychotherapy, and control conditions. *American Journal of Psychiatry, 159*(8), 1354–1360. https://doi.org/10.1176/appi.ajp.159.8.1354
Clark, D. A. (2013). Cognitive restructuring. In S. G. Hofmann (Ed.), *The wiley handbook of cognitive behavioral therapy* (pp. 1–22). John Wiley & Sons, Ltd.. https://doi.org/10.1002/9781118528563.wbcbt02
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum. https://doi.org/10.4324/9780203771587
Cristea, I. A., Huibers, M. J. H., David, D., Hollon, S. D., Andersson, G., & Cuijpers, P. (2015). The effects of cognitive behavior therapy for adult depression on dysfunctional thinking: A meta-analysis. *Clinical Psychology Review, 42*, 62–71. https://doi.org/10.1016/j.cpr.2015.08.003
Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry, 58*(7), 376–385. https://doi.org/10.1177/070674371305800702
Cuijpers, P., Noma, H., Karyotaki, E., Cipriani, A., & Furukawa, T. A. (2019). Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: A network meta-analysis. *JAMA Psychiatry, 76*(7), 700–707. https://doi.org/10.1001/jamapsychiatry.2019.0268
Cuijpers, P., van Straten, A., Warmerdam, L., & Andersson, G. (2008). Psychological treatment of depression: A meta-analytic database of randomized studies. *BMC Psychiatry, 8*(1), 36. https://doi.org/10.1186/1471-244X-8-36
Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication, 71*, 10–49. https://doi.org/10.1016/j.specom.2015.03.004
Davey, G. C. L., Sired, R., Jones, S., Meeten, F., & Dash, S. R. (2013). The role of facial feedback in the modulation of clinically-relevant ambiguity resolution. *Cognitive Therapy and Research, 37*(2), 284–295. https://doi.org/10.1007/s10608-012-9480-5
De Jong, K., Moerbeek, M., & Van der Leeden, R. (2010). A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects. *Psychotherapy Research, 20*(3), 273–284. https://doi.org/10.1080/10503300903376320
Diedrich, A., Grant, M., Hofmann, S. G., Hiller, W., & Berking, M. (2014). Self-compassion as an emotion regulation strategy in major depressive disorder. *Behaviour Research and Therapy, 58*, 43–51. https://doi.org/10.1016/j.brat.2014.05.006
Diedrich, A., Hofmann, S. G., Cuijpers, P., & Berking, M. (2016). Self-compassion enhances the efficacy of explicit cognitive reappraisal as an emotion regulation strategy in individuals with major depressive disorder. *Behaviour Research and Therapy, 82*, 1–10. https://doi.org/10.1016/j.brat.2016.04.003
Ehret, A. M., Joormann, J., & Berking, M. (2015). Examining risk and resilience factors for depression: The role of self-criticism and self-compassion. *Cognition & Emotion, 29*(8), 1496–1504. https://doi.org/10.1080/02699931.2014.992394
Ehret, A. M., Joormann, J., & Berking, M. (2018). Self-compassion is more effective than acceptance and reappraisal in decreasing depressed mood in currently and formerly depressed individuals. *Journal of Affective Disorders, 226*, 220–226. https://doi.org/10.1016/j.jad.2017.10.006
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146
First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2016). *SCID-5-CV: Structured clinical interview for DSM-5 disorders: Clinician version*. American Psychiatric Association.
Frost, R. O., & Green, M. L. (1982). Velten mood induction procedure effects. *Personality and Social Psychology Bulletin, 8*(2), 341–347. https://doi.org/10.1177/0146167282082024
Gilbert, P., & Procter, S. (2006). Compassionate mind training for people with high shame and self-criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 13*(6), 353–379. https://doi.org/10.1002/cpp.507
Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature Communications, 12*(1), 861. https://doi.org/10.1038/s41467-020-20649-4
Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry, 23*(1), 56–62. https://doi.org/10.1136/jnnp.23.1.56
Härter, M., Sitta, P., Keller, F., Metzger, R., Wiegand, W., Schell, G., Stieglitz, R.-D., Wolfersdorf, M., Felsenstein, M., & Berger, M. (2004). Stationäre psychiatrisch-psychotherapeutische Depressionsbehandlung Prozess- und Ergebnisqualität anhand eines Modellprojekts in Baden-Württemberg [Psychiatric-psychotherapeutic inpatient treatment for depression. Process and outcome quality based on a model project in Baden-Wurttemberg]. *Nervenarzt, Der, 75*(11), 1083–1091. https://doi.org/10.1007/s00115-004-1705-8
Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research, 4*(4), 383–395. https://doi.org/10.1007/BF01178214
Hönig, F., Batliner, A., Nöth, E., Schnieder, S., & Krajewski, J. (2014). Automatic modelling of depressed speech: Relevant features and relevance of gender. *Proc. Interspeech, 2014*, 1248–1252. https://doi.org/10.21437/interspeech.2014-313
Hughes, J., & Beiner, D.. reghelper: Helper functions for regression analysis. R package version 1.1.0. https://cran.r-project.org/web/packages/reghelper.
James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., Accrombessi, M. M. K., … Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet, 392*(10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7
Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication, 88*, 106–126. https://doi.org/10.1016/j.specom.2017.01.011
Karyotaki, E., Efthimiou, O., Miguel, C., Bermpohl, F. M. G., Furukawa, T. A., Cuijpers, P., Riper, H., Patel, V., Mira, A., Gemmil, A. W., Yeung, A. S., Lange, A., Williams, A. D., Mackinnon, A., Geraedts, A., van Straten, A., Meyer, B., Björkelund, C., Knaevelsrud, C., … Forsell, Y. (2021). Internet-based cognitive behavioral therapy for depression: A systematic review and individual patient data network meta-analysis. *JAMA Psychiatry, 78*(4), 361–371. https://doi.org/10.1001/jamapsychiatry.2020.4364
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
Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual Review of Public Health, 34*, 119–138. https://doi.org/10.1146/annurev-publhealth-031912-114409
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13). https://doi.org/10.18637/jss.v082.i13
Michalak, J., Mischnat, J., & Teismann, T. (2014). Sitting posture makes a difference-embodiment effects on depressive memory bias. *Clinical Psychology & Psychotherapy, 21*(6), 519–524. https://doi.org/10.1002/cpp.1890
Michalak, J., Rohde, K., & Troje, N. F. (2015). How we walk affects what we remember: Gait modifications through biofeedback change negative affective memory bias. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 121–125. https://doi.org/10.1016/j.jbtep.2014.09.004
Michalak, J., Troje, N. F., Fischer, J., Vollmar, P., Heidenreich, T., & Schulte, D. (2009). Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosomatic Medicine, 71*(5), 580–587. https://doi.org/10.1097/PSY.0b013e3181a2515c
Miller, I. W., Bishop, S., Norman, W. H., & Maddever, H. (1985). The modified Hamilton rating scale for depression: Reliability and validity. *Psychiatry Research, 14*(2), 131–142. https://doi.org/10.1016/0165-1781(85)90057-5
Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralts, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics, 20*(1), 50–64. https://doi.org/10.1016/j.jneuroling.2006.04.001
Mundt, J. C., Vogel, A. P., Feltner, D. E., & Lenderking, W. R. (2012). Vocal acoustic biomarkers of depression severity and treatment response. *Biological Psychiatry, 72*(7), 580–587. https://doi.org/10.1016/j.biopsych.2012.03.015
Pampouchidou, A., Simos, P. G., Marias, K., Meriaudeau, F., Yang, F., Pediaditis, M., & Tsiknakis, M. (2019). Automatic assessment of depression based on visual cues: A systematic review. *IEEE Transactions on Affective Computing, 10*(4), 445–470. https://doi.org/10.1109/TAFFC.2017.2724035
Price, T. F., & Harmon-Jones, E. (2015). Embodied emotion: The influence of manipulated facial and bodily states on emotive responses. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(6), 461–473. https://doi.org/10.1002/wcs.1370
Pustejovsky, J. E., & Tipton, E. (2018). Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business & Economic Statistics, 36*(4), 672–683. https://doi.org/10.1080/07350015.2016.1247004
Quatieri, T., & Malyska, N. (2012). Vocal-source biomarkers for depression: A link to psychomotor activity. *Proc. Interspeech, 2012*, 1059–1062. https://doi.org/10.21437/Interspeech.2012-311
Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., … Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141–1152. https://doi.org/10.1111/2041-210X.13434
Schnall, S., & Laird, J. D. (2007). Facing fear: Expression of fear facilitates processing of emotional information. *Social Behavior and Personality: International Journal, 35*(4), 513–524. https://doi.org/10.2224/sbp.2007.35.4.513
Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry, 17*(2), 197–208. https://doi.org/10.1177/1073858410386727
Skinner, B. F. (1965). *Science and human behavior*. Free Press.
Taylor, D. J., Walters, H. M., Vittengl, J. R., Krebaum, S., & Jarrett, R. B. (2010). Which depressive symptoms remain after response to cognitive therapy of depression and predict relapse and recurrence? *Journal of Affective Disorders, 123*(1–3), 181–187. https://doi.org/10.1016/j.jad.2009.08.007
Teasdale, J. D., & Barnard, P. J. (1993). *Affect, cognition, and change: Re-modeling depressive thought*. Psychology Press.
Tomicic, A., Martinez, C., & Krause, M. (2015). The sound of change: A study of the psychotherapeutic process embodied in vocal expression. Laura rice's ideas

revisited. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 25*(2), 263–276. https://doi.org/10.1080/10503307.2014.892647

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy, 6*(4), 473–482. https://doi.org/10.1016/0005-7967(68)90028-4

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioral therapy's effects. *Journal of Consulting and Clinical Psychology, 75*(3), 475–488. https://doi.org/10.1037/0022-006x.75.3.475