The efficacy of a transdiagnostic emotion regulation skills training in the treatment of binge-eating disorder—Results from a randomized controlled trial

Matthias Berking | Eva Eichler | Eva Naumann | Jennifer Svaldi

1University of Erlangen-Nuremberg, Erlangen, Germany
2University of Tuebingen, Tuebingen, Germany

Abstract

Objectives: Deficits in emotion regulation (ER) have been shown to be associated with binge-eating disorder (BED). To further clarify the causal nature of this association, we tested whether systematically enhancing ER skills would reduce symptoms of BED.

Methods: We randomly allocated $N = 101$ individuals meeting the criteria for BED to a transdiagnostic ER skills training or to a waitlist control condition (WLC). Primary outcome was the reduction in binges during the treatment-vs.-waiting period as assessed with the Eating Disorder Examination (EDE) interview.

Results: Mixed-model ANOVAs indicated that the average pre-to-post decrease in binges assessed with the EDE was significantly greater in the ER skills training condition than in the WLC ($d = 0.66$). These effects were stable over the 6-month follow-up period ($d = 0.72$). Remission rates at post/follow-up were 34.4/45.0% in the skills training and 7.5/20.0% in the WLC. Additionally, we found a greater reduction in general eating disorder psychopathology, of food consumption in a bogus taste test and of depression in the ER skills training condition. Moreover, the greater reduction in binge-eating episodes in the training condition was (partially) mediated by a greater increase in ER skills.
**INTRODUCTION**

Binge-eating disorder (BED) is one of the most prevalent eating disorders (Galmiche et al., 2019) and is defined by recurrent binge-eating episodes (BE) in the absence of compensatory behaviour, accompanied by feelings of loss of control and marked distress (American Psychiatric Association, 2013). BED is associated with overweight and obesity, high levels of psychological comorbidity, substantial medical, psychological and psychosocial sequelae, reduced quality of life and an increased mortality risk (Agh et al., 2016; Galmiche et al., 2019; Mitchell, 2016; Wilfley et al., 2016). Fortunately, psychological treatments for BED have been shown to be effective—however, around 20% of patients drop out of treatment and 30% to 50% fail to sustain initial treatment gains in the long-term (Harvey & Gumport, 2015; Hilbert et al., 2019; Vocks et al., 2010; Wilson et al., 2010). Thus, there is a significant need for more effective and accessible treatments for BED.

In the attempt to derive such treatments from evidence-based theories on BED, several authors have argued that BE might be seen as a dysfunctional way of avoiding undesired affective states such as anger/frustration, anxiety, sadness/depressed mood, disappointment, or loneliness (Dingemans et al., 2017; Fairburn et al., 2003; Heatherton & Baumeister, 1991; Kenardy et al., 1996; Peterson et al., 2020; Safer et al., 2009). Evidence for this hypothesis comes from numerous retrospective self-report studies in which individuals suffering from BED reported negative mood to be a significant antecedent of BE (Arnow et al., 1992; Binford et al., 2004; Bruce & Agras, 1992; Masheb & Grilo, 2006; Mitchell et al., 1999; Tachi et al., 2001; Zeeck et al., 2011), from cross-sectional studies showing BE to be associated with perceived deficits in ER (Prefit et al., 2019; Whiteside et al., 2007) and from self-report studies indicating that deficits in ER mediate the association between antecedents of BE and BE (El Archi et al., 2021; Kukk & Akkermann, 2020; Shakory et al., 2015). Further evidence comes from prospective field studies demonstrating that undesired affective states predict subsequent BE (DI et al., 2001; Greeno et al., 2000; Hilbert & Tuschen-Caffier, 2007; Munsch et al., 2012; Stein et al., 2007; Wild et al., 2007) and from ecological momentary assessment (EMA) studies in which arguably dysfunctional emotion regulation (ER) strategies, such as rumination or repetitive negative thinking, predict subsequent binge eating.

**Conclusions:** The findings provide further support for the assumed importance of deficits in ER as a maintaining factor and, hence, as a target in the treatment of BED. As ER skills trainings have been shown to also reduce other kinds of psychopathology, they might be considered a promising transdiagnostic add-on component to disorder-specific interventions.

**KEYWORDS**

binge-eating disorder, emotion regulation, mediation, randomized controlled trial, skills training

**Practitioner points**

- Emotion regulation training reduces binging in BED
- The training also reduces general eating psychopathology and depression
- Effects on binge eating are mediated by changes in emotion regulation skills
- A transdiagnostic intervention can be used to reduce symptoms of BED
(Svaldi et al., 2019) and other eating disorder symptoms (Sala et al., 2019), whereas the use of arguably adaptive ER strategies, such as reappraisal, acceptance or self-compassion, predict subsequent abstinence (Svaldi et al., 2019). Additionally, EMA studies showed an increase in negative affect prior to BE and a decrease on negative affect after BE (Schaefer et al., 2020), which is consistent with the assumption that BE is reinforced by a short-term reduction in negative affect after BE. Other EMA studies found a subsequent increase in negative affect, which is consistent with the assumption that the BE-induced relief from negative emotions is short-lived and instantly followed-up by negative self-evaluations cueing negative affective states (which increase the risk of further BE and so forth (Deaver et al., 2003; Haedt-Matt & Keel, 2011)). Finally, several experimental studies have shown that the induction of negative mood increases subsequent food intake and BE in individuals suffering from BED (Agras & Telch, 1998; Chua et al., 2004; Klatzkin et al., 2018; Laessle & Schulz, 2009; Svaldi et al., 2010). In sum, these findings suggest that binge eating can be conceptualized as a dysfunctional ER strategy that is used in the absence of more adaptive strategies (Dingemans et al., 2017; Leehr et al., 2015). As a lack of adaptive ER skills increases the likelihood of BE being seen as the only available way out of painful emotions (particularly in times of distress when action is guided more strongly by short-term than long-term goals (Tice et al., 2001)), it can be hypothesized that a lack of adaptive ER contributes to the maintenance of BED.

Following this line of reasoning, it can be further hypothesized that BED might be successfully (prevented and) treated with interventions that work to enhance adaptive ER skills. Preliminary evidence for this hypothesis comes from studies demonstrating that psychotherapeutic treatments that aim to foster ER skills significantly reduce BED symptoms (Agras et al., 2017; Clyne & Blampied, 2004; Masson et al., 2013; Peterson et al., 2020). However, all these treatments also target other potentially relevant mechanisms in addition to ER (e.g. meal planning, dysfunctional beliefs on weight/shape issues, or interpersonal skills). Thus, it is presently unclear to what extent the focus on ER is responsible for the effects of these treatments. More specific evidence comes from a study conducted by Svaldi et al. (2014) who found that individuals meeting the criteria for BED consumed less food in a bogus taste test after negative mood induction if they had received a short laboratory-based training in the use of adaptive ER skills (i.e. cognitive reappraisal; as opposed to expressive suppression in the control condition). However, this study did not investigate whether participants continued to use the skills and whether this would lead to a sustained reduction in BED symptoms. Recently, Holmqvist Larsson et al. (2020) reported a significant pre-vs.-post reduction of symptoms in individuals meeting criteria for various eating disorders who had received a group-based ER training in addition to inpatient treatment as usual. Similarly, Glisenti et al. (2021) reported positive effects of emotion-focused therapy on BED. However, because of the lack of (sufficiently powered) control conditions in these studies, it is unclear to what extent these changes were caused by the intervention. Thus, in spite of encouraging preliminary evidence, it is still unclear to what extent systematically enhancing adaptive ER might reduce symptoms of BED.

To clarify this question, we tested the hypothesis that an intervention, which exclusively focuses on enhancing adaptive ER skills would significantly reduce binges in BED. Additionally, we tested whether the training would also reduce general symptoms of eating disorders (e.g. shape and weight concerns), symptoms of depression and food intake after the induction of negative effect. Finally, we aimed to clarify whether the ER skills training would enhance ER skills and whether such an enhancement would mediate the training effect on symptoms of BED.

**METHODS**

**Design, participants and procedure**

We evaluated the effects of an ER skills training on symptoms of BED in a multicenter, prospective, 2-armed, randomized controlled trial (see Figure 1). Participants were recruited among individuals applying for psychotherapeutic treatment at one of the three outpatient treatment centres where the study
was conducted. Additional recruitment measures included advertising in local newspapers and sending study information to health insurance companies and general practitioners in private practices. In all forms of recruitment, participation in a study evaluating a “new treatment for BED” was offered in order to avoid preselection of participants particularly open to interventions addressing ER skills. Participants were offered a financial reimbursement ranging between €25 and €75, depending on the rate of complete data. Potential participants were prescreened by phone and—if still eligible—invited to an in-depth screening at the respective study center. The in-depth screenings included an interview conducted by a senior psychotherapist and structured clinical interviews conducted to further verify diagnoses and assess eating disorder symptom severity (see below). To be included in the study, participants had to meet the following criteria: (a) diagnosis of BED according to DSM-IV criteria, (b) 18 ≤ age ≤ 69, (c) BMI < 45 and (d) ability and willingness to provide informed consent. Patients were excluded if they met any of the following criteria: (a) high risk of suicide, (b) co-occurring psychotic, bipolar disorders, alcohol/substance dependence within the past six months and/or borderline personality disorder, (c) medical disorders that would affect the weight or the ability to participate, (d) insufficient German language skills (as assessment and treatment were conducted in German), (e) current participation in psychotherapeutic treatment other than those offered in the study, (f) current participation in a weight control program, (g) taking medication that would affect weight and (h) pregnancy. Subjects meeting all of the inclusion and none of the exclusion criteria were offered to participate in the study. Consenting subjects met with a study staff member to review study details and to obtain written informed consent. At this stage, potential participants were informed that the study would investigate whether enhancing ER skills would help overcome BED. If eligible and consenting, participants were enrolled in the study and completed the baseline assessment (self-report measures and the bogus taste test as described below). Subsequently, participants were randomly assigned to an intensive ER skills

![CONSORT flowchart of participants](image-url)

**FIGURE 1** CONSORT flowchart of participants. *Note*: BED, binge-eating disorder; ERST, emotion regulation skills training; WLC, wait list control condition; ITT, intention-to-treat
training or to a waitlist control condition (WLC). During the 8-week ER skills training vs. WLC phase, questionnaire-based assessments were scheduled on a weekly basis. Within a time period of 7 days after these 8 weeks, participants completed the post-treatment assessment (identical to the baseline assessment) and, 6 months later, the 6-month follow-up assessment (identical to the baseline assessment except for the bogus taste test, which was no longer included). The project was approved by the local ethics committees (blinded for review) and registered at ClinicalTrials.gov (blinded for review).

Measures

Diagnostics

BED criteria were assessed with the German version of the Eating Disorder Examination interview (EDE) (Fairburn & Cooper, 1993; Hilbert & Tuschen-Caffier, 2006). The EDE is a structured interview that determines the presence or absence of BED criteria with the help of specific questions and specified cut-off criteria (e.g. absence of self-induced vomiting at least once a week for 3 months). Evidence for its validity is provided by several studies (Berg et al., 2012; Hilbert et al., 2004). The EDE interviews were conducted by trained, blinded interviewers holding at least a B.Sc. degree in psychology (these interviewers also conducted the phone-based pre-screening). All interviews were videotaped and 5% of the interviews were rated by a second assessor. A Cohen's Kappa (Cohen, 1960) of .83 indicated a very good inter-rater reliability with regard to the presence vs. absence of a BED diagnosis. To assess comorbid disorders, we used the Structured Clinical Interview for DSM-IV (as this was the current version when the study commenced) Axis I and II Disorders (SCID) (Wittchen et al., 1997) conducted by the blinded and trained interviewers who also conducted the EDE. To evaluate the inter-rater reliability of the SCID interviews, a random sample of 5% of these interviews was videotaped and rated by two assessors. A Cohen's Kappa of 1 indicated hundred-percent accordance with regard to the diagnosis of comorbid disorders.

Binge-eating symptoms

We used the number of binges in the past 28 days, as assessed in the EDE, as a primary outcome of the study. We also analyzed the total score of the EDE as a measure of general eating disorder psychopathology (i.e. restraint eating, shape/eating/weight concerns, in addition to binge eating episodes). Previous studies have reported excellent psychometric properties for this measure with an internal consistency of Cronbach's $\alpha = .93$ (Hilbert & Tuschen-Caffier, 2006). In the present study, Cronbach's $\alpha$ was .87 for the total score.

For further analyses of BED symptom severity, we used the German version of the Eating Disorder Examination Questionnaire (EDE-Q) (Fairburn & Beglin, 1994; Hilbert et al., 2007). The EDE-Q assesses the same symptoms of BED as the EDE. In previous research, the EDE-Q displayed excellent internal consistency for the total score (Cronbach's $\alpha = .95$) (Hilbert et al., 2007). In the present study, we found an internal consistency ($\alpha = .90$) for the total score.

To obtain a standardized behavioural indicator of increased food intake in response to negative affective states cues, we used the bogus taste test (Agras & Telch, 1998; Chua et al., 2004; Laessle & Schulz, 2009). In this paradigm, participants were told that they would help clarify whether negative mood affects food taste, whereas, in reality, the paradigm focuses on assessing the amount of food consumed after a negative mood induction. More specifically, participants watched a sad video clip and were then exposed to four identical pre-weighed bowls filled with high caloric snack foods: cookies (448 kcal/100 g), milk chocolate (501 kcal/100 g), potato chips (410 kcal/100 g) and salted peanuts (627 kcal/100 g). They were left alone for 20 minutes with the instruction to taste the food and to evaluate the taste of each food on a questionnaire. Subsequently, they were asked what they considered
the purpose of the experiment. The amount of food in each bowl was measured before and after the taste test, and the caloric content of the consumed food was calculated after the task was completed. In these analyses, we excluded all participants who had guessed the right purpose of the test at either pre- or post-treatment assessment. Evidence for the validity of this procedure comes from several studies (Dingemans et al., 2009; Robinson et al., 2017)).

Emotion regulation skills

To assess the availability of adaptive responses towards challenging feelings, we used the Emotion Regulation Skills Questionnaire (ERSQ (Berking & Znoj, 2008)). The ERSQ is a 27-item self-report instrument that assesses respondents' adaptive ER skills in the previous week. Items are based on the adaptive coping with emotions model (Berking & Lukas, 2015; Berking & Whitley, 2014) and are clustered in the following nine subscales (each consisting of three items): Awareness, Sensations, Clarity, Understanding, Acceptance, Resilience, Readiness to Confront, Self-Support and Modification. In addition to the subscales, the ERSQ provides an overall indicator of ER skills computed as the average of all 27 items. Good reliability and validity of the ERSQ have been demonstrated in previous studies (Berking et al., 2008; Berking et al., 2010; Berking et al., 2011; Berking et al., 2012; Berking et al., 2013; Berking et al., 2019; Berking & Znoj, 2008; Diedrich et al., 2017; Ebert et al., 2017; Hopfinger et al., 2016; Lukas et al., 2018; Radkovsky et al., 2014; Schwarz et al., 2017; Wirtz et al., 2014). In the present study, the ERSQ showed an internal consistency of Cronbach's $\alpha = .94$ for the total score and a range of Cronbach's $\alpha$ between .74 and .89 for the ERSQ subscales.

Depressive symptoms

For the assessment of depressive symptoms, we used the Beck Depression Inventory-II (BDI-II) (Hautzinger et al., 2006). This commonly used self-report measure contains 21 items assessing somatic, behavioural, emotional and cognitive symptoms of depression and has demonstrated good psychometric properties in various studies (Hautzinger et al., 2006). In the present study, the BDI-II showed an internal consistency of $\alpha = .91$.

Treatment

Emotion regulation skills training

As an intervention that exclusively aims to enhance ER skills, we used the Affect Regulation Training (ART) (Berking, 2008; Berking & Whitley, 2014). ART is a transdiagnostic, group-based training aiming to enhance general ER skills in individuals who meet the criteria for mental disorders or are at risk of developing mental health problems (Berking & Lukas, 2015). The highly standardized training has so far been found to enhance ER skills and reduce symptoms of psychopathology in non-clinical samples (Berking et al., 2010; Lotfi et al., 2020), to improve health in patients suffering from physical symptoms (when used as a complementary component for traditional CBT) (Kleinstäuber et al., 2019), to reduce psychopathological symptoms in a diagnostically heterogeneous sample of inpatients treated for mental health problems (as an add-on-treatment) (Berking et al., 2008), to be effective as an add-on (Berking et al., 2013) and as a stand-alone (Berking et al., 2019) treatment for depression. ART synthesizes elements from cognitive behavioural therapy (CBT), dialectical behavioural therapy (Linehan, 1993), emotion-focused therapy (Greenberg, 2002), mindfulness-based interventions (Kabat-Zinn, 2003), neuro-psychotherapeutic translational approaches (Grawe, 2017), compassion-focused therapy (Gilbert, 2010),
problem-solving therapy (D'Zurilla & Nezu, 2010) and strength-focused interventions (Duckworth et al., 2005; Grawe, 2004). ART starts with a psychoeducational module that provides participants with information about the functions of affective states, about their biological and psychological mechanisms and about their relevance to mental and somatic health. It is explained that short-term affective states are rarely a significant health risk whereas long-lasting negative effect likely leads to unwanted mental and somatic consequences. Subsequently, seven ER skills are derived from this model, each aiming to help disengage from processes contributing to the maintenance of negative affective states: (1) muscle relaxation, (2) breathing relaxation, (3) non-judgmental perception and description of one's feelings, (4) acceptance and tolerance towards undesired feelings, (5) compassionate self-support when confronted with undesired feelings, (6) analysis of antecedents of undesired feelings and (7) active modification of one's feelings. To strengthen these skills, patients are then taught a specific set of skill-building exercises and strongly encouraged to develop their own daily training regimen to practice these exercises for at least 20 minutes per day. To facilitate the daily training, participants are provided with audio-based instructions guiding them through the exercise presently focused upon in the group. Additionally, participants can choose to receive text messages that suggest 140 short exercises throughout the training period. Following the guidelines for the intensive ART format, we scheduled one 180-minute session per week in training groups including four participants for a period of six consecutive weeks followed by another 2 weeks of independent skills practice (with one 90-minute booster session in the 8th week). This format resulted from extended (unpublished) feasibility trials and was considered an acceptable compromise between the attempt to minimize the chance that group members miss a training session (which increases proportionally to the total number of group-based sessions) and the attempt to prevent overtaxing participants' concentration spans in overly long sessions. The standard ART manual was modified only with regard to the introductory session in which the ART trainers presented the idea that BE is often used as a way to avoid unwanted feelings and followed up with a discussion on whether participants shared this assumption. The implementation of ART followed standardized treatment protocols (Berking, 2008; Berking & Whitley, 2014) and was conducted by doctoral candidates in clinical psychology who had completed or were in advanced stages of their clinical training. All therapists were intensely trained in ART and received weekly supervision from experienced ART trainers. To control adherence, a random sample of 5% of all sessions was videotaped and rated with regard to adherence to the training protocol. In case participants still met the criteria for BED (or any other mental disorder), they were offered individual CBT after completing the last assessment.

Waitlist control condition

Participants in the WLC received no treatment within the study but were offered to participate in ART or commence individual CBT after completing all assessments.

Statistical analyses

To compare differences between the ER skills training and WLC with regard to the reduction in BE as a primary outcome, we used mixed-model ANOVAs with time x condition as the independent and changes in the number of BE per week (as assessed in the EDE) from pre- to post-treatment as a dependent variable. In these analyses, we estimated missing values with the help of multiple imputations as suggested by Twisk and de Vente (Twisk & de Vente, 2002) and used the intention-to-treat (ITT) sample for our primary level of analyses. To evaluate the sensitivity of findings towards the effects of missing data, we also conducted completer analyses. To assess potential mid- to long-term effects, we used the same ANOVA-based method to compare changes in BE between baseline and the 6-month follow-up. The same approach was used to compare conditions with regard to general eating disorder
psychopathology (as assessed with the EDE) and depression (assessed with the BDI-II) at baseline, post-treatment and follow-up, and for testing differences regarding pre-post changes in caloric intake in the bogus taste test. To test differences between the ER skills training and WLC regarding remission rates at post and follow-up assessment, we used X (American Psychiatric Association, 2013)-tests. To test differences between conditions with regard to weekly self-reports of binge eating, general eating disorder psychopathology and adaptive responses to undesired feelings, we used multilevel model (MLM) analyses testing the effects of the treatment condition (ER skills training vs. WLC; coded as a level-2, dummy-coded variable) on the slopes of the total scores of the EDE-Q and the ERSQ, respectively.

To further clarify whether enhancing adaptive responses to undesired feelings is a significant mechanism responsible for the potential effects of ART on BE, we tested whether changes in ER would mediate potential differences in the rate of change of BE between ART and WLC. For this purpose, we used a path model in which the treatment condition is assumed to predict ER skills at follow-up (controlled for pre-treatment scores), which is assumed to predict BE at follow-up (controlled for pre-treatment scores). Following procedures suggested by Preacher (Preacher, 2015), we tested the indirect effect in this model with the help of bias-corrected bootstrapped 95% confidence intervals (number of bootstrap draws = 1000). In all analyses, we used one-sided tests for testing directed hypothesis and two-sided tests for testing non-directed hypotheses. This decision was based on previous evidence provided by the large number of studies showing that active bona fide psychological interventions are superior to waitlist control conditions (de Jong & DeRubeis, 2018) and previous research on ART, which showed consistently that ART was superior to WLCs and other control conditions (Berking et al., 2008; Berking et al., 2010; Berking et al., 2013; Berking et al., 2019).

All computations except for the MLMs were performed in IBM SPSS Statistics 26. MLMs were performed in R 3.1.3 (R Core Team, 2018), using the package lme4 (Bates et al., 2014). Simple slopes, model comparisons using X^2-tests, confidence intervals and p-values were computed with the packages lsmeans (Lenth, 2016), car (Fox & Weisberg, 2011) and lmerTest (Kuznetsova et al., 2017), which use Satterthwaite-adjusted degrees of freedom. Mediation models were computed using the SPSS package PROCESS (Hayes, 2012). We powered the study to be able to detect small-to-moderate differences between the conditions (d = 0.5) with β = .80 and critical alpha set to 5%.

RESULTS

Recruitment and patient characteristics

As shown in Figure 1, out of 390 screened individuals, 101 met the study criteria and were randomly allocated to the treatment or the WLC. Over the course of the study, serious difficulties occurred with regard to recruiting participants as quickly as necessary to avoid drop-outs due to participants’ unwillingness to wait until eight individuals were recruited for randomization. To solve the problem, we switched from the original block randomization to a procedure in which the next patient eligible for the study was allocated to the treatment condition until the minimum group size of four was reached whether this was necessary to prevent participants from dropping out of the study. Because assessors were blind to this decision and because this decision was made prior to the baseline assessment, allocation depended exclusively on the rank of the participant in the recruitment order and was therefore as random as it would have been if the combination of the rank in the recruitment order and the block-randomized indicator of study condition had been the basis for the allocation. Rate of study completers was 87% in the ER skills training and 83% in the WLC. Sociodemographics and information on diagnoses are reported in Table 1. Mean and standard deviation for all outcomes and time points are reported in Tables 2, 3 and 4. The conditions did not differ significantly regarding sociodemographics, comorbidity and baseline levels of any of the outcome variables.
Consistent with our primary hypothesis, there was a significant decrease in binge eating episodes over the last 28 days as assessed with the EDE (primary outcome) in the ER skills training condition (pre: 16.29 vs. post: 3.61; t = 9.77; p < .001; d = 1.35). More importantly, a significant time x condition interaction in the mixed model ANOVA showed that the reduction in interview-assessed BE (primary outcome) during the ER skills training vs. WLC period was significantly stronger in the training condition than in the WLC (F(1, 99) = 10.69; p < .001; d = 0.66). A significant interaction for the pre- vs. follow-up assessment (F(1, 99) = 12.72; p < .001; d = 0.72) provides evidence for significant mid-to-long-term effects of the training (see Figure 2).

With regard to the percentage of patients remitted (zero BE during the last 28 days in the EDE), we found significant differences between conditions with 34.4% remitted patients in the ER skills training and 7.5% remitted patients in WLC at post-treatment (χ² (1) = 10.11, p < .001).
TABLE 2  Mean, standard deviation and ANOVA test statistics for EDE, bogus taste test and BDI-II at pre-treatment, post-treatment and 6-month follow-up

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Pre</th>
<th>Post</th>
<th>Follow-up</th>
<th>Pre vs. post</th>
<th>Pre vs. follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M (SD)</td>
<td>T</td>
<td>df</td>
<td>p</td>
<td>M (SD)</td>
</tr>
<tr>
<td><strong>BE_{EDE}</strong></td>
<td>ERST</td>
<td>16.29 (11.30)</td>
<td>1.4</td>
<td>97</td>
<td>.18</td>
<td>3.61 (4.76)</td>
</tr>
<tr>
<td></td>
<td>WLC</td>
<td>13.42 (7.82)</td>
<td></td>
<td></td>
<td></td>
<td>7.17 (6.15)</td>
</tr>
<tr>
<td><strong>EDE_{total}</strong></td>
<td>ERST</td>
<td>2.66 (1.05)</td>
<td>0.97</td>
<td>97</td>
<td>.33</td>
<td>1.66 (0.94)</td>
</tr>
<tr>
<td></td>
<td>WLC</td>
<td>2.47 (0.86)</td>
<td></td>
<td></td>
<td></td>
<td>2.14 (0.81)</td>
</tr>
<tr>
<td><strong>Caloric intake</strong></td>
<td>ERST</td>
<td>315.49 (449.85)</td>
<td>0.63</td>
<td>71</td>
<td>.51</td>
<td>298.46 (197.07)</td>
</tr>
<tr>
<td></td>
<td>WLC</td>
<td>204.37 (106.75)</td>
<td></td>
<td></td>
<td></td>
<td>465.88 (612.35)</td>
</tr>
<tr>
<td><strong>BDI-II</strong></td>
<td>ERST</td>
<td>19.20 (10.23)</td>
<td>1.38</td>
<td>97</td>
<td>.17</td>
<td>11.16 (9.15)</td>
</tr>
<tr>
<td></td>
<td>WLC</td>
<td>16.13 (11.59)</td>
<td></td>
<td></td>
<td></td>
<td>12.94 (10.64)</td>
</tr>
</tbody>
</table>

BE, binge eating episodes; BDI-II, Beck Depression Inventory-II; EDE, Eating Disorder Examination; ERST, emotion regulation skills training; WLC, wait list control condition.

*We excluded 20 participants who had guessed the true purpose of the experiment.

*p < .10; *p < .05; **p < .01; ***p < .001.
difference was still significant after the 6-month follow-up period (ERST: 45.9%, WLC: 20.0%, \(\chi^2(1) = 7.89, p = .003\)). Similarly, we found a significant reduction in BE as assessed by self-reports in the ER skills training condition (EDE-Q; pre: 13.16 vs. post: 5.31; \(t(60) = 8.18; p < .001; d = 1.04\)). Consistently, we found a significant time x condition interaction indicating a significantly stronger decrease of self-reported binges in the skills training condition than in the WLC (\(\beta = 0.04, t[558.19] = 1.76, p = .04\)). However, this effect was no longer significant in the pre-vs.-follow-up comparison (\(F(1, 99) = 1.84, p = .09\)).

### General eating disorder psychopathology

As anticipated, the EDE interview total score showed a significant reduction during the ER skills training (pre: 2.65 vs. post: 1.65; \(t(60) = 7.95, p < .001, d = 1.03\)). Again more importantly, we found a significantly stronger decrease of the EDE interview total score in the ER skills training than in the WLC from pre- to post-treatment (\(F(1, 99) = 12.51, p < .001, d = 0.71\)) and from pretreatment to follow-up (\(F(1, 99) = 12.81, p < .001, d = 0.72\)). Consistently, we found a significant time x condition interaction indicating a significantly stronger decrease of the self-reported EDE total score in the ER skills training than in the WLC (\(\beta = 0.09, t[561.29] = 5.39, p < .001\)). This effect was also significant in the pretreatment-vs.-follow-up comparison (\(F(1, 99) = 7.44, p = .004, d = 0.56\)). To explore differential effects on specific domains of eating psychopathology, we also compared changes in EDE and EDE-Q subscales between conditions (see Table 4 for results). For the analyses of the caloric intake in the bogus taste test the mixed-model ANOVA showed a significant time x condition interaction (\(F(1, 55) = 4.34, p = .02, d = 0.56\)) with a stable caloric intake in the skills training condition and an increase in the WLC between the pre- and the post-assessment.

### Emotion regulation skills and depression

Further MLM analyses showed that there was a significantly greater increase in self-reports on the availability of adaptive ER skills in the ER skills training than in the WLC (ERSQ\_total; \(\beta = -0.13, p < .001\)).
EMOTION REGULATION SKILLS TRAINING IN BED

\[ t[558.72] = -5.86, p < .001 \]. This effect was also significant in the pretreatment-vs.-follow-up comparison (\( F(1, 97) = 10.19, p < .001, d = 0.65 \)). With regard to the subscales of the ERSQ, the greatest differences were found for **Acceptance** (\( \beta = -0.14, t[558.23] = -6.39, p < .001 \)), **Understanding** (\( \beta = -0.13, t[558.72] = -5.86, p < .001 \)) and **Tolerance** (\( \beta = -0.11, t[558.34] = -5.28, p < .001 \)). Participants in the ER skills training condition reported a significantly greater reduction in depressive symptoms as assessed with the BDI-II (pre-vs.-post: \( F(1, 97) = 7.19, p = .005 \); pre-vs.-follow-up: \( F(1, 97) = 5.43, p = .01 \)) than did participants in the WLC.

**Mediation analysis**

The results of the mediation analyses are illustrated in Figure 3. In line with the analyses reported above, group status was associated with BE and the EDE total score controlled for pretreatment scores. Additionally, there was a significant association between the study condition and the ERSQ total score. Group status was also associated with the ERSQ total score. Finally, ERSQ scores were significantly associated with both BE and the EDE total score. With regard to the final criterion for mediation, in both models the indirect effect was significant as the 95% CIs of the bootstrap analyses did not contain zero (BE: \( \beta = 0.95 \); 95% CI: 0.07–3.00; EDE_total: \( \beta = 0.19 \), 95% CI = 0.06–0.42). This evidence for a mediation effect suggests that the ER skills training affects BE symptoms and general eating disorder psychopathology by enhancing participants’ ER skills. As the direct path was still significant when the indirect path was controlled for, the results indicate partial mediation (see Figure 2).

**Additional analyses**

We conducted complementary completer analyses for all outcomes, which differed only slightly in \( p \)-values but led to the same conclusions regarding changes in dependent variables (e.g. number of BE pre-post ER skills training-vs.-WLC: \( F(1, 75) = 5.93, p = .01 \); remission rates at post-treatment: ER skills training 40%, WLC 7.1%). To test the sensitivity of our primary ANOVA approach against deviations from the normal distribution, we also computed pre-vs.-post and pre-vs.-follow-up changes for each participant for the primary outcome and compared differences regarding these changes between conditions.
with the Kruskal-Wallis test (Kruskal, 1952). The results are consistent with those found in the primary analytical approach (EDE binge eating episodes pre-vs.-post: \( H(1) = 8.98, p = .002 \); EDE binge eating episodes pre-vs.-follow-up: \( H(1) = 10.93, p < .001 \)). Finally, according to our original power analyses, we sought to recruit a total sample size of \( N = 102 \) to ensure a power of at least \( \beta = .80 \) with a critical alpha of 5% on the basis of an anticipated effect size of \( d = 0.5 \) for the primary outcome. Empirically, the resulting effect size was larger than the one the power calculation was based upon (\( d = 0.66 \) vs. 0.5). Thus, the actual power was \( \beta = .94 \).

**DISCUSSION**

The primary aim of the present study was to evaluate the efficacy of an intervention that exclusively focuses on enhancing ER skills with regard to reducing binge eating in individuals suffering from BED. To attain this aim, we compared the rate of symptom reduction between transdiagnostic, group-based ER skills training with a WLC in a sample of \( N = 101 \) participants meeting the criteria for BED. Results indicate that participating in the ER skills training was associated with a greater reduction in binges as assessed in the EDE interviews, of general eating pathology, of food consumption in a bogus taste test and of depression than was participating in the WLC. Results also indicate that differences across conditions regarding BED symptom reduction were partially mediated by changes in ER skills. As findings from the study show that systematically varying ER skills with the help of a specific intervention is associated with the assumed changes in BE (and other symptoms of BED), we consider the results as further evidence for the hypothesis that deficits in ER are a significant maintaining factor in BED.

From a practical perspective, it is of note that the pre-vs.-post/follow-up effect sizes for the number of BE episodes (\( d = 0.66/0.71 \)), for general eating psychopathology (\( d = 0.71/0.72 \)) and for the remission rate at follow-up (46%) all fell within the 95% confidence intervals of respective effect sizes reported for comparisons between disorder-specific BED treatments and WLCs in a recent meta-analyses conducted by Hilbert et al. (2019) (average across 12 studies: \( d_{\text{binge eating}} = 0.83 \), CI = 0.45–1.20; rate of remission: 53%, CI = 45–61, \( d_{\text{general eating psychopathology}} = 0.65 \), CI = 0.37–0.93). If these findings should be replicated in future studies, they would suggest that transdiagnostic ER training might not be significantly inferior towards disorder-specific treatments of BED and that, hence, such transdiagnostic approaches might be considered whenever they offer significant advantages over disorder-specific approaches (e.g., as they are easier to deliver in group-based formats).

However, before such a claim can be made, future research will need to overcome the limitations of the present study. For example, the present study compared the ER skills training exclusively with a WLC. Such a comparison does not control for unspecific effects associated with a therapeutic alliance, nor for a rational explaining of symptoms and pointing out ways to overcome the disorder, nor for positive expectations. Thus, the study does not allow final conclusions about changes in ER being the pivotal mechanism of change. Hence, there is still a need for a study in which the ER skills training is tested against a condition providing comparable social support, treatment rationale and expectancies. Ideally, such a condition would not involve other active mechanisms of change but exclusively control unspecific factors.

From a practical perspective, the lack of control for unspecific factors might not be considered a significant limitation—as long as the intervention demonstrates clinically significant effects. Instead, from this perspective, it would be more important to compare transdiagnostic ER training with present gold standard treatment alternatives. Potential candidates for such a comparison include individual disorder-specific CBT (Hilbert et al., 2019), group-based disorder-specific CBT (Hilbert et al., 2019), group-based and individual interpersonal therapy (Miniati et al., 2018) and non-disorder-specific treatments that focus on transdiagnostic factors other than ER—such as problem-solving trainings (Vanderlinden et al., 2007), social skills trainings (Duchesne et al., 2012), acceptance and commitment therapy (Hill et al., 2015), mindfulness-based approaches (Katterman et al., 2014), or self-control trainings (Peterson et al., 2020). In the present study, we did not compare the ER skills training with any of these treatments.
### Table 4
Mean, standard deviation and ANOVA test statistics for EDE- and EDE-Q subscales at pre-treatment, post-treatment and 6-month follow-up

<table>
<thead>
<tr>
<th>Variable and group</th>
<th>Pre</th>
<th>Post</th>
<th>Follow-up</th>
<th>ANOVA pre vs. post</th>
<th>ANOVA pre vs. follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EDE</td>
<td>EDE-Q</td>
<td>EDE</td>
<td>EDE-Q</td>
<td>EDE</td>
</tr>
<tr>
<td>Eating concern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERST</td>
<td>1.74 (1.28)</td>
<td>2.78 (1.46)</td>
<td>0.81 (0.99)</td>
<td>1.54 (1.60)</td>
<td>0.65 (0.85)</td>
</tr>
<tr>
<td>WLC</td>
<td>1.42 (1.24)</td>
<td>2.51 (1.27)</td>
<td>1.35 (1.00)</td>
<td>2.13 (1.37)</td>
<td>1.14 (1.10)</td>
</tr>
<tr>
<td>Weight concern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERST</td>
<td>3.24 (1.17)</td>
<td>3.76 (1.25)</td>
<td>2.24 (1.13)</td>
<td>2.73 (1.45)</td>
<td>1.84 (1.18)</td>
</tr>
<tr>
<td>WLC</td>
<td>3.27 (1.07)</td>
<td>3.78 (1.05)</td>
<td>2.72 (0.89)</td>
<td>3.28 (1.19)</td>
<td>2.64 (1.24)</td>
</tr>
<tr>
<td>Shape concern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERST</td>
<td>3.49 (1.28)</td>
<td>4.22 (1.44)</td>
<td>2.38 (1.25)</td>
<td>3.07 (1.68)</td>
<td>2.00 (1.31)</td>
</tr>
<tr>
<td>WLC</td>
<td>3.31 (0.93)</td>
<td>4.10 (1.05)</td>
<td>2.88 (0.99)</td>
<td>3.71 (1.26)</td>
<td>2.70 (1.16)</td>
</tr>
<tr>
<td>Restraint eating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERST</td>
<td>1.93 (1.38)</td>
<td>2.49 (1.47)</td>
<td>1.43 (1.37)</td>
<td>1.58 (1.60)</td>
<td>1.38 (1.35)</td>
</tr>
<tr>
<td>WLC</td>
<td>1.79 (1.23)</td>
<td>2.09 (1.33)</td>
<td>1.88 (1.23)</td>
<td>2.10 (1.36)</td>
<td>1.60 (1.03)</td>
</tr>
</tbody>
</table>

Abbreviations: ERST, emotion regulation skills training; EDE, Eating Disorder Examination; EDE-Q, Eating Disorder Examination Questionnaire; WLC, wait list control condition.

*+p < .10; *p < .05; **p < .01; ***p < .001.
as (to ensure sufficient power) such an approach would have meant to test the training against only one of various equally important benchmarks. Instead, we used the comparison with a WLC to generate an effect size that can be compared with the significant number of studies in which bona fide treatments for BED were compared with WLC. However, now that there is evidence that the effects of ER skills training fall within the range of what is commonly achieved with gold standard treatments, direct comparisons with the most promising of these treatments should be conducted.

Other limitations of the present study result from the fact that about 13.86% of the recruited sample were male although approximately 40% of BED cases in the community are male (Swanson et al., 2011), that we had excluded individuals meeting the criteria for several co-occurring disorders/conditions (see exclusion criteria) and that only individuals open for psychological interventions...
participants in the study. Thus, the findings should not be carelessly generalized to the entire population of individuals suffering from BED. A third limitation of the present study is the lack of a valid assessment of the frequency and duration of participants’ engagement in their independent skills training. Thus, future studies are needed to clarify to what extent the intensity of the independent skills training is associated with treatment outcomes. Future studies should also vary the extent to which ER skills training utilize cognitive, mindfulness-based, or compassion-based techniques in order to clarify which of these foci are most effective for enhancing ER skills in individuals suffering from mental disorders.

Additionally, it is of note that in the present study participants in the WLC reported a slight decrease in symptoms after they had been included in the study. Given that this was also found in several other WLC-controlled studies (Stunkard & Allison, 2003), this could point to the importance of unspecific factors such as receiving social support from clinical experts during the diagnostic measures (Flückiger et al., 2020), increased hope and remoralization because of the anticipated start of treatment (Constantino et al., 2019; Constantino et al., 2021; Thiruchselvam et al., 2019), or enhanced self-reflection prompted by the diagnostic measures (Truijens et al., 2021). It is also of note that self-reports on BE at follow-up failed to reach significant alpha although there were significant group differences for BE as assessed with the EDE interview. Potentially, this indicates that ER skills training have stronger long-term effects on BE as operationalized in clinical interviews than on subjective BE as commonly referred to by participants. As neither the active ingredients of the WLC nor reasons for the discrepancy between interview- and self-report-based assessment of BE can be clarified in the present study, they should be further investigated in future research. Finally, it is of note that significant evidence for the importance of deficits in ER for the development and maintenance of the disorder does not only exist for BED, but also for various other forms of psychopathology, such as other eating disorders (Racine & Wildes, 2015), depression (Radkovsky et al., 2014), anxiety disorder (Lukas et al., 2018), obsessive-compulsive disorder (Yap et al., 2018), substance use disorders (Berking et al., 2011), personality disorders (Fitzpatrick et al., 2020) and even for psychotic disorders (Ludwig et al., 2019). The ER skills training found to be effective for BED in the present study has also been shown to be effective in diagnostically heterogenous inpatients (Berking et al., 2008), depressed inpatients (Berking et al., 2013) and depressed outpatients (Berking et al., 2019). Furthermore, it has displayed promise in the reduction in anxiety (Lotfi et al., 2020) and the reduction in medically unexplained symptoms (Kleinstäuber et al., 2019). Thus, there is reason to believe that transdiagnostic ER skills training might indeed be an effective treatment for several disorders. Future studies should continue to test this hypothesis and evaluate training focusing exclusively on ER skills with regard to their efficacy in the treatment of a broad range of mental disorders.

AUTHOR CONTRIBUTIONS
Matthias Berking: Conceptualization; investigation; methodology; resources; software; supervision; validation; visualization. Eva Eichler: Methodology; visualization. Eva Naumann: Investigation. Jennifer Svaldi: Conceptualization; investigation; methodology; resources.

ACKNOWLEDGEMENT
This work was supported by the German Research Foundation with grant BE 4510/5-1 to MB and grant SV 83/2-1 to JS. Other than funding the study, the German Research Foundation was not involved in any part of the research described in the manuscript.

CONFLICT OF INTEREST
All authors declare that they have no conflict of interest.
OPEN RESEARCH BADGES

This article has been awarded Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. Data is available at Open Science Framework

This article has been awarded Open Materials Badge for making the components of the research methodology needed to reproduce the reported procedure and analysis publicly available at Open Science Framework

This article has a preregistered research design available at Open Science Framework

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES


---

**How to cite this article:** Berking, M., Eichler, E., Naumann, E. & Svaldi, J. (2022). The efficacy of a transdiagnostic emotion regulation skills training in the treatment of binge-eating disorder—Results from a randomized controlled trial. *British Journal of Clinical Psychology, 61*, 998–1018. [https://doi.org/10.1111/bjc.12371](https://doi.org/10.1111/bjc.12371)