JOURNAL OF CLINICAL CHILD & ADOLESCENT PSYCHOLOGY
https://doi.org/10.1080/15374416.2022.2079130

Efficacy of Emotion-Focused Parenting Programs for Children's Internalizing and Externalizing Symptoms: A Randomized Clinical Study

Nadia Ansar[a,b], Helene A. Nissen Lie[a], Rune Zahl-Olsen[c], Thomas B. Bertelsen[c], Robert Elliott[d], and Jan Reidar Stiegler[b]

[a]Department of Clinical Psychology, University of Oslo; [b]Department of Research and Development, The Norwegian Institute of Emotion-Focused Therapy; [c]Department of Child and Adolescent Mental Health, Sorlandet Hospital; [d]School of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK

ABSTRACT

Objective: Emotion-Focused Skills Training (EFST) is a 12-week parental program based on Emotion-Focused Therapy, developed to improve children and adolescents' mental health problems.

Methods: In a randomized clinical dismantling study, including parents of 236 children and adolescents (ages 6–13, M_{age} 8.9, 60.6% boys, 95.8% Caucasian) with externalizing and/or internalizing problems within clinical range, we examined the efficacy of two versions of EFST: one *experiential* condition (n = 120) involving emotionally evocative techniques and two-chair interventions, and one *psychoeducational only* condition (n = 116) involving didactic teaching of emotion skills. Both groups received a 2-day group training and 6 hours of individual supervision. Outcomes were parent- and teacher-reported symptoms at baseline, posttreatment, and 4-, 8-, and 12-month follow-up. Analyses were conducted using multilevel growth curve modeling and Bayesian post hoc analysis.

Results: EFST showed efficacy in reducing parent-reported externalizing ($b = -1.72$, $p < .001$, $d = 1.0$) and internalizing ($b = -1.71$, $p < .001$, $d = 0.9$) symptoms, and teacher-reported externalizing ($b = -.96$, $p < .001$, $d = 0.4$), but not internalizing ($b = -.13$, $p > .05$, $d = 0.2$) symptoms. Multilevel analyses showed nonsignificant differences between conditions (all p's > .05), although a Bayesian longitudinal sensitivity analysis indicated a better outcome for the *experiential* condition.

Conclusion: EFST showed efficacy in symptom reduction for children and adolescents with internalizing and externalizing symptoms. Outcomes were maintained over 12 months for both conditions, supporting EFST as a transdiagnostic parental approach for early intervention.

Data Transparency Statement

The data reported in this manuscript have not been previously published but are part of a larger data collection. A qualitative study, with a reflexive thematic analytic approach, has been conducted based on 15 parents who participated in this study. The manuscript (Ansar et al., 2021) focused on exploring parents' experiences on receiving the program, their perceived changes in everyday life, and their views on the effect of the program on their child. A secondary outcome study is in progress, examining whether parents' own retention of the program mediates the magnitude of program effects on children and adolescents.

Public Health Significance

This study suggests that the mental health symptoms of children and adolescents can be reduced with a manualized, transdiagnostic, easy-to-implement parental intervention based on emotion-focused principles. Since the interventions were administered to parents, they may contribute to less stress and stigma for children and adolescents. Additionally, the data showed significant reductions in child symptoms after 12 months.

Children's mental health problems are recognized as a high priority intervention target due to their high cost for society and those affected (World Health Organization[(WHO], 2005). Externalizing and internalizing problems in childhood and adolescence are widespread, increasing, and have long-term implications for mental health in adulthood (Catalano & Kellogg, 2020; Fanti & Henrich, 2010).

Even though evidence-based treatment interventions are available to clinicians, they are not widely used (Higa & Chorpita, 2008). Such protocols are typically designed for single disorder-specific treatments and may fail to address the reality in everyday clinical practice where

CONTACT Nadia Ansar ansar@ipr.no University of Oslo, Forskningsveien 3A, Oslo 0373
Supplemental material for this article can be accessed online at https://doi.org/10.1080/15374416.2022.2079130

multiple comorbid diagnoses are present (B.F. Chorpita et al., 2013). Recent studies have revealed that internalizing and externalizing problems tend to co-occur in primary school years (Wang & Liu, 2021). Due to the commonalities in etiology, latent structure, and phenomenology, research on child mental health difficulties suggests that diagnostic comorbidity is the rule rather than the exception (Weisz et al., 2017). Hence, a growing attention toward common core treatment elements (Garland et al., 2008) has led to the call for transdiagnostic approaches applicable to the prevention of several related disorders (Barlow et al., 2010). Difficulties in emotion regulation are commonly recognized as a transdiagnostic component in the development and maintenance of a broad range of internalizing and externalizing problems (Compas et al., 2017). Thus, a core component for the development of psychopathology may be rooted in maladaptive emotion regulation strategies, and a broader focus on emotion dysregulation might be beneficial for a wide range of mental health disorders (Ehrenreich et al., 2018).

Several treatments are being tested for their applicability for a variety of child disorders, i.e., cognitive and behavioral oriented parent training, such as MATCH-ADTC (B.F. Chorpita et al., 2013) and the abbreviated version (FIRST; Cho et al., 2021), and social learning-theory-based Unified Protocols for transdiagnostic treatment of emotional disorders (UP-C/A; Ehrenreich et al., 2018). These interventions are predominantly delivered to the child in combination with parental feedback. However, only half of children fulfilling criteria for mental disorders receive mental health care (Whitney & Peterson, 2019), and about half of those offered treatment, drop out (De Haan et al., 2013). Considering the extraordinarily high prevalence of emotional disorders, this discrepancy between prevalence and access to treatment is of great concern (Ehrenreich et al., 2018).

Arguably, the least intrusive method for alleviating these difficulties is for parents to learn new skills by participating in parenting programs (Buchanan-Pascall et al., 2018). As parents play a crucial role in the mental health of developing child, they are a good target for prevention and early intervention (Morris et al., 2017). Additionally, emerging evidence supports that family-based interventions are successful in improving parenting, and in turn, in a range of long-term outcomes for children (Buchanan-Pascall et al., 2018). A common assumption is that children's affective system is developed and matured in interaction with significant others (Mackler et al., 2015). When parents become emotionally dysregulated (under- or overregulated), they lose their ability to respond adequately to their child's needs (Ford et al., 2018). Thus, working with emotion dysregulation and avoidance of intense emotional experience in parents may contribute to better mental health outcomes in children (Schore, 1999).

Emotion-Focused Skills Training (EFST) for parents is a recently developed, intensive, skills-oriented program that aims to increase parents' capacity to respond adaptively to the child's emotions as well as working with parents' own emotional understanding and expression (Dolhanty & Greenberg, 2009; Greenberg, 2015; A. H. V. Hagen et al., 2019). EFST is based on Emotion-Focused Family Therapy (EFFT), which was originally developed for treating eating disorders in children. EFFT focus mainly on parental self-efficacy and reducing emotional impasses between parents and their children (Dolhanty & Greenberg, 2009; Robinson et al., 2015). While maintaining these important elements, EFST was further developed as a transdiagnostic approach to address a broader range of mental health difficulties. EFST focuses on helping parents at dealing with their own and their children's emotions, repairing relationship ruptures between parent and child and on helping parents set sound boundaries for their child. EFST is grounded in emotion theory and research (Damasio, 2000; Izard, 2002; LeDoux, 1993; Tomkins, 1963), Emotion-Focused Therapy for couples and individuals (EFT; Greenberg et al., 1997 and EFT-C; Goldman et al., 2006), emotion coaching programs (Gottman et al., 1997; Havighurst et al., 2013; Katz et al., 2012) and client-centered humanistic psychotherapy (Cain, 2010; Elliott et al., 2004).

In EFST, evocative experiential techniques, such as imaginative two-chair dialogs, evocative empathy, and interventions focused on connecting with emotions, are utilized to help parents facilitate and engage in deeper processing of unpleasant emotions, such as shame, fear, anger, or sadness (Dolhanty & Greenberg, 2009). The underlying assumption is that in order to regulate emotion, one needs to experience being in and enduring the emotional pain, and evocative interventions are assumed to facilitate such experience (Greenberg et al., 1997). Research on EFT suggests that the use of these evocative interventions aids emotional processing and adds to the effect of psychotherapy (Goldman et al., 2006; Greenberg et al., 2008; Stiegler et al., 2018). Although it is assumed that this also applies to EFST, studies have not yet explored whether adding evocative techniques enhances the treatment effect of parenting programs, as compared to a psychoeducational format based on the same theoretical principles. There are currently no efficacy studies in EFST, but there have been some studies in EFFT that have many similarities with EFST. Only one previous study has examined the effect of EFFT on child symptom reduction (Foroughe et al., 2019). In this

process-outcome study, 124 parents received a 2-day intensive intervention and results showed significant reduction in child mental health symptoms. Additionally, one Norwegian pilot study found that children's oppositional defiant problems significantly decreased after a 2-day program in EFFT (Wilhelmsen et al., 2019).

This randomized clinical dismantling study compared two different versions of EFST: one containing both didactic psychoeducation and explicit facilitation of emotion activation, and the other comprising entirely of a didactic psychoeducational format without the explicit use of experiential techniques (i.e., facilitation of emotion activation). Both conditions contained instruction in the basic principles of EFST, which aim to enhance parental emotion understanding, and involves training in four core parental skills: *validating emotion; working with emotional difficulties; relationship repair; and boundary setting* (see below for a more detailed description). This study investigated whether emotionally evocative techniques were particularly helpful in reducing internalizing and externalizing symptoms in children, as reported by both parents and teachers, when delivering EFST. Our primary outcome was parent- and teacher-rated internalizing and/or externalizing child problems. Based on prior research and theory, the following hypotheses were posed: (1) Both versions of EFST are effective in alleviating symptomatic distress (externalizing and internalizing symptoms) in children; and (2) The condition containing evocative experiential components will outperform the psychoeducational-only condition. Due to ethical reasons involving severity of child symptoms, a waiting control group was not used to examine these hypotheses.

Method

Study Design and Procedure

The treatment was administrated at two different outpatient mental health clinics in Norway (Bergen and Oslo). Participants were recruited through advertisements and clinical referrals offering a free of charge program for parents of children aged 6–13. Caregivers who responded to the advertisement were screened by phone by study staff. Inclusion criteria were children's externalizing (EXT) and/or internalizing (INT) difficulties within clinical range, estimated using Brief Problem Monitor for parents (BPM-P; Achenbach et al., 2011) with scores at or above 1.5 *SD* on INT and/or EXT (*T* score ≥ 65), defined as high enough to be considered as dysfunctional, thereby meeting the inclusion criteria (Achenbach et al., 2011). All outcomes were completed by primary caregivers (i.e., the caregiver who spent the most time with the child). Exclusion criteria were serious physical, mental or chronic difficulties in the parent or child, or other ongoing psychotherapy.

The study was conducted as a randomized clinical trial that compares two active conditions. Baseline assessments (T0) were conducted prior to the randomization process. Each participant who appeared eligible for treatment following the phone screen, were invited to the clinic where they completed the initial assessment comprising a face-to-face interview and an online symptom questionnaire (BPM-P; Achenbach et al., 2011). Parents of children who were deemed eligible consented to study participation and were randomly assigned by an independent research assistant to one of the two conditions using an online number generator. Participants were masked for (i.e., unaware of) the treatment conditions. A total of 236 children were eligible for and included in the trial, and their parents were randomly allocated to one of two treatment conditions: *experiential* (EXP, n = 120) or *psychoeducational* (PE, n = 116). The first post-assessment was conducted 1 week after the group training (T1), the second assessment was conducted 1 week after program completion (group training + individual supervision: (T2). Follow-up assessments were conducted 4 (T3), 8 (T4), and 12 (T5) months after program completion. No financial compensation was offered, and the program was free of charge for all those participating. Participant flow is illustrated in CONSORT (Figure 1) following requirements for randomized clinical trials (Clinical Trials registration NCT03807336). Participants were enrolled over a five-month period from August to December 2018, and the data were collected from August 2018 to May 2020. All procedures were approved by the Regional Committee for Medical and Health Research Ethics (case 2018/754).

Participants

Youth, Parents, and Teachers

Participants were parents of children recruited from Oslo and Bergen, but parents from other urban and rural areas could also participate. The sample size comprised 236 children who were aged between 6 and 13 (M_{age} = 8.9, *SD* = 2.2; 143 boys [60.6%] and 93 girls [39.4%]) and 313 parents (M_{age} = 40.5, *SD* = 6.5). Parents participated either individually (234 mothers [74.8%] and two fathers [0.6%]) or in joint parent sessions (n = 77, 24.6%). Since the parent reports of their child's symptoms might be biased, owing to their participation in the treatment, the primary teachers of all included children were recruited as independent raters of change. However,

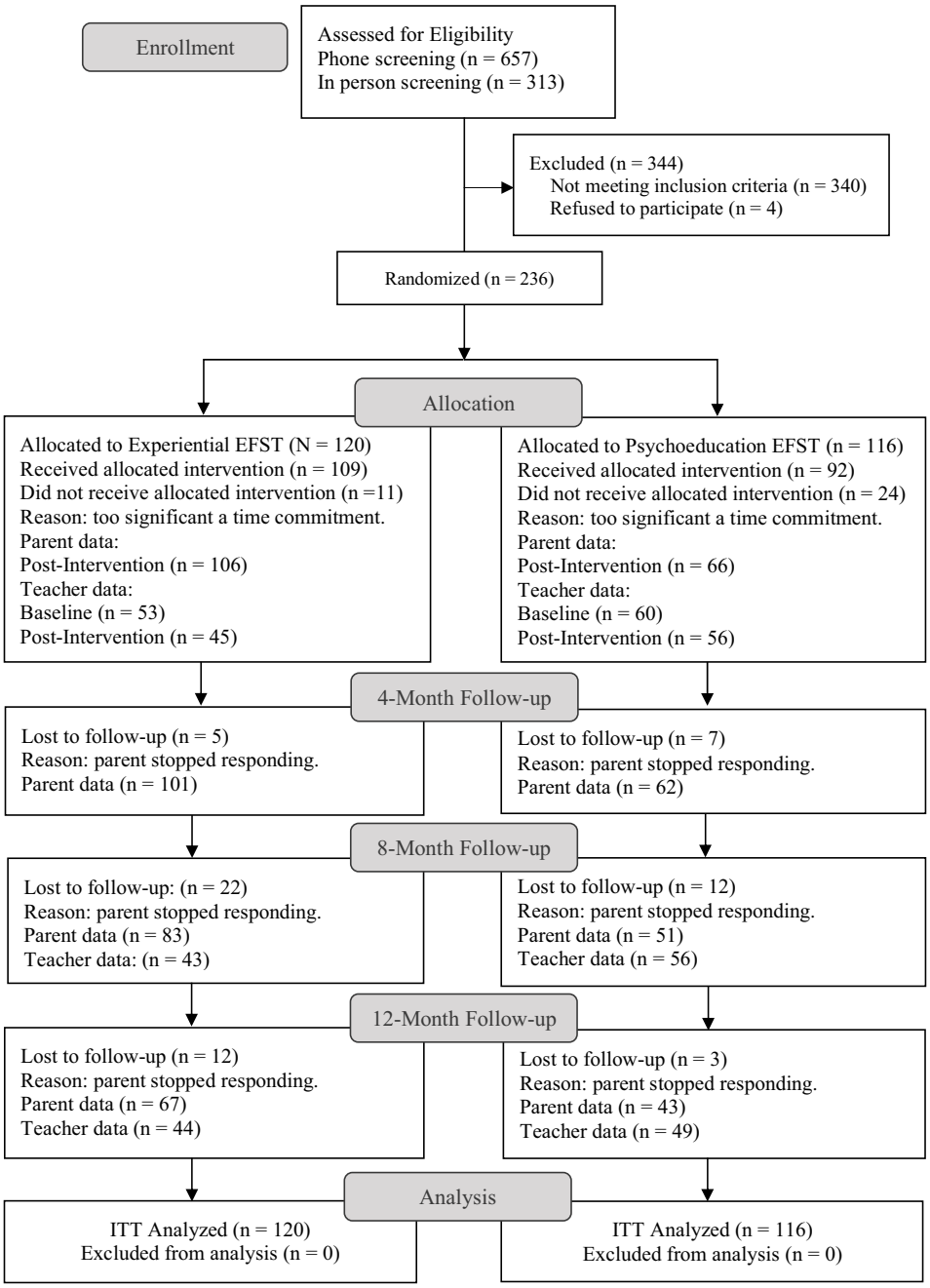

Figure 1. Consolidated standards of reporting trials (CONSORT) flow diagram of participants during enrollment, allocation, follow-up, and data-analysis. ITT intention-to-treat.

only about half of the children's teachers chose to participate (n = 113, 47.9%, 82% female). No other data on teachers were collected.

Therapists

A total of 13 therapists, comprised of 2 men and 11 women, provided both treatment conditions. Of these, 11 were clinical psychologists, and 2 were family therapists (one psychomotor physiotherapist and one clinical social worker), all having 5–15 years of clinical experience (M = 10.4, SD = 3.8). The group training was led by 4 out of the 11 therapists. All therapists were trained and certified in EFST, which involved a minimum of a 4-day group training, 15 hrs of supervision, and at least 1 year of experience using the program. Prior to the trial, all therapists received an extensive full-day training in the intervention. Each therapist provided both conditions to prevent unintended allegiance effects. During the study, the therapists received weekly supervision from two of

the authors (N.A. & J.R.S.), and if the therapists had any concerns or questions between these meetings, they could receive supervision in person. All therapists received at least 10 hrs of supervision from EFST trainers during the study. All sessions were videotaped, and after every session, the therapist filled in a session-form to check whether they had completed all components of the session.

Treatments

The manualized program for both conditions comprised two parts: a 2-day group-training program and 6 hrs of one-on-one parental supervision (20 hrs total). The 2-day group training was delivered in groups of 20–36 participants, and the individual supervision was delivered within a 3–4 weeks' time frame. The entire program was applied within 12 weeks.[1] The rationale for the program, group training, supervision plan, condition-specific treatment and examples were described in a treatment manual for the therapists (A. H. V. Hagen et al., 2019). Participants received separate condition-specific group training and supervision, and the EFST treatment manual was used for both conditions. The group training consisted of training in four core parenting skills in the following order:

In the first phase, *validation of emotion*, parents were given psychoeducation about the working of emotions and different emotional states, where one emotion (for instance, anger) might be helpful (primary adaptive), with regards to past trauma (primary maladaptive) or serve a protective function (secondary). Based on this understanding, they were assigned tasks to respond adequately to their child's emotions to meet their needs. The second phase, *enhancing motivation*, involved working with the problematic emotions of the parents that can disrupt the parent–child interaction (i.e., rejecting anger, fear of aggravating the situation or for the child's future, or shame/ self-blame for their child's difficulties). The goal was to alleviate problematic emotions and help the parent into a healthier emotional state, enhance their emotion-processing capacity and increase their motivation to meet their child's emotional needs. In the third phase, *resolving interpersonal injuries*, the goal was to help parents repair past relationship ruptures, thus liberating the child from emotional entanglement or feelings of being responsible, rejected, or dismissed by the parent. This was done by providing a genuine apology by (a) expressing their regret and attending to the child's vulnerable feelings; (b) validating their child's reactions; (c) taking radical responsibility and ownership for the situation; and (d) expressing heartfelt apology and committing to doing things differently henceforth. The final stage of the group training, *boundaries*, focused on strengthening new learning and increasing parental capacity to guide their child by setting sound and flexible boundaries. By accessing adaptive emotions, such as self-affirmation, assertiveness, and increased understanding of parents' own boundaries, the goal was to enhance the understanding of parents on their unique position to guide their child, and on the simple principle that changing their own boundaries could change their child's behavior.

In both conditions, the therapists were instructed to deliver the same EFST program, using basic humanistic principles from emotion theory and client-centered therapy including empathy, genuineness, and unconditional positive regard. After group training, parents were assigned individual supervision where they could choose to work with one of the four core skills that best suited their situation to help them implement and adjust the skills to their child's emotional needs. Each session commenced by evaluating the plan from the previous session, making adaptations, when necessary, before choosing another skill to work on. Therapists were taught to address all four parenting skills during the supervision period. The differences between the conditions are outlined below.

Experiential Condition (EXP)

In EXP, the interventions (both group training and supervision) were integrated within an experiential framework. This included experiential tasks such as (a) *focusing*; to assist parents in attending to their internal experience and to obtain bodily felt sensations and awareness of the issues they are struggling with (Gendlin, 1996); (b) *evocative empathy*, such as emphatic conjecture (imaginative entry into the experience of the other), acknowledging and expressing the pain and distressing feelings associated with the emotional injury; and (c) *two-chair dialogs* to access the problematic emotions of the parents, such as anger, fear, shame/self-blame, or sadness to help parents process unresolved feelings. In the group training, therapist self-disclosure of shameful experiences as a parent was shared with the group. In addition, parents were encouraged to write down a heartfelt apology to one of their own children, and to form and read out loud an apology they wished they had got from their own parent. The goal was to foster emotional processing by increasing the emotional arousal and ability of the parents to allow,

[1]The program was planned to be delivered within 10 weeks but was postponed due to practical challenges. Data collection or study activities were not affected by COVID-19, (i.e., all parents participated in person in the program).

stay in contact with, process, and make sense of their own emotions, and thereby get a deeper sense and meaningful understanding of the emotional reactions of their child. Imaginative two-chair dialogs with "parts of the self" or the "imaginary child" were used for each of the four core parenting skills. For example, in the *enhance motivation* task, parents were encouraged to have a dialogue with other parts of their self that either scared or criticized them, to get a deeper sense of how their own emotions can get in the way of attending properly to their child. The goal of the EXP tasks was to form a deeper connection with their intuitive parental skills and the imagined experience of the child, which in turn could alter their way of relating to their child's expressed difficulties.

Psychoeducational Condition (PE)

In PE, parents received the same EFST program as in EXP, except not using experiential tasks, such as bodily awareness, evocative empathy, focusing tasks, and two-chair dialogs. In the group training, participants were informed of the treatment rationale didactically with active therapeutic engagement. Rather than self-disclosing, the therapists shared case examples from other parents' struggle and parents were given homework assignments, for example, to practice an apology toward their child or to practice validation. In supervision, parents were encouraged to bring back situations where they experienced difficulties with regards to their child and the therapists gave concrete guidance about how to deal with such situations, such as a rationale for validation of emotions. The therapists provided active guidance and supervision without experiential tasks to increase the understanding of parents on the working of emotions, awareness, and acceptance of their own and their child's emotions. Therapists in the PE intervention were explicitly instructed to refrain from using EXP interventions (i.e., chair work). Adherence checks indicated that they managed to keep these distinctions, see below.

Adherence

Adherence was indirectly assessed through therapist supervision and therapist self-reports. During the trial, all therapists had condition-specific small-group meetings that were conducted on a weekly basis, where supervisors were instructed to guide the therapists to adhere to the particular condition. Adherence to the treatment manual was monitored by supervisors and group-leader reports on a task-specific intervention adherence form, which specifies the four core skills in EFST. The therapists completed the measure form after each session and stated whether they deviated from the protocol. The form comprised adherence questions (e.g., *"To what extent did you deliver the treatment in accordance with its design and manual?"*) and therapist-perceived competence (e.g., *"How well did you feel that you delivered this treatment condition?")*. Deviations from the protocol were examined by analyzing the therapists' video-recordings. The study did not include independent ratings of the sessions to be correctly classified in accordance with treatment adherence criteria, and video-recordings were primarily used to detect deviations from the manual. In cases where deviation was discovered (n = 11), either supervision based on video-recordings were given where there was room for adjustments (n = 4), or cases were eliminated from the study altogether (n = 7). However, the participants were allowed to complete the treatment.

Alliance Measure

Since nonspecific or common therapeutic factors can vary between treatment conditions, an integrated instrument was applied to measure alliance in the group training and individual supervision. The current questionnaire is based on an adapted version of WAI-SF (Horvath & Greenberg, 1989) and WAI-SR (Group version[2]), with integrated questions regarding the supervisory alliance, parents' description of the content of their treatment as well as overall satisfaction with the program. The inventory consisted of 14 items with scores from a 7-point Likert scale comprising seven items to rate the alliance and satisfaction in the group training (based on the WAI-SR, e.g., "*My participation in this group gives me new ways of looking at my problem*"), six items to rate the alliance and satisfaction in supervision (based on WAI-SF, e.g., "*The supervision gave me new ways to look at my problem*"), and one to rate the content of their treatment *("What was it that you mainly focused on today?"*). The participants filled out this form after group training and after each supervision (T0-T2). The instrument was found to have high degree of internal consistency (α = .88 [95% CI .87; .89]). Since all measures were statistically different from normal assumptions (all Shapiro–Wilk tests < .01), a confirmatory factor analysis with diagonally weighted least squares showed that one-factor model was a good fit for data (χ^2 = 161.99, df = 54, p < .001; CFI = .97, TLI = .096, SRMR = .082, RMSEA = .066; p < .001) according to Hu and Bentler's cutoff criteria for fit indices (Hu & Bentler, 1999). See Supplemental material (*S1*) for details.

[2]https://wai.profhorvath.com/downloads.

Symptom Measures

Brief Problem Monitor for Ages 6–18 (Achenbach et al., 2011). The children's behavior and symptomology were assessed by caregivers (Brief Problem Monitor for parents: BPM-P) and teachers (Brief Problem Monitor for teachers; BPM-T; Achenbach & Rescorla, 2001). Both BPM-P and BPM-T are widely used instruments that assess a broad array of behavioral and emotional manifestations of psychopathology and consist of 19 items – six for Internalizing problems (INT; anxiety, depression, withdrawal, and somatic problems), seven for Externalizing problems (EXT; aggression, and rule-breaking behavior), and six for attention/hyperactivity problems (B. C. Chorpita et al., 2010). Items are rated as *not true* (0), *somewhat or sometimes true* (1), or *very true or often true* (2). BPM scales in terms of *T* scores are based on multicultural norms for a child's age and gender. The validity and reliability of these scales are well documented (Achenbach & Rescorla, 2001). An established measure translation was used, applicable for the current sample (Richter, 2015). For the present sample on the BPM-P scale, we found test–retest reliability to be $r = 0.81$ ($p < .001$), and internal consistencies (Cronbach's *alpha*) to be $\alpha_{INT} = 0.80$ and $r_{INT} = 0.83$ ($p < .001$), and $\alpha_{EXT} = 0.88$, $r_{EXT} = 0.87$ ($p < .001$). For the BPM-T scale, we found $r_{INT} = 0.86$ ($p < .001$), $\alpha_{INT} = 0.80$ and $r_{EXT} = 0.88$, $\alpha_{EXT} = 0.88$ ($p < .001$). BPM-P was assessed at baseline, after group training, at post-intervention, and at 4, 8, and 12 months' follow up (T0–T5; all six occasions), while BPM-T was assessed at T0, T2, T4, and T5 (four occasions). Participants were asked to register their responses online within 1 week from each assessment point.

Statistical Analyses

All data were analyzed with the principle of intention-to-treat, i.e., all cases included were analyzed. Analyses were conducted using SPSS Statistics version 28.0 (IBM Corp, 2021) and The R software[3] to perform Bayesian analyses and visualize data with the following packages: brms (Bürkner, 2017) and ggplot (Wickham, 2016). Effect sizes for absolute *r* were interpreted as follows: > 0.10, small; > 0.30, medium; > 0.50, large (Cohen, 1988) and 95% CI calculations (Lenhard & Lenhard, 2016). Chi-square (χ^2) tests were applied for categorical data. ANOVA and t-tests were used for demographic and clinical differences at baseline.

We investigated hypothesis 1 (proposing that both versions of EFST are effective in alleviating symptomatic distress in children); and hypothesis 2 (proposing that the EXP condition will outperform the PE condition), using multilevel modeling (MLM). Because repeated measures (Level 1) were nested within children/patients (Level 2), we used a two-level hierarchically nested growth model to analyze the effect of EFST and to investigate whether there were differences in outcomes depending on treatment condition. An important reason to use these methods is to account for non-independence in the data due to nesting. Failure to account for data dependency could result in an underestimation of the standard errors which could lead to in an inflated Type I error rate (Raudenbush & Bryk, 2002). MLM is also robust in allowing for missing observations, which is a typical problem in longitudinal, naturalistic psychotherapy research (Hox, 2010). Following the procedures recommended by Singer and Willett (2003), we built models successively adding variables and interaction terms according to our hypotheses. These variables acted as covariates in the adjusted regression models. Model 1 (the "null model") included only a fixed intercept, a fixed slope for time (representing all measurement points including follow-up data for 4, 8 and 12 months after the intervention). In model 2, we added "treatment condition" as a dummy coded variable (with EXP = 1, PE = 0) to investigate the potential differential effect of outcomes between conditions (using an interaction term; treatment condition*slope). Model fit was evaluated with the Akaike's information criterion (AIC; Akaike, 1974), with reducing scores indicating less deviance and hence improved model fit. An alpha level of 0.05 was chosen for all tests, and all tests were two-tailed. Log-linear and quadratic modeling of time were also tested, however, none of these increased the model fit. Hence, we opted for linear modeling of time to ease interpretation. The time intervals between the measurement waves were identical from post-treatment to 12 months follow up (4 months). Random intercepts and slopes were also included but only random intercepts were significant, hence we discarded random slopes in the models. We tested the following models for both parent and teacher ratings of BPM.

Model 1 (Null-Model)

Level-1
(repeated measures):-

$$Outcome_{ij} = b_{oj} + b_{ij}(Time_{ij}) + r_{ij}.$$

Level-2 (patient level): $b_{0j} = g_{00} + u_{0j}$

$$b_{1j} = g_{10} + u_{1j}$$

[3]https://www.organizingcreativity.com/2020/08/citing-r-and-rstudio.

Model 2 ("Treatment Condition")

Level-1 (repeated measures): $Outcome_{ij} = b_{0j} + b_{ij}(Time_{ij}) + r_{ij}$.

Level-2 (patient level): $b_{0j} = g_{00} + g_{01}(treatment\ \ condition_j) + u_{0j}$

$$b_{1j} = g_{10} + g_{11}(treatment\ condition_j) + u_{1j}$$

Reliable change Index (RCI) was used to calculate whether the amount of change from pre- to posttreatment was large enough at an individual level (as reported by parents), compared to the variability of the pretreatment groups, considering the reliability of the measure used,[4] thus allowing to ascertain whether clinical meaningful improvement was made (Jacobson & Truax, 1991). The groups were divided into four categories of clinical outcome; *improved* (RCI larger than 1.96); *deteriorated* (RCI less than 1.96); *crossed clinical cutoff* (crossed the symptom threshold of < 65, but not improved or recovered), and *recovered* (both (a) improved and (b) crossed clinical cutoff; Beurs et al., 2016).

Attrition and Missing Data

The number of respondents in both groups and measurement occasions is presented in the CONSORT diagram (Figure 1). In the sample, 90.8% in the EXP group, and 79.3% in the PE group completed the intervention. The overall missingness in the EXP group was 25.6% (T2: 11.7, T3: 15.8, T4: 30.8 and T5 44.2), and 46.6% in the PE group (T2: 20.7, T3: 46.6, T4: 56 and T5: 63). As the between-group difference in dropout rates across the study period was statistically significant (χ^2 [1, 208] = 58.22, $p < .001$), missingness was further analyzed according to the procedure outlined by Buuren (2018) and assumed missing at random (MAR), with condition, time, and previous severity of symptom as predictors. The Kaplan–Meyer test, which is a nonparametric statistic for survival analysis (Dudley et al., 2016), was used to analyze the missing data. Gender and age were not significant predictors of outcome variable missingness (all p's > 0.8). However, condition, time, and symptoms at previous time points were significant predictors (all p's < .001). Thus, we cannot refute the assumption of missing not at random (MNAR) as the pattern of missingness.

To ensure the validity of our findings, and to assess whether certain missing data conditions would change the interpretation of our findings, we expanded our intended analysis with sensitivity checks for missing data patterns to analyze the same model under different assumptions of missingness, as outlined by Linero and Daniels (2018). Sensitivity post hoc analysis, including multiple forms of data imputation, was conducted under three assumptions that the missing values can be (a) predicted by intra-individual factors, (b) reasonably estimated by the last occasion measured, or (c) predicted by values of other individuals at the same time point. A Bayesian approach was employed to reasonably reflect the inherent uncertainty in such a sensitivity analysis (Smid et al., 2020) and the results were reported using the Bayes Factor (*BF*), which describes the weight of evidence for the alternative over the null hypothesis. In these sensitivity analyses, an imputation model was defined first. For model (a), data were imputed based on prediction of gender, age, and pretreatment symptoms. Model (b) imputed data based on last available measurement of the outcome variable. For model (c), data were imputed using predictions of the group-level mean outcome measure at the missing time point. For each model, 10 imputed data sets were created. These imputed data sets were visually compared with each other and the observed data. As expected, these data sets differed from each other. Therefore, rather than interpreting them individually, the key is to see whether all point in the same direction. All described models were estimated using Hamiltonian No-U-Turn sampling, with 10,000 iterations and 1000 burn-in iterations across three chains in the brms package (Bürkner, 2017). Priors were diffuse to maintain estimate uncertainty (prior for regression estimates was normal distribution with mean 0 and SD 2). Bayes factors (*BF*s) for all model parameters were calculated using the Savage–Dickey ratio (Bürkner, 2017). Regardless of the missing data mechanisms, while the ranges that exclude values above 3 were considered evidence of a substantial difference, ranges that exclude values below 3 were considered evidence of substantial lack of difference (Linero & Daniels, 2018).

Statistical Power

Statistical power calculations were performed using the principle of G*power for the primary aims of evaluating the equivalence and the efficacy relative to a benchmark on the minimum clinical difference

[4] $RC = (X_2.X_1)/S_{diff}$; X_2 = posttreatment score, X_1 = pretreatment score, $S_{diff} = \sqrt{(2SE)^2}$, $SE = SD\sqrt{(1-r)}$, SD = standard deviation of the "control group,", and r = the test–retest reliability of the measure.

between the two treatment conditions (Faul et al., 2009). These were a priori power calculations used for sample size planning. Based on a power analysis for the analysis of main effect, we concluded that a sample size of 98 participants per treatment group will detect an effect size of 0.4, provided power of .80, at an alpha level of .05. Since an expected drop-out of 20% was accounted for, the recruitment of N = 236 children provided adequate power for the analyses of both equivalence and superiority (Faul et al., 2009). The recruitment was managed according to plan, although the follow-up dropout rate was larger than expected (see below).

Results

Demographical and Clinical Characteristics

The descriptive statistics are presented in Table 1. The sample size was 95.8% white and 4.2% non-Western (which is below the community from which the sample was drawn, with 10.9% non-Westerns[5]). The average annual family income was above average (M = $73,808, SD = $11,355). The educational level was high; among half of the parents (48.6%) had completed 3–4 years of college degree and 38.3% had a post-college degree. Earlier treatment had been received by 119 (37.3%) of the participants (1–7 times), more than half of them (n = 163, 69.1%) showed both INT and EXT within clinical range, and a moderate significant correlation was found between the two sets of symptoms (T ≥ 65, r = .453, p = .001), reflecting the comorbidity in the sample. No significant baseline differences were observed between the EXP and PE samples with regards to age, gender, income, social status, or symptom severity (for details, see, Table 1).

Effect Sizes

The pre, post, and follow-up mean, standard deviations, effect sizes, 95% confidence intervals (CI), and p-values are displayed in Table 2, without missing data imputation. Parent reports (BPM-P) indicated large effect sizes on reduced symptoms in both INT and EXT for both conditions: d_{INT} = 1.09, 95% CI [0.5, 1.26] in EXP and d_{INT} = 0.76, 95% CI [0.31, 0.90] in PE and d_{EXT} = 1.21, 95% CI [0.55, 1.28] in EXP and d_{EXT} = 0.79, 95% CI [0.09, 1.05] in PE. Teacher reports (BPM-T) indicated no effect on INT in EXP, but a small effect in PE, d = 0.28, 95% CI [0.13, 0.63], and medium effects on EXT for both conditions, d = 0.46, 95% CI [0.32, 1.33] in EXP and d = 0.38, 95% CI [0.08, 0.89] in PE. Parent-reported EXT differed significantly at T3 and T5 (p's < .02); the largest absolute difference was t = 3.26 at 12-month follow-up (0.33 SD of the BPM). Parent-reported INT did not differ significantly across the 4 post-treatment assessments; the largest absolute difference was t = 2.04 at 4-month follow-up (0.20 SD of the BPM). The differences between conditions at T5 was small on INT, Hedges' g = .28 (95% CI 0.106–0.663) and medium on EXT, Hedges' g = .55 (95% CI 0.969–0.131).

Therapeutic Alliance

Both conditions reflected good alliance and satisfaction, and the independent t-test revealed no statistically significant differences between groups, t [313] = 1.432, p = .181. On a 7-point Likert scale (where 1 = *not at all* and 7 = *always*), parents rated the alliance to be M = 5.9 (n = 286, SD = 0.94) after group training and M = 5.6 (n = 282, SD = 1.38) after supervision, indicating high satisfaction.

Adherence

Based on the therapist's self-reported adherence measures, the average fidelity for both group-leaders and supervisors were 94.6% for EXP and 96.7% for PE. The average measure of perceived therapist competency was 94.7%. General treatment adherence over the group training and after all supervision sessions was 96.1% for PE and 94.2% for EXP. Treatment differentiation (the presence of sufficient differences between the two treatment conditions) is achieved if more than 90% of sessions were correctly classified in accordance with the treatment adherence criterion recommended by Leeuw et al. (2009). The difference on adherence between conditions was nonsignificant, t [828] = 88.74, p = .076. This result is, however, based on the therapists' self-reports and was not examined by independent raters.

Multilevel Growth Curve Modeling

The starting point on the BPM-P (fixed intercept) in the sample was 65.92 (SD = 10.4, 95% CI 64.81–67.44) on EXT symptoms. For the INT symptoms, the fixed intercept was 70.18 (SD = 9, 95% CI 69.09–71.14). Both differed significantly at baseline (p < .001), indicating

[5]https://www.ssb.no/en/statbank/table/09817/.

Table 1. Baseline characteristics of children and parents in both conditions.

Parent demographics	EXP (n = 160)	PE (n = 153)	Total (n = 313)	p
Parent age *n* (*m*, SD)	160 (40.6, 6.7)	153 (40.5, 6.2)	313 (40.5, 6.5)	.680
Parent *n* (%) female	120 (75)	115 (75.2)	235 (74.8)	.764
Parent *n* (%) male	40 (25)	39 (25.4)	79 (25.2)	.905
Single *n* (%)	20 (12.5)	17 (11.1)	37 (11.8)	
Married/cohabitants	119 (74.4)	115 (75.2)	234 (74.8)	
Separated/divorced	21 (13.1)	21 (13.7)	42 (13.4)	.267
Diploma 1–3 years	18 (11.3)	23 (15)	41 (13.1)	
College/University 3–4 years	83 (51.9)	69 (45.1)	152 (48.6)	
Post-College degree	59 (36.8)	61 (39.9)	120 (38.3)	.567
Norwegian	141 (88.1)	139 (90.9)	280 (89.5)	
European	11 (6.9)	8 (5.2)	19 (6.1)	
Other nationalities	8 (5)	6 (3.9)	14 (4.5)	.357
Christian	62 (38.8)	58 (37.9)	120 (38.4)	
Non-religious	90 (56.2)	87 (56.9)	177 (56.5)	
Other religions	8 (5)	8 (5.2)	16 (5.1)	.291
Heterosexual	149 (93.1)	141 (92.2)	290 (92.7)	
Homosexual/lesbian/bisexual	11 (6.9)	12 (7.8)	23 (7.3)	.128
Full-time employed	119 (74.4)	108 (70.6)	227 (72.5)	
Part-time employed	25 (15.6)	24 (15.7)	49 (15.7)	
Unemployed/sick leave	9 (5.6)	8 (5.2)	17 (5.4)	
Student/ /resident	7 (4.4)	13 (8.5)	20 (6.4)	.227
No earlier treatment*	149 (93.1)	138 (90.2)	287 (91.7)	
Earlier treatment ≤ 1–3 times	7 (4.4)	13 (8.5)	20 (6.4)	
Earlier treatment ≤ 4–7 times	4 (2.5)	2 (1.3)	6 (1.9)	.484
Children demographics	**EXP (n = 120)**	**PE (n = 116)**	**Total (n = 236)**	**p**
Youth age *n* (*m*, SD)	120 (8.9, 2.1)	116 (8.9, 2.0)	236 (8.9, 2.1)	.921
Youth *n* (%) female	48 (40)	46 (39.7)	94 (39.8)	
Youth *n* (%) male	72 (60)	70 (60.3)	142 (60.2)	.103
INT *n* (%, *m*, SD)	19 (15.8, 70.18, 8.8)	21 (18.1, 70.03, 9.2)	40 (16.9, 70.11, 9)	.710
EXT *n* (%, *m*, SD)	18 (15, 65.6, 7.2)	15 (12.9, 66.2, 6.8)	33 (14, 65.9, 7)	.906
INT and EXT *n* (%, m, SD)	83 (69.2, 69.8, 8.9)	80 (69, 69.2, 8.6)	163 (69.1, 69.5, 8.4)	.202
No earlier treatment* *n* (%)	75 (62.5)	73 (62.9)	148 (62.7)	
Earlier treatment 1–3 times	23 (19.2)	24 (20.7)	47 (19.9)	
Earlier treatment 4–7 times	22 (18.3)	19 (16.4)	41 (17.4)	.071

EXP, experiential condition, PE, psychoeducational condition. INT internalizing symptoms and EXT externalizing symptoms (*T* scores 65 or higher on Brief Problem Monitor for parents; Achenbach & Rescorla, 2001). *Due to mental health issues.

a significant within-person variance in symptoms at pre-treatment. The linear effect in model 1 indicated a significant reduction of symptoms across treatments, i.e., the estimated fixed slope was $b_{INT} = -1.71$ ($SE = 0.145$, $p < .001$) and $b_{EXT} = -1.72$ ($SE = 0.113$, $p < .001$). The negative estimate of the linear growth coefficient indicates that, on average, parents reported a reduction in the participants' INT and EXT symptom levels over time, and statistically significant improvements for both conditions. In model 2, we added "treatment condition" as a dummy coded variable (with EXP = 1, PE = 0) to investigate the potential differential effect of outcomes between conditions (assessed using an interaction term with the fixed slope). Model 2 shows that the variation in intercepts between conditions was nonsignificant [$est._{INT} = 0.30$, $SE = 0.496$, $p = .304$ and $est._{EXT} = 0.53$, $SE = 0.759$, $p = .530$], suggesting that the randomization process was successful (i.e., symptomatic distress at baseline was the same across conditions). Results from the growth curve model showed that the interaction between slope and condition for both INT and EXT were nonsignificant [$b_{INT} = -1.32$, $SE = 0.217$, $p = .076$, and $b_{EXT} = -1.51$, $SE = 0.184$, $p = .165$]. For a visual inspection, see line graphs, *S2*.

For teachers' feedback (BPM-T), we built models that were identical to the parental feedback (see above), adding variables and interactions in accordance with our research questions to investigate the linear effects of time and treatment condition. Model 1 ("null model") indicated a statistically significant reduction of symptoms across treatment in EXT [$b = -0.96$, $SE = 0.202$, $p = .001$], but not in INT [$b = -0.13$, $SE = 0.178$, $p = .445$]. Model 2 exhibited nonsignificant interactions between slope and condition for both EXT [$b = 0.76$, $SE = 0.396$, $p = .987$] and INT [$b = -0.23$, $SE = 0.192$, $p = .367$], indicating that the treatment condition was not a significant predictor of children's symptom reduction (for details, see Table 3).

Clinically Significant Change

The results from clinically significant and reliable change calculations on EXT and INT symptoms for each condition were based on parent reports. Both improvement and recovery rates appeared to be higher in the EXP condition. For the PE group, 20.3% was improved and 15.2% was recovered on INT, and 18.6%

Table 2. Mean scores and effect sizes (Cohen's d) of differences in means across assessment waves.

	Baseline-T0		After group training-T1		Post-Intervention-T2		4-Month FU-T3		8-Month FU – T4		12-Month FU – T5		*d* [95%CI]	
	Mean (*n*, *SD*)	*p*	Mean (*n*, *SD*)	*p*	Mean (*n*, *SD*)	*p*	Mean (*n*, *SD*)	*p*	Mean (*n*, *SD*)	*p*	Mean (*n*, *SD*)	*p*	T2; T0	T0; T5
Internalizing														
Parent reports														
PE	70.09 (116;9.2)		66.24 (92; 5.9)		66.95 (59; 7.8)		65.33 (62; 5.8)		63.45 (58; 5.9)		64.70 (49; 6.2)		0.53 [0.69–0.10]	0.76 [0.31, 0.90]
EXP	70.27 (120;8.8)	.710	65.60 (109;5.1)	.503	65.70 (106;7.9)	.503	63.29 (101;5.3)	.332	63.03 (79; 5.3)	.115	62.99 (53; 6.1)	.901	0.77[0.99–0.46]	1.09 [0.50, 1.26]
Teacher reports														
PE	60.01 (60; 5.5)		_		60.27 (56; 5.8)		_		59.74 (56; 6.4)		60.07 (49; 5.9)		0.20 [0.61–0.10]	0.28 [0.13, 0.63]
EXP	59.23 (53; 5.9)	.252	_		57.53 (45;6.4)	.047*	_		58.82 (43; 6.2)	.667	57.37 (44; 6.3)	.057	0.01[0.42–0.39]	0.05 [0.38, 0.49]
Externalizing														
Parent reports														
PE	65.72 (116; 10.1)		61.15 (92; 6.0)		61.38 (59; 9.0)		60.02 (62; 7.3)		56.84 (58; 5.9)		61.09 (49; 5.6)		0.74 [0.98–0.37]	0.79 [0.09, 1.05]
EXP	66.12 (120;10.6)	.906	60.10 (109;6.6)	.293	59.47 (106;8.1)	.170	57.32(101;6.2)	.019**	57.28 (79; 6.1)	.679	57.83 (53; 6.1)	.011**	0.99 [1.14–0.58	1.21 [0.55, 1.28]
Teacher reports														
PE	61.35 (60; 6.4)		_		59.55 (56; 6.0)		_		58.44 (56; 5.7)		58.94 (49; 7.2)		0.28 [0.70-0-03]	0.38 [0.08, 0.89]
EXP	61.83 (53;6.0)	.686	_		60.11 (45;6.4)	.619	_		57.42 (43; 6.0)	.453	59.18 (44; 6.8)	.853	0.28 [0.78–0.06]	0.46 [0.32, 1.33]

n sample size, SD standard deviation, *d* Cohen's d, 95% confidence intervals (CI) with lower and upper bound and *p*-values of differences between conditions. Primary outcome on internalizing and externalizing symptoms in children measured by parents with Brief Problem Monitor (BPM-P) and teachers with Brief Problem Monitor for teachers (BPM-T). FU follow-up, PE psychoeducation, EXP experiential.

was improved and 27.1% was recovered on EXT. For the EXP group, 31.3% was improved and 28.3% was recovered on INT, and 41.5% was improved and 36.7% was recovered on EXT. For the EXP group, 28.3% of the INT and 24.6% of the EXT falls into the category "Uncertain or unimproved." For the PE group, this applies to 39% on INT and 34.6% on EXT (Table *S3*).

Bayesian Longitudinal Analysis

All models' *BFs* are reported in Supplemental material (*S4*), to show how missing data mechanisms affect model interpretations. Based on parent reports, model 1 was robust to different assumptions about missingness. However, Bayesian sensitivity analyses under different assumptions of missingness indicated a substantial statistical interaction between slope and condition for both INT [*BF* 7.47, 150], and EXT [*BF* 73, 150] symptoms. When compared to INT symptoms, EXT symptoms were less affected by missing data mechanisms. Sensitivity analyses for the interaction of slope and condition on INT all indicated a substantial effect (lowest *BF* 7.47). Similarly, the interaction of slope and condition on EXT showed an effect regardless of how the missing variables were modeled (lowest BF 73). This suggests that symptom reduction differed between conditions in favor of EXP, although establishing the magnitude of this difference depends on assumptions about missing data mechanisms (S4). Based on teacher reports, the Bayesian sensitivity analyses identified that the results varied depending on assumed mechanism of missing data: i.e., BF for slope ranged from 0.05 to 150 for EXT, the first indicating evidence of no difference while the second indicating evidence of difference between EXP and PE conditions. Hence, interpretation of any difference between EXP and PE should be taken with caution.

Table 3. Growth curve model for primary outcome variables across T0–T5.

Growth curve model						
	Model 1			Model 2		
Variable	Est.	S.E.	95% CI	Est.	S.E.	95% CI
Parent ratings						
Internalizing						
Fixed effects						
Intercept	70.18***	.374	[69.09, 71.14]	69.64***	.583	[68.49, 70.78]
Slope	−1.71***	.127	[−1.97, −1.47]	−1.69***	.136	[−1.95, −1.42]
Condition				.304	.759	[−1.94, .46]
Condition*Slope				.320	.260	[−1.08, .054]
Model fit						
AIC	6262.236			6190.550		
Externalizing						
Fixed effects						
Intercept	65.92***	.413	[64.81, 67.44]	64.91***	.637	[63.66, 66.17]
Slope	−1.72***	.113	[−1.94, −1.50]	−1.51***	.184	[−1.87, −1.15]
Condition				.530	.834	[−2.16, 1.1]
Condition*Slope				.165	.233	[−.78, .13]
Model fit						
AIC	6170.130			6001.813		
Teacher ratings						
Internalizing						
Fixed effects						
Intercept	59.62***	.549	[58.39, 60.56]	57.44**	.657	[54.14, 60.74]
Slope	−.13	.178	[−.48, .21]	−.23	.192	[−1.31, .84]
Condition				.20	.637	[−.75, 3.57]
Condition*Slope				.853	.367	[−.64, .77]
Model fit						
AIC	2438.568			2442.567		
Externalizing						
Fixed effects						
Intercept	61.59***	.561	[59.91, 62.11]	60.38***	1.782	[56.86, 63.89]
Slope	−.96***	.202	[−1.34, −.57]	− 1.13	.599	[−2.31, .05]
Condition				.71	1.172	[−1.87, 2.74]
Condition*Slope				.76	.396	[−1.67, 2.52]
Model fit						
AIC	2449.268			2445.569		

*p < .05; **p < .01; ***p < .001. Estimates for primary outcome measures of internalizing and externalizing symptoms in children measured by parents with Brief Problem Monitor (BPM-P) and teachers with Brief Problem Monitor for teachers (BPM-T). Estimates (Est.) 95% confidence intervals (CI) with lower and upper bound and standard error (S.E.). AIC Akaike information criterion. Model 1: Intercept and slope, no covariates. Unstructured covariance matrices show significant differences in growth curves between internalizing and externalizing symptoms and slope reported by parents, and significant differences in externalizing symptoms, but not on internalizing symptoms, reported by teachers. Model 2: adding treatment condition improve model fit on both internalizing and externalizing symptoms based on both parent and teacher reports.

Discussion

This is the first study to systematically investigate the efficacy of an intensive skills-oriented parental program based on principles from EFT that compares two active conditions (one experiential; EXP, and one psychoeducational; PE). Clinical trials are commonly criticized for their failure to replicate the comorbidity and complexity that clinicians encounter in real-world settings. As a remedy, the inclusion criteria chosen for this trial were liberal and involved a broad range of clinical symptoms to allow heterogeneity among children. Training clinicians to deliver a transdiagnostic protocol that can target comorbid states can prove to be cost-effective because clinicians adhere to core strategies that can be flexibly applied to a range of emotional experiences. In this study, EFST showed efficacy for reducing parent-reported INT and EXT problems, with large improvements in both conditions. The mean EXT and INT symptoms were below clinical range at 1-year follow-up. These findings correspond well with those obtained in other studies in EFFT, including Foroughe et al. (2019) and Strahan et al. (2017). Other emotion coaching parental programs show comparable results (e.g., Gottman et al., 1997; Havighurst et al., 2013; Katz et al., 2012). Nevertheless, as the effects of EFST were not compared with another active treatment intervention, conclusions about causality are not possible.

Against our prediction, the multilevel analyses did not indicate significant differences in remission rates between conditions. Since the EXP condition was considered emotionally evocative, as it involves imaginative dialogs with "parts of the self" or the "imaginary child," we expected this condition to lead to a significantly larger reduction in symptoms than the PE condition. There are several possible explanations for why this was not found. One rationale could be the relatively similar treatment content, and levels of group training and supervision, which make it difficult to detect possible differences between them. The lack of time in active treatment to process and consolidate the experienced effects could be another explanation. Some studies point to the importance of ongoing repetition of key interventions to provide opportunities to internalize change (Greenberg & Pascual-Leone, 2001). The short-term nature of this format may have masked a potential difference between the conditions, since the intervention is too brief to facilitate profound changes (Pascual-Leone & Greenberg, 2007). Additionally, parents in both conditions might have received sufficient aid to apply these skills in situations with their children. Finally, vast amounts of evidence prove that different treatment models perform equally well in psychotherapy treatment, a finding which has been referred to as the dodo bird verdict (Luborsky et al., 1975; Wampold & Imel, 2015).

Despite the similar outcome in the two conditions, there was a higher level of dropout across T2–T5 in the PE when compared to the EXP condition. The post hoc analysis of missingness showed significant difference between conditions. Given that the PE group lost 20% more participants compared to the EXP, we may run the risk of overestimating the effect of PE when using growth curve modeling (Smid et al., 2020). When accounting for different mechanisms of missingness, all Bayesian models indicated that EXP was superior to PE in reducing parent-rated child symptoms. Since the participants did not know which condition they received, this difference may imply that the EXP group did better. On the other hand, therapist adherence was very good across conditions, and there were no significant differences in parent satisfaction or alliance. In addition, despite significant improvements in both conditions, most mean T scores were still > 60 on INT at 1-year follow-up ($M = 63.8$). Regarding the clinical significance, most children (> 50%) in each condition did not reach the RCI criterion for being clinically improved in parent-reported INT nor EXT, leading toward the other direction, i.e., both groups equally benefitted from the interventions, although some children in both groups were still above the clinical cutoff and may benefit from additional health care services.

Teacher reports showed significant effects in EXT on both conditions. Other comparable trials have found similar effects (A. K. Hagen et al., 2011; Havighurst et al., 2013). The lack of significant effects on INT might be explained by the primary school context, where EXT problems tend to be more observable than INT problems (Bongers et al., 2004). Analysis of missingness showed no significant differences between the conditions. Results from Bayesian analyses suggested that results on teacher-rated outcomes varied depending on the mechanism of missing data. For future research, achieving higher N may be beneficial to account for missingness and to get a better understanding of teachers' feedback.

Since the research in EFST is at a rudimentary stage, alternate pathways that may mediate the long-term effects of such parenting programs, need to be investigated. In a qualitative study, which was part of the current research, participants reported that they gained access to their practical wisdom as parents, experienced a new calmness and were less afraid. Many parents described a changed attitude of willingness to do what is needed for their child, providing them with feelings of self-efficacy and agency (Ansar et al., 2021). Several studies have demonstrated associations between improved parental self-efficacy and child outcomes, enabling parents to

maintain the learned skills (e.g., Jones & Prinz, 2005). Additional assessments of the parents' maintenance of EFST learning across posttreatment, which is subject to a secondary outcome study, may give valuable information about how parents' own program retention predicts the magnitude of program effects on children.

Strengths and Limitations

Parenting interventions focused on emotional dimensions of parent–child relationships have been shown to be beneficial, yet this literature is still burgeoning and greater evidence for how such programs work is needed. Hence, it is valuable to determine which aspect(s) of a multidimensional program conveys the benefits, as in this dismantling approach. The repeated follow-up over a full year enables close tracking of program efficacy. A thorough and novel comparison of two versions of EFST, their effect on children's INT and EXT, and the teachers' inclusion as independent reporters, are other strengths. In addition, the use of different methods of analyzing efficacy gives a nuanced picture of the findings. Due to missing data, the study may have been prone to failing to reject the null hypothesis of no differences between conditions, exemplifying the importance of the Bayesian add-on addressing missing data issues. By using the same therapists in both conditions, we reduce the potential for allegiance effects, although we cannot rule out that allegiance effects may have led to lower dropouts and better effects in the EXP condition.

There are also some important limitations to consider. First, the lack of a no-treatment condition (be it waitlist or TAU) or other more established emotion-oriented programs, makes it difficult to ensure whether the reduction in child symptoms is due to regression to the mean or a result of the program. Second, the missing data was not random but systematically confounded with treatment conditions. Due to uncertainty around missingness, the sensitivity of findings was assessed to various possible causes. A provisional interpretation of the findings was possible in light of these sensitivity analyses. However, future research should focus on deciphering why differences in missingness occur and the implications for clinical practice. Third, the outcome measures were based on parental self-report assessments. Although independent observations (from teachers) were used, direct clinical observations of the interactions between parent and child would most likely have captured nuances in the child's change process. The reliance on self-reported (i.e., lack of external) adherence checks is another study limitation.

Given the combination of strengths and limitations, the current findings provide evidence for the efficacy of an emotion-focused parental program on child symptom reduction. Our findings also provide support for the efficacy of a PE format, which may be better suited for implementation to community-based healthcare settings where families seek care, and therapists lack extensive training. As the EXP condition requires rather extensive training, experience and supervision, it may support program fidelity and standardization within the study – however, it may challenge its dissemination and likelihood for clinical services to adopt it and implement it effectively. Moreover, EFST without the evocative component could be more cost-efficient and easier to implement for training and service delivery, which has implications for likely effectiveness of EFST in the field.

Disclosure Statement

J.R.S has written several books on EFT and receives royalties for them. N.A., R.E., and J.R.S. designed the study and contributed to data analysis and the article. H.A.N.L, R.Z.O, and T.B.B. analyzed the data and contributed to the article. N.A and J.R.S. discloses a conflict of interest because they work in an organization that conducts training in the method.

Funding

We gratefully acknowledge funding from Research Council of Norway and The Norwegian Institute of Emotion Focused Therapy (NIEFT) (case ES621974/project 285852) and Institute of Psychological Counseling; Norges Forskningsr?d [case ES621974/project 285852]; Norwegian Institute of Emotion Focused Therapy [case ES621974/project 285852].

ORCID

Nadia Ansar http://orcid.org/0000-0003-2326-2110
Helene A. Nissen Lie http://orcid.org/0000-0003-2197-5942
Rune Zahl-Olsen http://orcid.org/0000-0002-3747-4184
Robert Elliott http://orcid.org/0000-0002-3527-3397
Jan Reidar Stiegler http://orcid.org/0000-0003-1007-6810

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school age forms and profiles*. Research center for children, youth, and families, University of Vermont.

Achenbach, T. M., McConaughy, S. H., Ivanova, M. Y., & Rescorla, L. A. (2011). *Manual for the ASEBA brief problem monitor (BPM)*. University of Vermont.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Ansar, N., Hjeltnes, A., Stige, S. H., Binder, P.-E., & Stiegler, J. R. (2021). Parenthood – lost and found: Exploring parents' experiences of receiving a program in emotion focused skills training. *Frontiers in Psychology*, *12*(559188). https://doi.org/10.3389/fpsyg.2021.559188

Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., & May, J. T. E. (2010). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide*. Oxford University Press. http://dx.doi.org/10.1093/med:psych/9780199772667.001.0001

Beurs, E., Barendregt, M., Heer, A., Duijn, E., Goeree, B. H., & Merks, A. (2016). Comparing methods to denote treatment outcome in clinical research and benchmarking mental health care. *Clinical Psychology and Psychoth*, *23*(4), 308–318. https://doi.org/10.1002/cpp.1954

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, *75*(5), 1523–1537. https://doi.org/10.1111/j.1467-8624.2004.00755.x

Buchanan-Pascall, S., Gray, K. M., Gordon, M., & Melvin, G. A. (2018). Systematic review and meta-analysis of parent group interventions for primary school children aged 4–12 years with externalizing and/or internalizing problems. *Child Psychiatry Human Development*, *49*(2), 244–267. https://doi.org/10.1007/s10578-017-0745-9

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Buuren, S. V. (2018). *Flexible imputation of missing data* (2nd ed.). Chapman & Hall/CRC Interdisciplinary Statistics.

Cain, D. J. (2010). The therapy process (Chap 4). In *Person-centered psychotherapies*. Washington, DC: American Psychological Association.

Catalano, R. F., & Kellogg, E. (2020). Fostering healthy mental, emotional, and behavioral development in children and youth: A national agenda. *Journal of Adolescent Health*, *66*(3), 265–267. https://doi.org/10.1016/j.jadohealth.2019.12.003

Cho, E., Bearman, S. K., Woo, R., Weisz, J. R., & Hawley, K. M. (2021). A second and third look at FIRST: Testing adaptations of a principle-guided youth psychotherapy. *Journal of Clinical Child & Adolescent Psychology*, *50*(6), 919–932. https://doi.org/10.1080/15374416.2020.1796678

Chorpita, B. C., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., & Krull, J. L. (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology*, *78*(4), 526–536. http://doi.org/10.1037/a0019602

Chorpita, B. F., Weisz, J. R., Daleiden, E. L., Schoenwald, S. K., Palinkas, L. A., Miranda, J., Higa, M. C. K., Nakamura, B. J., Austin, A., & Gibbons, R. D. (2013). Research network on youth mental health. Long-term outcomes for the child STEPs randomized effectiveness trial: A comparison of modular and standard treatment designs with usual care. *Journal of Consulting and Clinical Psychology*, *81*(6), 999–1009. https://doi.org/10.1037/a0034200

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum.

Compas, B. E., Jaser, S. S., Bettis, A. H., Watson, K. H., Gruhn, M. A., Dunbar, J. P., & Thigpen, J. C. (2017). Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychological Bulletin*, *143*(9), 939–991. http://dx.doi.org/10.1037/bul0000110

Damasio, A. (2000). *The feeling of what happens*. Harvest, Harcourt Inc.

de Haan, A. M., Boon, A. E., de Jong, J. T., Hoeve, M., & Vermeiren, R. R. (2013). A meta-analytic review on treatment dropout in child and adolescent outpatient mental health care. *Clinical Psychology Review*, *33*(5), 698–711. https://doi.org/10.1016/j.cpr.2013.04.005

Dolhanty, J., & Greenberg, L. (2009). Emotion-focused therapy in a case of anorexia nervosa. *Clinical Psychology & Psychotherapy*, *16*(4), 366–382. https://doi.org/10.1002/cpp.624

Dudley, W. N., Wickham, R., & Coombs, N. (2016). An introduction to survival statistics: Kaplan–Meier analysis. *Journal of the Advanced Practitioner in Oncology*, *7*(1), 91–100. http://doi.org/10.6004/jadpro.2016.7.1.8

Ehrenreich, M. J., Kennedy, S. M., Sherman, J. A., Bilek, E. L., Buzzella, B. A., Bennett, S. M., & Barlow, D. (2018). *Unified protocols for transdiagnostic treatment of emotional disorders in children and adolescents: Therapist guide*. Oxford University press.

Elliott, R., Watson, C. W., Goldman, R., & Greenberg, L. (2004). *Learning emotion-focused therapy: The process-experiential approach to change*. American Psychological Association.

Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: Findings from the national institute of child health and human development study of early childcare. *Developmental Psychology*, *46*(5), 1159–1175. https://doi.org/10.1037/a0020659

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Ford, Q. B., Lam, P., John, P. O., & Mauss, B. I. (2018). The psychological health benefits of accepting negative emotions and thoughts: Laboratory diary and longitudinal evidence. *Journal of Personality and Social Psychology*, *115*(6), 1075–1092. https://doi.org/10.1037/pspp0000157

Foroughe, M., Stillar, A., Goldstein, L., Dolhanty, J., Goodcase, E. T., & Lafrance, A. (2019). Brief emotion focused family therapy: An intervention for parents of children and adolescents with mental health issues. *Journal of Marital and Family Therapy*, *45*(3), 410–430. https://doi.org/10.1111/jmft.12351

Garland, A. F., Hawley, K. M., Brookman-Frazee, L., & Hurlburt, M. S. (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*(5), 505–514. https://doi.org/10.1097/CHI.0b013e31816765c2

Gendlin, E. T. (1996). *The practicing professional. Focusing-oriented psychotherapy: A manual of the experiential method*. The Guilford Press.

Goldman, R. N., Greenberg, L. S., & Angus, L. (2006). The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. *Psychotherapy Research*, *16*(5), 537–549. https://doi.org/10.1080/10503300600589456

Gottman, J. M., Katz, L. F., & Hooven, C. (1997). *Meta-emotion: How families communicate emotionally*. Erlbaum.

Greenberg, L., Rice, L., & Elliott, R. (1997). *Facilitating emotional change. The moment-by moment process*. Guilford Publications.

Greenberg, L., & Pascual-Leone, J. (2001). A dialectical constructivist view of the creation of personal meaning. *Journal of Counseling Psychology*, *14*(3), 165–186. https://doi.org/10.1080/10720530125970

Greenberg, L., Warwar, H. S., & Malcolm, M. W. (2008). Differential effects of emotion-focused therapy and psychoeducation in facilitating forgiveness and letting go of emotional injuries. *Journal of Counseling Psychology*, *55*(2), 185–196. https://doi.org/10.1037/0022-0167.55.2.185

Greenberg, L. (2015). Coaching for emotional intelligence in parenting (Chap 14). In *Emotion focused therapy: Coaching clients to work through their feelings* (2nd ed., pp. 307–328). Amer Psychol Association.

Hagen, A. K., Ogden, T., & Bjørnebekk, G. (2011). Treatment outcomes and mediators of parent management training: A one-year follow-up of children with conduct problems. *Journal of Clinical Child & Adolescent Psychology*, *40*(2), 165–178. https://doi.org/10.1080/15374416.2011.546050

Hagen, A. H. V., Austbø, B., Hjelmseth, V., & Dolhanty, J. (2019). *Emosjonsfokusert ferdighetstrening for foreldre*. Gyldendal.

Havighurst, S. S., Wilson, K. R., Harley, A. E., Kehoe, C., Efron, D., & Prior, M. R. (2013). "Tuning into kids": Reducing young children's behavior problems using an emotion coaching parenting program. *Child Psychiatry & Human Development*, *44*(2), 247–264. https://doi.org/10.1007/s10578-012-0322-1

Higa, C. K., & Chorpita, B. F. (2008). Evidence-based therapies: Translating research into practice (Chap 4). In R. G. Steele, T. D. Elkin, & M. C. Roberts, (Eds.), *Handbook of Evidence-Based Therapies for Children and Adolescents. Issues in Clinical Child Psychology* (pp. 45–61). Boston, MA: Springer. https://doi.org/10.1037/0022-0167.36.2.223

Horvath, A. O., & Greenberg, L. (1989). Development and validation of the working alliance inventory. *Journal of Counseling Psychology*, *36*(2), 223–233. https://doi.org/10.1037/0022-0167.36.2.223

Hox, J. (2010). *Multilevel analysis. Techniques and applications*. Erlbaum.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

IBM Corp. (2021). *IBM SPSS statistics for MacOS, version 28.0*.

Izard, C. E. (2002). Translating emotion theory and research into preventive interventions. *Psychological Bulletin*, *128*(5), 796–824. https://doi.org/10.1037/0033-2909.128.5.796

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical Psychology Review*, *25*(3), 341–363. http://dx.doi.org/10.1016/j.cpr.2004.12.004

Katz, L. F., Maliken, A. C., & Stettler, N. M. (2012). Parental meta-emotion philosophy: A review of research and theoretical framework. *Child Development Perspectives*, *6*(4), 417–422. https://doi.org/10.1111/j.1750-8606.2012.00244.x

LeDoux, J. E. (1993). Emotional networks in the brain. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 109–118). Guilford Press.

Leeuw, M., Goossens, M. E., de Vet, H. C. W., & Vlaeyen, J. W. (2009). The fidelity of treatment delivery can be assessed in treatment outcome studies: A successful illustration of behavioral medicine. *Journal of Clinical Epidemiology*, *62*(1), 81–90. https://doi.org/10.1016/j.jclinepi.2008.03.008

Lenhard, W., & Lenhard, A. (2016). *Computation of effect sizes*. Psychometrica. https://doi.org/10.13140/RG.2.2.17823.92329

Linero, R. A., & Daniels, M. J. (2018). Bayesian approaches for missing not at random outcome: The role of identifying restrictions. *Statistical Science*, *33*(2), 198–213. https://doi.org/10.1214/17-STS630

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies. Is it true that "everyone has won and all must have prizes"? *Archives of General Psychiatry*, *32*(8), 995–1008. https://doi.org/10.1001/archpsyc.1975.01760260059004

Mackler, J. S., Kelleher, R. T., Shanahan, L., Calkins, S. D., Keane, P. S., & O´Brien, M. (2015). Parenting stress, parental reactions, and externalizing behavior from ages 4 to 10. *Journal of Marriage and the Family*, *77*(2), 388–406. https://doi.org/10.1111/jomf.12163

Morris, A. S., Houltberg, B. J., Criss, M. M., & Bosler, C. D. (2017). Family context and psychopathology: The mediating role of children's emotion regulation (Chap 18). In *The Wiley handbook of developmental psychopathology* (pp. 365–389). Wiley & Sons. https://doi.org/10.1002/9781118554470.ch18

Pascual-Leone, A., & Greenberg, L. (2007). Emotional processing in experiential therapy: Why "the only way out is through. *Journal of Consulting and Clinical Psychology*, *75*(6), 875–887. https://doi.org/10.1037/0022-006X.75.6.875

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. SAGE Publications Inc.

Richter, J. (2015). Preliminary evidence for good psychometric properties of the Norwegian version of the brief problems monitor. *Nordic Journal of Psychiatry*, *69*(3), 174–178. https://doi.org/10.3109/08039488.2014.951070

Robinson, A. L., Dolhanty, J., & Greenberg, L. (2015). Emotion-focused family therapy for eating disorders in children and adolescents. *Clinical Psychology & Psychotherapy*, *22*(1), 75–82. https://doi.org/10.1002/cpp.1861

Schore, A. N. (1999). *Affect regulation and the origin of the self: The neurobiology of emotional development*. Classic Edition.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. Oxford University Press.

Smid, S. C., McNeish, D., Miočević, M., & Schoot, V. D. R. (2020). Bayesian versus frequentist estimation for structural equation models in small sample contexts: A systematic review. *A Multidisciplinary Journal*, *27*(1), 131–161. https://doi.org/10.1080/10705511.2019.1577140

Stiegler, J. R., Molde, H., & Schanche, E. (2018). Does the two-chair dialogue intervention facilitate processing of emotions more efficiently than basic Rogerian conditions? *European Journal of Psychotherapy & Counselling*, *20*(3), 337–355. https://doi.org/10.1080/13642537.2018.1495245

Strahan, J. E., Stillar, A., Files, N., Nash, P., Scarborough, J., & Lafrance, A. (2017). Increasing parental self-efficacy with emotion-focused family therapy for eating disorders: A process model. *Person-Centered and Experiential Psychotherapies*, *16*(3), 256–269. https://doi.org/10.1080/14779757.2017.1330703

Tomkins, S. (1963). *Affect, imagery and consciousness: The negative affects* (Vol. 1). Springer.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

Wang, Y., & Liu, Y. (2021). The development of internalizing and externalizing problems in primary school: Contributions of executive function and social competence. *Child Development*, *92*(3), 889–903. https://doi.org/10.1111/cdev.13462

Weisz, J., Bearman, S. K., Lauren, C., & Jensen-Doss, S. A. (2017). Initial test of a principle-guided approach to transdiagnostic psychotherapy with children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *46*(1), 44–58. https://doi.org/10.1080/15374416.2016.1163708

Whitney, D. G., & Peterson, M. D. (2019). US national and state-level prevalence of mental health disorders and disparities of mental health care use in children. *JAMA Pediatrics*, *173*(4), 389–391. https://doi.org/10.1001/jamapediatrics.2018.5399

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wilhelmsen, L. A., Aardal, H., Hjelmseth, V., Fyhn, K. H., & Stige, S. H. (2019). An EFFT workshop for parents with children 6-12 years increased parental self-efficacy. *Emotional & Behavioral Difficulties*, *25*(1), 29–45. https://doi.org/10.1080/13632752.2019.1655921

World Health Organization (WHO). (2005). *Child and adolescent mental health policies and plans*. Mental Health Policy & Service Guidance Package.