Behaviour Research and Therapy 176 (2024) 104520

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Optimizing indicated cognitive behavioral therapy to prevent child anxiety and depression: A cluster-randomized factorial trial

Carina Lisøy [a,b,*], Simon-Peter Neumer [a], Frode Adolfsen [d], Jo Magne Ingul [c], Lene-Mari Potulski Rasmussen [d], Tore Wentzel-Larsen [a], Joshua Patras [a], Anne Mari Sund [c], Kristin Ytreland [c], Trine Waaktaar [b], Solveig Holen [a], Anne Liv Askeland [a], Ida Mari Haug [d], Elisabeth Valmyr Bania [c], Kristin Martinsen [b,a]

[a] Regional Centre for Child and Adolescent Mental Health, Eastern and Southern Norway (RBUP Øst og Sør, Postboks 4623 Nydalen, 0405, Oslo, Norway
[b] Department of Psychology, University of Oslo, Harald Schjelderups hus Forskningsveien 3A, 0373, Oslo, Norway
[c] Regional Centre for Child and Youth Mental Health and Child Welfare – Central Norway, Department of Mental Health, Norwegian University of Science and Technology (NTNU, RKBU Midt-Norge, 7491, Trondheim, Norway
[d] Regional Centre for Child and Youth Mental Health and Child Welfare – Northern Norway, UiT - The Arctic University of Norway (UiT Norges Arktiske Universitet, Det helsevitenskapelige Fakultet, RKBU Nord, Varemottak-MH, Plan 6, Sykehusv. 44 9019 Tromsø, Norway

ARTICLE INFO

Keywords:
Anxiety
Depression
Cognitive behavioral therapy
Factorial design
Indicated prevention
Children

ABSTRACT

Identifying effective components can lead to interventions that are less resource-intensive and better suited for real-world needs. In this 2×2×2 cluster-randomized factorial trial (clinicaltrials.gov NCT04263558), we investigated the effects of three components of an indicated, transdiagnostic CBT intervention for children: 1) Intervention Delivery Format (child group format versus a blended format with group sessions and automated web-based sessions), 2) Parental Involvement in the intervention (group-based versus psychoeducational brochure), and 3) a Measurement Feedback System (MFS; on versus off). The intervention was delivered at schools in a group-based format. The participants ($N = 701$ children) were school children (age 8–12 years) with elevated symptoms of anxiety or depression, and their parents. The main outcomes were self-reported ($N = 633$) and parent-reported ($N = 725$) symptoms of child anxiety and depression post-intervention. The secondary outcome was children's user satisfaction with the intervention. We did not find significant main or interaction effects of Delivery Format, Parental Involvement, or MFS on children's symptom levels. There were no significant effects on children's user satisfaction. Results were compatible with retaining the least resource intensive combination (i.e., blended format, parental brochure, no MFS) in an optimized intervention.

1. Introduction

Anxiety and depression are common in children and adolescents (Polanczyk et al., 2015). To reduce the long-term burden of disease, effective prevention and early intervention is important. Research suggests that prevention should begin in primary school years and that early intervention should focus on children with elevated symptoms, rather than waiting until a diagnosable disorder is evident (Mulraney et al., 2021). Yet, few school-age children and adolescents receive services to reduce symptoms associated with two of the most prevalent mental disorders (Kern et al., 2017). Various approaches aim to improve the reach and implementation of evidence-based interventions, like transdiagnostic cognitive behavioral therapy (CBT). Nevertheless, the dissemination and implementation of such interventions are challenging (McHugh & Barlow, 2010). One obstacle may be their complexity. Typically, most attention has been paid to evaluating these interventions as bundled treatment packages to demonstrate effectiveness. This leaves a limited understanding of the active ingredients. Integrating a focus on identification of components that improve interventions can lead to more effective and less resource-intensive multicomponent

* Corresponding author. RBUP Øst og Sør Postboks 4623 Nydalen 0405 Oslo, Norway.
E-mail addresses: carina.lisoy@r-bup.no (C. Lisøy), simon-peter.neumer@r-bup.no (S.-P. Neumer), frode.adolfsen@uit.no (F. Adolfsen), jo.m.ingul@ntnu.no (J.M. Ingul), lene-mari.p.rasmussen@uit.no (L.-M. Potulski Rasmussen), tore.wentzel-larsen@r-bup.no (T. Wentzel-Larsen), joshua.patras@r-bup.no (J. Patras), anne.m.sund@ntnu.no (A.M. Sund), kristin.ytreland@ntnu.no (K. Ytreland), trine.waaktaar@psykologi.uio.no (T. Waaktaar), solveig.holen@r-bup.no (S. Holen), ala@r-bup.no (A.L. Askeland), ida.mari.haug@uit.no (I.M. Haug), elisabeth.bania@ntnu.no (E.V. Bania), kristin.martinsen@psykologi.uio.no (K. Martinsen).

https://doi.org/10.1016/j.brat.2024.104520
Received 15 March 2023; Received in revised form 12 December 2023; Accepted 13 March 2024
Available online 18 March 2024

interventions (Collins, 2018). This could enhance the uptake of evidence-based interventions and help close the gap between the need for health interventions and the provision of such efforts.

The Multiphase Optimization Strategy (MOST; Collins, 2018) is an innovative engineering-inspired framework which guides the optimization of interventions. In MOST, optimization is the process of identifying an intervention that provides the best expected outcome obtainable while balancing needs for efficiency, economy, or scalability. Interventions that balance effectiveness against these real-world criteria will be easier to implement as intended and could lead to a sustainable impact on a more efficient budget. A fundamental idea of MOST is that interventions can and should be optimized to meet desired criteria. MOST consists of three phases: preparation, optimization, and evaluation. Main activities in the preparation phase include selection of the components that are candidates for inclusion in the intervention, and the development of a conceptual model of the processes to be intervened on. A component, in this context, is any aspect of an intervention that can be singled out and set to different levels for testing. In the optimization phase of MOST, the aim is to build an optimized intervention by evaluating components and component levels. The selection is based on empirical data obtained via an optimization trial and chosen selection criteria, such as only including active components. Once an intervention is expected to be sufficiently effective, one moves on to the evaluation phase of MOST, where the effectiveness of the optimized intervention is assessed.

The present trial, Echo, was undertaken to optimize Emotion (Emotion: "Coping Kids" Managing Anxiety and Depression; Martinsen et al., 2017b), an indicated group CBT intervention for children with elevated symptoms of anxiety or depression. Emotion is school-based and intended for use in first-line services, including school health services. Although the effectiveness of Emotion was documented in a previous study (the Coping Kids trial; Martinsen et al., 2019), health services reported that offering Emotion was resource demanding and time-consuming (Rasmussen et al., 2020). We thus endeavored to optimize Emotion by investigating the effects of three candidate components to build a less resource-intensive intervention with maintained or enhanced effectiveness. The conceptual model depicting how the components should work is previously described by Ingul et al. (2021). This study presents results from the optimization trial.

The first component we examined was the intervention's delivery format, as it may impact the scalability and cost of psychosocial interventions. Individual face-to-face, group, and blended (combining face-to-face delivery with web-based delivery) formats are all robust intervention approaches for anxiety and depression (Crowe & McKay, 2017; Cuijpers et al., 2019; Rasing, 2021). A recent meta-analytic study of school-based preventive interventions found no significant effects of delivery format (face-to-face or digital) in effects for depression or anxiety programs (Werner-Seidler et al., 2021). From the health services perspective, group-based formats and blended formats are often less resource demanding per recipient than individual face-to-face interventions. Thus far, research on blended interventions has focused on the combination of *individual* face-to-face sessions combined with web-based sessions. Little is known about the potential for blended *group* format, although some documentation for adults exists (Schuster et al., 2018). A blended group format combined with an indicated transdiagnostic intervention for children has, to our knowledge, not previously been tested.

The second component we examined was parental involvement in the intervention. Child-focused interventions for anxiety and depression often involve parents because parental involvement components are expected to improve child outcomes (Wei & Kendall, 2014). However, the effect of parental involvement in child-focused treatment for internalizing problems, compared to child-alone interventions, is unclear. Several reviews of the effectiveness of active parental involvement have noted mixed findings (e.g., Breinholst et al., 2012; Crowe & McKay, 2017; Dardas et al., 2018; Thulin et al., 2014). The inconsistencies may be explained by various factors, such as whether the most important parental factors associated with child symptoms were targeted. Additionally, variations in how parents were involved, heterogeneity in the parental samples, and differences in children's age and problem type and severity may also contribute. We are not aware of any studies specifically investigating the effect of parental involvement in preventive interventions. The lack of studies highlights the importance of examining this factor. Integrating a separate parental involvement component to child-focused interventions demands extra time and resources from both healthcare services and parents, despite the potential benefits of the support parents offer. Therefore, such an inclusion may hinge on demonstrating its ability to yield enhanced child outcomes.

The third component we examined was a Measurement Feedback System (MFS). MFS typically consists of client self-assessment measures to regularly track processes and outcomes during an intervention, and update clinicians with reports that summarize the results. This systematic feedback is intended to facilitate clinical decision-making and individual tailoring of the intervention. In specialist care settings, MFS has improved therapy success rates for adults, especially for those who were either not improving or deteriorating while in therapy (Rognstad et al., 2022). Studies have also indicated that MFS can support service providers in enhancing co-operation with youth and help them stay focused on the aims of the intervention (Bickman et al., 2016; Tollefsen et al., 2020). A recent meta-analysis (Rognstad et al., 2022) found an effect-size for MFS of $d = 0.29$ for studies involving children and adolescents in therapy, but there are few studies where MFS has been applied in mental health settings with children, group-interventions, or in primary care settings and prevention.

The overall aim of the present study was to examine the individual and combined effects of the three candidate components (Intervention Delivery Format, Parental Involvement, and MFS) to optimize Emotion. The main outcomes were child- and parent-reported symptoms of children's anxiety (Multidimensional Anxiety Scale for Children; MASC-C/MASC-P; March et al., 1997) and depression (The Short Mood and Feelings Questionnaire; SMFQ-C/SMFQ-P; Angold et al., 1995). It is recommended to use data from several informants to obtain a fuller picture of children's symptoms because research has shown discordance in reports from youth and parents. Furthermore, both parents and children provide useful information about symptoms (Silverman & Ollendick, 2005). The study's secondary outcome was children's user satisfaction with the intervention (The Stigma and Evaluation sheet; Rapee et al., 2006). We used an "all active components" optimization criterion (Collins, 2018), where main effects of statistical and practical significance in the desired direction would entail a component being included in the intervention.

2. Method

2.1. Study design

The Echo study had a 2×2×2 full factorial design with between-cluster randomization at the school-level. As each of the three fully crossed factors had two levels (high/on versus low/off), there were eight

Table 1
The three two-level study factors forming the eight experimental conditions.

Condition	Factor 1	Factor 2	Factor 3
	Format	Parental Involvement	MFS
1	Group (+)	High (+)	Yes (+)
2	Group (+)	High (+)	No (−)
3	Group (+)	Low (−)	Yes (+)
4	Group (+)	Low (−)	No (−)
5	Blended (−)	High (+)	Yes (+)
6	Blended (−)	High (+)	No (−)
7	Blended (−)	Low (−)	Yes (+)
8	Blended (−)	Low (−)	No (−)

possible experimental conditions (see Table 1). A factorial design was chosen to obtain information about which of the candidate components (factors) to include in an optimized intervention. The design allows for performance assessment of the individual factors (i.e., the main effects), and whether the factors affect each other's performance (i.e., interaction effects; Collins, 2018). More specifically, as recommended for the optimization phase in MOST, a factorial 2^k design (k factors with 2 levels), was chosen. In this design, the power to detect main effects and interaction effects is about the same (Cohen, 1988). In addition, the design is economical with respect to subject requirements.

2.2. Intervention

The intervention Emotion (Martinsen et al., 2017a) is an indicated school-based CBT group intervention intended for children aged 8–12 years. Emotion is transdiagnostic and directed at changeable risk factors and theorized common underlying change mechanisms of child symptoms of anxiety or depression (Kendall et al., 2014; Martinsen et al., 2019). Based on service providers' feedback from the Coping Kids trial (Rasmussen et al., 2020), the intervention length in the optimization trial was abbreviated (20 child sessions reduced to 16, and 7 parent meetings reduced to 5). In the EMOTION program, there are two child sessions per week. The sessions follow a standard CBT structure with setting an agenda, chat time, and review of homework before focusing on the main activity for the session. All sessions end with new homework being assigned. The first half of EMOTION focuses on psychoeducation and learning new coping skills. This includes learning about feelings and seeing the connection between thoughts, feelings, bodily reactions, and actions. Coping strategies are then introduced to improve the children's ability to regulate emotions when feeling sad or worried. The second half emphasizes practicing newly learned coping skills, and employs behavioral activation and exposure-based experiments to address situations that were previously avoided or stressful (Martinsen, Kendall, Stark, O'Neil & Arora, 2019). Additionally, the program's second half targets improving children's self-esteem by establishing a more differentiated self-concept. During parental sessions, parents largely learn the same strategies as their children. Moreover, the parental session content emphasizes doing positive activities as a family, positive parenting, the effect of rewards, and guidance on supporting children during exposures and behavioral experiments. An overview of sessions can be found in Table 2. The therapeutic content within the CBT-intervention (e.g. psychoeducation, exposure, problem solving), which are also commonly referred to as CBT components, were not tested in the present study. When we talk about components in this paper, we are exclusively referring to the three study factors.

The first component we tested was the delivery format of child sessions. One level entailed all 16 sessions (two per week) delivered in a group format. The other level also involved 16 sessions (two per week), but half of the sessions were delivered face-to-face in groups (1 session per week, 8 sessions in total); the other half were delivered as automated web-based sessions (1 session per week, 8 sessions in total). Both the group version and the blended version included the most common practice elements from the best-established treatments for anxiety and depression, such as exposure, cognitive techniques, relaxation, psychoeducation, modelling, and problem solving (Collins & Dozois, 2008; Higa-McMillan et al., 2016). The digital sessions aimed to teach the same psychoeducational, cognitive, and behavioral skills to target the same mechanisms of change as the group version (see Ingul et al., 2021). The web-sessions were child-friendly and provided examples and possibilities for rehearsal of coping strategies through activities that required little writing. To allow for sustained and positive alliance between the group leaders and the children, about every other session was group and every other was digital (see also Table 2). The digital sessions took about 10–30 min to complete by children individually at home, hence the duration of each session was shorter than group sessions. Furthermore, the sessions could be completed when it suited the child between group sessions, and children could complete the web-based sessions as many times as they wanted. The intention behind the blended format was to reduce the number of hours required of health services to deliver the intervention, but still maintain the program's intensity from the child perspective. The choice to provide a blended

Table 2
Overview of the content for child sessions and parent sessions.

Content of child sessions	Group Format (16 group sessions)	Blended Format (8 group + 8 digital sessions)	Content of Parental group sessions
1: Introduction, establishing rules	Group	Group	
2: House of change/ conceptual model	Group	DIGGI	
3: Recognizing feeling, setting goals	Group	Group	1: Motivation/ goalsetting (parents only) Facilitate parent-child relationship
4: Emotion focused coping	Group	DIGGI	
5: Problem solving	Group	Group	
6: Thoughts influences feelings	Group	DIGGI	
Problem solving in real situations	(Not in group version)	Group	2: Positive parenting and reinforcement (parents and children)
7: Problem solving applied to anxiety	Group	DIGGI	
8: Cognitive change/ Behavioral experiments	Group	Group	
9: Cognitive change/ Behavioral experiments Positive self-concept	Group	DIGGI	
10: Cognitive change/ Behavioral experiments Positive self-concept	Group	Group	3: House of change/ behavioral experiments Recognition of emotions (parents only)
11: Cognitive change/ Behavioral experiments Positive self-concept	Group	DIGGI	
12: Cognitive change/ Behavioral experiments Positive self-concept	Group	DIGGI	
13: Cognitive change/ Behavioral experiments Positive self-concept	Group	Group	4: Cognitive restructuring/ behavioral experiments Parental engagement in problem-solving (parents and children)
14: Integrating knowledge, Behavioral experiments, Positive self-concept	Group	DIGGI	
15: Integration of coping skills Behavioral experiments	Group	(Not in DIGGI version)	
16: Closing up	Group	Group	5: Experiencing parental modeling behavior, summarizing the program (parents and children)

intervention, where every other session was face-to-face in a group rather than a more minimal form of support, was also driven by the aim of enhancing participant engagement and reducing dropout. There are challenges with engaging young people in digital programs (Werner--Seidler et al., 2021). Dropout rates in internet-based programs remain significant, albeit mitigatable with therapist support (Clarke et al., 2015; Hollis et al., 2017). There were some differences between group sessions and digital sessions, such as in-vivo exposure in face-to face sessions, while exposure tasks in the digital sessions were in sensu. While group leaders ran the face-to-face group meetings, children completed the web-based sessions on their own. This meant that group leaders were not available to children during digital sessions. Finally, all child groups used Virtual Reality (VR) headsets with 360-degree movies to facilitate behavioral experiments. VR was used during the face-to-face group meetings, which also allowed post-video discussions between children and group leaders. The VR library comprised 12 videos covering six scenarios, with between one and three levels of difficulty for graded exposure. The scenarios addressed social insecurity, rejection, assertiveness, fear-related symptoms, and sorrow. In the videos, there were pauses for participants to reflect on their emotional responses to the scenarios. As the VR movie library was limited, it is unlikely to have matched every participant's presenting problem, but it served as introductory step to practicing in vivo experiments.

The second component was parental involvement in the intervention. High parental involvement consisted of five parent group sessions, of which three were together with their children (session 2, 4 and 5). Meetings were focused on psychoeducation, positive parenting, and skills-training to support children in difficult situations. Low parental involvement consisted of parents receiving a self-help brochure based on the parent workbook (Martinsen et al., 2017c). For a person with average reading abilities, reading this brochure typically takes approximately 15 min. All groups were also recommended to hold a start-up parent information meeting to provide parents with practical information about participation and, if relevant, how to give children access to the digital resources. A start-up meeting presentation was provided by the project group. It did not contain any psychoeducational or therapeutical content, to separate the start-up meeting from the parent group sessions.

The third and final component tested was MFS. One level of the factor included MFS during the program, and the other level did not include MFS. An MFS application (the MittEcho app) was created for the current study. Children made personal goals, reported weekly symptom levels, and reported on personal goal progression using the app. Group leaders accessed this information by logging into a visual dashboard that tracked progress for each child in their group. Group leaders received a 2-h training in how to use MFS before the intervention started. Here, an important topic was child symptom development during the intervention and how decisions can be guided by children's feedback, inspired by the MATCH program (Chorpita & Weisz, 2009). An instruction manual and videos were also available for group leaders, and issues regarding MFS were also addressed in supervision. Group leaders were instructed to review the children's feedback once a week before group meetings, and especially monitoring for not-on-track cases. The group leaders introduced the MittEcho app in the groups and supported the children in formulating individual goals. More information about each study factor can also be found in Neumer et al. (2021) and Ingul et al. (2021).

2.3. Participants and procedure

A total of 58 public schools in urban and rural areas from 29 municipalities across Norway took part in the present study. The participating children ($N = 701$) attended fourth through sixth grade.

The recruitment of participants took place over five recruitment waves in the beginning of each school semester from spring 2020 to spring 2022. Importantly, as the intervention was indicated (i.e., targeting high-scorers of depression or anxiety symptoms), children were encouraged to sign up if they often felt sad or anxious. Children and parents at participating schools received oral and written information about the study. Children with valid parental consent completed electronic T1 screening surveys at school. Children scoring $\geq$1 standard deviation above the expected population mean for self-reported depressive (SMFQ-C inclusion $\geq$7 points; Angold et al., 2002; Rhew et al., 2010) or anxious (MASC-C inclusion girls $\geq$61 points; boys $\geq$54; Olason et al., 2004; Villabø et al., 2012) symptoms, or both, were eligible for the study and invited to join a child group. Child group sessions took place at school, either during school-hours or just after school-hours. Each session lasted approximately 45–60 min and were led by two group leaders. Group leaders were instructed to collaborate with the school when sessions occurred during school hours to ensure that participants would not miss out on the same school subject repeatedly. Group size was set to 3–7 children. If more than seven children were eligible, seven children were selected to be invited via an automated randomization procedure. Exclusion criteria were few; instances where children were not likely to benefit from an intervention in group format (e.g., language problems or severe developmental challenges) were considered individually. Parents of children who accepted the invitation to group participation received links to electronic parental surveys via email and SMS. Two parents could respond per child. Where there were parent group sessions, these were held after regular work hours. Children attended groups over approximately 8–10 weeks. After participation, children completed an electronic T2 post-intervention survey at school and their parents received links to an electronic T2 post-intervention survey.

The Echo study group leaders primarily worked in municipal first-line health services, such as school health services and pedagogical psychological services. They attended a 3-day training prior to delivering the intervention. Whilst running groups, group leaders received regular supervision. The supervisors, all at minimum therapists previously trained in CBT and in CBT supervision, participated in training for the Emotion program alongside group leaders to enhance uniform program understanding. They also received a supervisor manual, and they attended a minimum of three meetings with the other supervisors and project investigators during each semester of supervision to ensure similar focus on structure in supervision and to discuss supervision issues specific for Emotion. The aim was to enhance program implementation.

Procedures complied with the Helsinki Declaration, and the study was approved by the Regional Committees for Medical and Health Research Ethics (REK) - South East Norway (2019/1198) and The Norwegian Agency for Shared Services in Education and Research (152745). The study was registered with clinicaltrials.gov (NCT04263558). A study protocol was published which describes the study and procedures in further detail (Neumer et al., 2021).

2.4. Protocol deviation from study as planned

Commencement of child groups in the first recruitment waves coincided with Norwegian schools closing due to the Covid-19 pandemic in March 2020 (see also "study limitations"). As schools remained closed until May 2020, the study lost 30 groups (n = 190 children). Group leaders who attempted to deliver the intervention had to adapt it to the distancing rules (e.g., turning face-to face child groups into group video calls). The first participant wave was thus considered lost to the study (regarded as completely at random) and has not been included in data analyses or in the participant flow chart. Nine child groups (n = 53 children) recruited in the first participant wave postponed group start one full school semester and supplementary analyses were also conducted without these participants. The study started re-enrollment of participants for the second recruitment wave and recruited additional schools to amend for the loss of power. Data have been used as planned and with intent to treat (ITT) analyses for all participants from then on.

2.5. Randomization

Although the study's objectives pertained to the individual level, schools were the unit of randomization. This was due to the intervention's school and group-based nature and to avoid contamination effects. Upon recruitment of schools, they were randomized into an experimental condition, to which they recruited participants for child groups over the course of up to five recruitment waves. Schools were thus randomized after they were recruited but before any of the children were recruited, and group leaders were then notified about the allocation of the schools. Parents and children were given more information about their experimental condition in a start-up information meeting. This meeting was held after children scoring above the cut-off criteria had accepted the invitation to join a child group. Schools were assigned to one of eight conditions in a restricted manner to better preserve the balance property of the factorial design, as recommended by Nahum--Shani & Dziak, 2018. The randomization procedure was carried out in R, by one of the authors (T.W.-L.) and witnessed by two researchers. A function was specifically written for the Echo study for the randomization purpose (script at https://github.com/ToreWentzel-Larsen/ECHO-miscellaneous). Due to the nature of the intervention, it was not possible for the participants to be blinded to their intervention condition.

2.6. Sample size

Data were hierarchical, with cluster-randomization at the school level, which was considered in the power estimation. We assumed an ICC of 0.05 at school-level and set a two-sided 0.05 alpha with 80% power to detect differences with an effect size of $d = 0.25$ (see Cohen, 1988; Raudenbush & Liu, 2000). In the original sample size calculation (Neumer al., 2021), we estimated an average cluster size of 20, yielding a required sample size of 800 children within 40 schools. To mitigate the risk of schools not being able to recruit the presumed cluster size during the Covid-19 pandemic, we recruited additional schools, as increasing the number of clusters enhances efficiency in cluster-randomized trials (Hemming et al., 2017). With an assumed average cluster size of 12, the required sample size was 620 participants from 52 schools. Differential attrition between treatment conditions was not expected, so the allocation ratio between the eight conditions was equal.

2.7. Measures

Demographic variables. Basic demographic information (i.e., age, grade, sex) was provided upon completing parental consent forms. Parents of children who joined child groups after the baseline-screening completed surveys with more extensive demographic information, baseline measures, and questions regarding their and their child's participation in the study.

Multidimensional Anxiety Scale for Children (MASC; March et al., 1997*)* is a measure for symptoms of anxiety in children and youth aged 8–19 years, with 39 items rated on a 4-point Likert-scale. In the child self-report version (MASC-C) children report on their own anxiety symptoms, and in the parent-report version (MASC-P) parents report on their child's symptoms. MASC contains 4 subscales (Physical Symptoms, Harm Avoidance, Social Anxiety, and Separation/Panic) that aggregate into a Total Anxiety score. This total score can range from 0 to 117. Robust psychometric properties have been demonstrated internationally (e.g., Rynn et al., 2006), and in Norway (Martinsen et al., 2017c; Villabø et al., 2012). In the present study, Cronbach's alpha for the MASC-C subscales at T2 ranged from 0.67 (Harm Avoidance) to 0.87 (Physical Symptoms). Cronbach's alpha for the MASC-P subscales ranged from 0.72 (Separation/Panic) to 0.88 (Physical Symptoms and Social Anxiety).

The Short Mood and Feelings Questionnaire (SMFQ: Angold et al., 1995*)* is a measure to assess core symptoms of depression in children and youth 6–18 years. SMFQ is available in a child version (SMFQ-C) for children's self-report of depressive symptoms and in a parent version (SMFQ-P) for parents to report on their child's symptoms of depression. The SMFQ comprises 13 items rated on a 3-point Likert-scale. The total score can range from 0 to 26. SMFQ has high internal consistency (Thabrew et al., 2018) and is unidimensional (Lundervold et al., 2013; Messer et al., 1995). Cronbach's alpha in the present study was 0.89 for SMFQ-C, and 0.89 for SMFQ-P at T2.

The Stigma and Evaluation sheet (Rapee et al., 2006). The present study used five of the items from the user satisfaction subscale of The Stigma and Evaluation Sheet, which was part of the post-intervention (T2) child survey. These items are measured on a 10-point scale from 1 (=least) to 10 (=most) and are based on face validity to tap into satisfaction with a program: 1) Amount learned, 2) Program effectiveness, 3) Enjoyment of program, 4) Happy to do program, and 5) Would recommend to others. The mean of the five indicators was used to form an overall user satisfaction score; Cronbach's alpha for these items was 0.87.

Attendance and compliance to the group sessions were registered by electronic checklists, while user data were collected for the digital sessions and the MittEcho feedback app. Group leaders also reported estimated time spent on preparing and conducting the intervention after each group session. Parents in the low parental involvement conditions responded to questions regarding receiving and reading the parental brochure, as well as the usefulness of the brochure.

2.8. Statistical analyses

To assess within-group differences in symptoms from pre-to post-intervention, we used mixed effects models for children, which accounted for the nesting of participants both within time and within group. For parents, mixed effects models accounted for the nesting of participants within time, within family and within group.

We assessed differential attrition among children via logistic regression, where a dummy-coded "stayers versus leavers" variable for child dropout in the intervention was regressed on experimental condition.

In the primary analyses, we investigated main and interaction effects of the three study factors while accounting for hierarchical data using linear mixed effects models. Models for each of the primary outcomes, child-reported symptoms (MASC-C and SMFQ-C) and parent-reported child symptoms (MASC-P and SMFQ-P), and the secondary outcome (User Satisfaction) were run separately. The main independent variables were effect-coded factors for the three factors and their products, to enable estimation of the main and interaction effects, as described in more detail below. Random effects in models with child self-report data were initially included representing random differences between 1) municipalities, 2) schools within municipality, and 3) child intervention groups within school. The random structure was simplified for model stability, when necessary (e.g., Barr et al., 2013; Pinheiro & Bates, 2000). Random effects in models with parent-reported data were run in accordance with the models for child data, but as two parents could respond per child, it included an additional level of random effects to account for data nested within families. Since the main and interaction effects are represented by dichotomous variables, no random slopes were included. Unadjusted models were estimated first. We then adjusted the models in two steps, using the same random effect structure. In the first step, we controlled for sex and children's age (years and months). In the second step, we also included centered MASC and SMFQ baseline scores as control variables. To estimate effects, we used effect coding (see effect coded design matrix in Table 3) as described by Kugler et al. (2018). With effect coding, the fixed effects coefficients in the model are equivalent to the classically defined main effects and interactions, except for a scaling constant of 2 (which does not affect p-values). For factors with two levels, multiplying the regression coefficient by this scaling constant produces the estimated main and

Table 3
The 8 experimental conditions along with the effect-coded design matrix.

Condition	Components			Effect coding matrix						
	Factor 1 Format	Factor 2 Parents	Factor 3 MFS	Format	Parents	MFS	Format x Parents	Format x MFS	Parents x MFS	Format x Parents x MFS
1	Group	High	Yes	+1	+1	+1	+1	+1	+1	+1
2	Group	High	No	+1	+1	−1	+1	−1	−1	−1
3	Group	Low	Yes	+1	−1	+1	−1	+1	−1	−1
4	Group	Low	No	+1	−1	−1	−1	−1	+1	+1
5	Blended	High	Yes	−1	+1	+1	−1	−1	+1	−1
6	Blended	High	No	−1	+1	−1	−1	+1	−1	+1
7	Blended	Low	Yes	−1	−1	+1	+1	−1	−1	+1
8	Blended	Low	No	−1	−1	−1	+1	+1	+1	+1

interaction effects. The main effects, then, were interpreted as the estimate by which receiving the high level of a factor changes the outcome variable compared to receiving the lower level. Interaction effects were interpreted as the difference in the effect of a factor or a combination of factors across the levels of a second/third factor.

For mixed effects analyses, we used R version 4.2.1. and the package nlme. For other analyses, we used IBM SPSS v.28.0.

3. Results

Of 9461 children attending the grades informed about the study, 1364 (14%) children had informed parental consent to participate. The

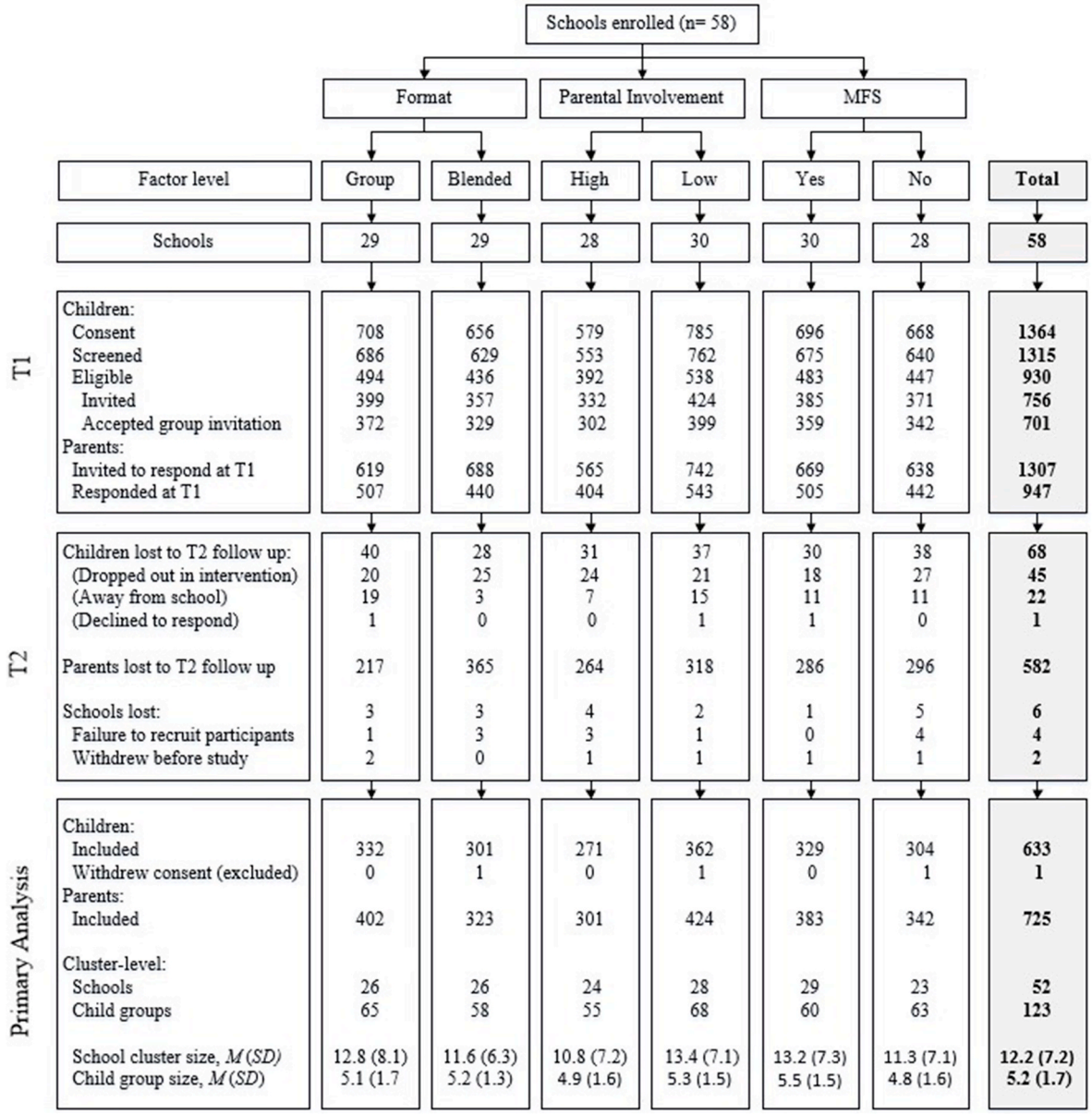

Fig. 1. Participant flow.

flow of participants can be seen in Fig. 1 (for participant flow by the eight experimental conditions, see Appendix, Figure A1). Of the 1364 children with consent, 1315 (96%) participated in T1 screening. Overall, 930 (71%) of the screened children scored above the cut-off criteria. Two children who scored above the cut-off criteria were excluded from the intervention due to intervention exclusion criteria: unlikely to benefit from a group intervention. As there was a group size limitation, 172 randomly selected eligible children were *not* invited to the intervention, while 756 of the eligible children were invited to the intervention. Of the invitees, 701 (93%) children accepted the invitation to join a group. During the intervention, 45 children (6%) dropped out of the study. The most frequent reason was that the child no longer wanted to participate. The drop-out rates did not significantly differ between the eight experimental conditions ($p = 0.774$). At T2, 90% of the children who accepted group invitation and enrolled in the intervention completed the primary outcome measures MASC-C ($n = 633$) and SMFQ-C ($n = 632$). A total of 1307 parents (parents of the 701 children who enrolled in the intervention) were asked to complete a parent survey. At T1, 947 (72%) parents responded to the survey, reporting on 622 of the children. At T2, 725 parents (55%) responded, reporting on 528 children.

3.1. Participant characteristics

The participants' age ranged from 8.5 to 12.9 years ($M = 10.6$, $SD = 0.7$; see Table 4). There were 393 (62%) girls and 240 (38%) boys. Primary analyses included 633 children. Of these, 614 (97%) of the children had at least one parent who provided baseline demographic information. The majority of children (96%) were born in Norway, and 69% lived with both their parents. As both caretakers could respond, the overall parental response rate was 75% (59% mothers, 40% fathers, and 1% other). Among parents who responded, the mean parental age was 41.8 years ($SD = 5.8$). The proportion of parents who were not employed (9%) was comparable to the proportion outside the workforce among this age group in the Norwegian population (13%; Statistics Norway, 2023). The proportion of parents with attained tertiary education (66%) was larger than that of the general population (50%).

3.2. Attendance and compliance

Children assigned to the group format attended 90% of their 16 group sessions on average ($M = 14.5$, $SD = 1.6$). In the blended format conditions, children attended 89% of their 8 group sessions ($M = 7.1$, $SD = 1.3$). In the high parental involvement conditions (5 parent meetings), the parents attended 79% of the sessions ($M = 3.9$, $SD = 1.1$). On average, mothers attended 1.9 sessions without the other caregiver ($SD = 1.7$) and fathers attended 0.6 sessions ($SD = 1.0$) without the other caregiver. Both caregivers were present in 1.5 sessions ($SD = 1.6$). The relative attendance rates were similar to the Coping Kids effectiveness trial where, on average, children participated in 90% of child sessions and parents participated in 80% of parent sessions (Martinsen, 2019). Parents in the low parental involvement conditions received a parental brochure, and 86% of the parents who responded to the survey reported that they had received it. Of those who reported having received the brochure, 98% reported reading it. The parents who reported receiving the brochure also responded to a question about whether they had gained a better understanding of their child's emotional challenges after reading it, where percentages endorsing the response categories were *Not at all* = 4%, *A little* = 39%, *Neither/nor* = 24%, *Somewhat* = 29%, and *Very* = 5%.

User data for the children with blended format indicated that children accessed 59% of the eight digital sessions ($M = 4.7$, $SD = 2.7$). The proportion of children who accessed each session declined throughout the intervention and ranged from 90% to 35% (90%, 80%, 71%, 64%, 54%, 42%, 35% and 35%, for session 1 through 8, respectively).

Of the 329 children assigned to use MFS, 285 (87%) set at least one

Table 4
Characteristics of Participants along with p-values from Mixed effects models for T1 and T2 differences within each factor level.

		Format		Parental Involvement		MFS	
	All	Group	Blended	High	Low	Yes	No
	($n = 633$)	($n = 332$)	($n = 301$)	($n = 271$)	($n = 362$)	($n = 329$)	($n = 304$)
Age	10.6 (0.7)	10.6 (0.7)	10.6 (0.7)	10.56 (0.7)	10.59 (0.7)	10.6 (0.71)	10.53 (0.7)
Years, *M* (*SD*)							
Sex, *n* (%)							
Boys	240 (38%)	128 (39%)	112 (37%)	113 (42%)	127 (35%)	131 (40%)	109 (36%)
Girls	393 (62%)	204 (61%)	189 (63%)	158 (58%)	235 (65%)	198 (60%)	195 (64%)
School-Grade, *n* (%)							
4th	111 (17%)	67 (20%)	45 (15%)	53 (20%)	58 (16%)	46 (14%)	65 (21%)
5th	328 (52%)	180 (54%)	153 (51%)	134 (49%)	202 (56%)	172 (52%)	156 (51%)
6th	194 (31%)	85 (26%)	103 (34%)	84 (31%)	102 (28%)	111 (34%)	83 (27%)
MASC-C	($n = 633$)	($n = 332$)	($n = 301$)	($n = 271$)	($n = 362$)	($n = 329$)	($n = 304$)
T1, *M* (*SD*)	69.70 (14.63)	68.02 (14.28)	71.55 (14.82)	69.92 (14.88)	69.54 (14.76)	69.22 (15.14)	70.22 (14.06)
T2, *M* (*SD*)	58.30 (17.22)	57.06 (16.75)	59.68 (17.66)	58.90 (17.09)	57.86 (17.33)	56.62 (17.41)	60.13 (16.86)
p		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
MASC-P	($n = 648$)	($n = 359$)	($n = 289$)	($n = 265$)	($n = 383$)	($n = 352$)	($n = 296$)
T1, *M* (*SD*)	51.73 (14.91)	51.84 (15.13)	51.60 (14.65)	52.28 (15.54)	51.36 (14.46)	51.0 (15.1)	52.62 (14.58)
	($n = 725$)	($n = 402$)	($n = 323$)	($n = 301$)	($n = 424$)	($n = 383$)	($n = 342$)
T2, *M* (*SD*)	48.81 (15.08)	48.91 (15.07)	48.69 (15.12)	49.92 (14.30)	48.02 (15.59)	48.9 (15.6)	48.72 (14.56)
p		<0.001	<0.001	0.001	<0.001	<0.001	<0.001
SMFQ-C	($n = 632$)	($n = 332$)	($n = 300$)	($n = 271$)	($n = 361$)	($n = 328$)	($n = 304$)
T1, *M* (*SD*)	11.53 (5.42)	11.25 (4.98)	11.84 (5.87)	11.27 (5.34)	11.73 (5.48)	11.59 (5.48)	11.47 (5.37)
T2, *M* (*SD*)	9.00 (6.06)	8.55 (7.95)	9.49 (6.35)	8.81 (5.70)	9.14 (6.32)	9.07 (5.94)	8.92 (6.20)
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
SMFQ-P	($n = 646$)	($n = 352$)	($n = 280$)	($n = 254$)	($n = 378$)	($n = 341$)	($n = 291$)
T1, *M* (*SD*)	6.62 (4.82)	6.38 (4.75)	6.85 (4.97)	6.7 (4.9)	6.52 (4.85)	6.7 (4.9)	6.4 (4.8)
	($n = 709$)	($n = 395$)	($n = 314$)	($n = 290$)	($n = 419$)	($n = 372$)	($n = 337$)
T2, *M* (*SD*)	5.64 (5.00)	5.64 (5.02)	5.65 (4.97)	5.7 (5.1)	5.6 (5.0)	6.0 (5.2)	5.3 (4.7)
p		< 0.001[a]	< 0.001[a]	< 0.001[a]	<0.001	< 0.001[a]	<0.001
User Satisfaction	($n = 628$)	($n = 330$)	($n = 398$)	($n = 269$)	($n = 359$)	($n = 326$)	($n = 302$)
T2, *M* (*SD*)	7.62 (1.94)	7.76 (1.92)	7.48 (1.96)	7.69 (2.03)	7.57 (1.87)	7.55 (2.05)	7.70 (1.80)

Note: MFS = Measurement Feedback System, MASC = Multidimensional Anxiety Scale for Children, SMFQ = Short Mood and Feelings Questionnaire.
[a] Mixed effects model with outcome nested within time, family and group was unstable; group level removed.

personal goal in the app, and 267 (81%) gave feedback using the app at least once. In total, an average of 2.3 ($SD = 0.8$) goals per child were set in the app. The app was used an average of 4.2 ($SD = 3.0$) times per child to respond to symptom measures. The intervention duration, or number of weeks to complete the intervention, varied somewhat between groups. Considering this, children only completed the symptom measures 45% of the possible number of times. Data on MFS group leader usage showed that they only viewed feedback in 52% of the weeks during the intervention.

Group leaders indicated how much time they estimated having spent preparing the sessions and how long it took to conduct the sessions at each participant wave (Table 5). As answers were given in response categories per session, we employed "proxy" values based on the midpoints of the categories, where: *less than 1 h* was interpreted as 0.5 h, *1–2 h* as 1.5 h, *2–3 h* as 2.5 h and *more than 3 h* as 3.5 h. Consequently, the reported hours spent on preparation and intervention delivery provide an approximate estimation of the time invested per group leader, rather than precise figures. Each group had a minimum of two leaders. Notably, group leaders in the blended format (8 group sessions) without parental sessions used the least amount of time on intervention provision, while group leaders in the format with 16 group sessions and 5 parental groups used the most amount of time.

3.3. Symptom levels at T1 and T2

Of the 633 children in the primary analyses, 421 (67%) were included based on both elevated self-reported anxiety and depression scores, while 114 (18%) were included based on depression scores alone, and 98 (16%) based on anxiety scores alone. Mean anxiety and depression levels are shown in Table 4. Notably, children reported higher mean symptom levels than parents reported for the child, particularly at T1. Both parent-reported and child-reported symptoms declined from pre-to post-intervention within each of the factor levels, and mixed effects models showed that the reduction was significant within each factor level. Mean levels of symptoms also declined within the eight experimental conditions (Appendix table A1). However, mixed effect models showed that the difference was significant within all conditions for child-reported symptoms (MASC-C/SMFQ-C), but not consistently within all conditions for parent-reported symptoms (MASC-P/SMFQ-P).

3.4. Effects on levels of anxiety and depression

In the mixed models examining main and interaction effects for MASC-C and SMFQ-C, the random differences between municipalities and schools had to be deleted to achieve model stability. Random differences between groups could be kept in the models. The ICC at group level was 0.069 for MASC-C and 0.096 for SMFQ-C. There were 633 observations across 123 groups for MASC-C. For SMFQ-C, there were 632 observations across 123 groups, as one respondent had not completed the survey. Models for parent-reported child outcomes (MASC-P and SMFQ-P) were attempted using random effects for group and family level. Random differences between groups could be kept in the model for MASC-P, but removed for SMFQ-P. For MASC-P, the ICC was 0.024 at the group level and 0.471 at the family level. For SMFQ-P, the ICC was. 495 at the family level. MASC-P had 725 parent observations of 528 children in 121 child groups. Of these, 646 parents of 486 individual children had reported baseline measures at T1 and could thus be included in the MASC-P analysis with baseline adjustment. SMFQ-P had 709 parental observations of 519 children. Of these, 632 parents had provided baseline measures. Therefore, 709 participants were included in unadjusted models and models adjusted for sex and age, while 632 parents were included in the SMFQ-P analysis with baseline adjustment. The effects can be seen in Table 6 and in Table 7.

There were no significant main or interaction effects. An absence of significant main effects can be interpreted as no evidence that the high/on (+) levels of the factors were superior to the low/off (−) levels, and non-significant interaction effects can be interpreted as no evidence that the components affected each other's performance. The point estimates represent the best estimates of the true differences. However, it is essential to note that non-significant findings do not necessarily mean that effects can be excluded, and as the 95% confidence bounds could represent effects of practical relevance, we discuss these. The lower confidence bounds represent uncertainty surrounding effects in the desired direction, as lower symptom-levels would be favorable, while the upper bounds represent uncertainty surrounding negative effects and higher symptom-levels.

The main effects for "Delivery Format" on child symptoms of anxiety and depression, were not statistically significant. As seen from the lower confidence bounds in the final, adjusted models for MASC-C and SMFQ-C (see Table 6), effects of practical relevance downwards (i.e., superiority of group format over blended format) could not be entirely excluded. However, effect sizes at the lower confidence bound would be considered small (MASC-C lower bound = −4.05, SMFQ-C lower bound = −1.64). For both parent-reported outcomes (MASC-P/SMFQ-P), effects of practical relevance downwards could be entirely excluded (Table 7).

For "parental involvement", the main effects on MASC-C/P and SMFQ-C/P were not statistically significant. As seen in Tables 6 and 7, positive effects of practical relevance of high parental involvement could be excluded altogether by the lower confidence interval bounds, for both child-reported and parent-reported symptoms.

Finally, there were no significant main effects for MFS. As seen in Tables 6 and 7, positive effects of practical relevance on anxious symptoms could be excluded by the lower confidence bound for parent-reported symptoms, but not for child-reported symptoms (SMFQ-C lower bound = −5.29). Positive effects of practical relevance on symptoms of depression could be entirely excluded for both child (SMFQ-C) and parent-reported (SMFQ-P) symptoms.

In supplementary analyses, we re-ran analyses without the 9 child groups with delayed study entry due to the Covid-19- pandemic ($n = 53$ children, 59 parents). There were some differences between models with and without these participants (see Appendix, Table A2 and A.3). In analyses that excluded the 59 parents of children who had a delayed

Table 5
Estimated average time used for preparing and conducting the sessions, per group leader.

	GHY	GHN	GLY	GLN	BHY	BHN	BLY	BLN
Format	+	+	+	+	−	−	−	−
Parental Involvement	+	+	−	−	+	+	−	−
MFS	+	−	+	−	+	−	+	−
Number of sessions[a]	21	21	16	16	13	13	8	8
Hours per group leader	63.0	60.9	46.4	46.4	37.7	45.5	23.2	24.8

Note: GHY = Group format + High Parental Involvement + MFS; GHN = Group format + High Parental Involvement; GLY = Group format + Low Parental Involvement + MFS; GLN = Group Format + Low Parental Involvement; BHY= Blended Format + High Parental Involvement + MFS; BHN= Blended Format + High Parental Involvement; BLY= Blended Format + Low Parental Involvement + MFS; BLN= Blended Format + Low Parental Involvement.

[a] Total number of sessions provided by health services in the experimental condition (child group sessions plus parent group sessions where applicable).

Table 6
Main and interaction effects for children's self-reported symptoms of anxiety (MASC-C) and depression (SMFQ-C), in models that were 1) unadjusted, 2) adjusted for age and sex, and 3) adjusted for age, sex and baseline scores.

	MASC-C (n = 633)			SMFQ-C (n = 632)		
	Estimate	95% CI	p	Estimate	95% CI	p
Unadjusted:						
Format	−2.63	[-5.78, 0.51]	0.100	−0.84	[-2.02, 0.34]	0.162
Parental Involvement	1.39	[-1.75, 4.53]	0.382	−0.25	[-1.43, 0.93]	0.680
MFS	−2.87	[-6.01, 0.27]	0.073	0.49	[-0.69, 1.67]	0.414
Format x Parental Involvement	0.01	[-3.14, 3.15]	0.997	0.74	[-0.44, 1.92]	0.217
Format x MFS	0.01	[-3.13, 3.16]	0.992	0.28	[-0.90, 1.46]	0.642
Parental Involvement x MFS	0.73	[-2.41, 3.87]	0.646	0.32	[-0.86, 1.50]	0.590
Format x Parental Involvement x MFS	−1.74	[-4.88, 1.41]	0.276	−0.53	[-1.71, 0.65]	0.376
Adjusted for age and sex:						
Format	−2.62	[-5.63, 0.39]	0.087	−0.82	[-1.10, 0.36]	0.171
Parental Involvement	1.82	[-1.19, 4.84]	0.234	−0.17	[-1.35, 1.00]	0.773
MFS	−2.44	[-5.45, 0.58]	0.112	0.52	[-0.65, 1.70]	0.380
Format x Parental Involvement	0.27	[-2.75, 3.28]	0.861	0.79	[-0.39, 1.97]	0.186
Format x MFS	−0.25	[-3.26, 2.76]	0.869	0.22	[-0.95, 1.40]	0.708
Parents x MFS	0.96	[-2.06, 3.98]	0.529	0.39	[-0.79, 1.57]	0.511
Format x Parental Involvement x MFS	−2.05	[-5.07, 0.97]	0.182	−0.61	[-1.79, 0.57]	0.305
Adjusted for age, sex and baseline scores:						
Format	−1.17	[-4.05, 1.70]	0.421	−0.60	[-1.64, 0.44]	0.256
Par. Involvement	1.45	[-1.42, 4.31]	0.320	−0.02	[-1.06, 1.02]	0.969
MFS	−2.42	[-5.29, 0.44]	0.097	0.40	[-0.64, 1.43]	0.449
Format x Parental Involvement	−0.62	[-3.48, 2.25]	0.672	0.73	[-0.31, 1.77]	0.166
Format x MFS	−0.18	[-3.04, 2.68]	0.903	0.30	[-0.74, 1.33]	0.571
Parental Involvement x MFS	1.04	[-1.82, 3.91]	0.473	0.38	[-0.66, 1.42]	0.469
Format x Parental Involvement x MFS	−1.52	[-4.39, 1.35]	0.296	−0.66	[-1.70, 0.38]	0.212

Estimate denotes estimated effects (2 times the unstandardized regression coefficient (*b**2)).
Format = Group format versus blended format, Parental Involvement = High Parental Involvement versus Low Parental Involvement, MFS = MFS (Measurement Feedback System) versus no MFS.
MASC-C = Multidimensional Anxiety Scale for Children - Child version, SMFQ = Short Mood and Feelings Questionnaire – Child version.

Table 7
Main and interaction effects for parent-reported child symptoms of anxiety (MASC-P) and depression (SMFQ-P), in models that were 1) unadjusted, 2) adjusted for age and sex, and 3) adjusted for age, sex and baseline scores.

	MASC-P			SMFQ-P		
	Estimate	95% CI	p	Estimate	95% CI	p
Unadjusted:		(N = 725)			(N = 709)	
Format	−0.11	[-2.81, 2.58]	0.933	0.06	[-0.80, 0.92]	0.898
Parental Involvement	1.81	[-0.88, 4.51]	0.185	0.14	[-0.72, 1.00]	0.750
MFS	0.02	[-2.68, 2.71]	0.990	0.73	[-0.13, 1.59]	0.097
Format x Parental Involvement	−0.23	[-2.92, 2.47]	0.867	0.40	[-0.46, 1.26]	0.364
Format x MFS	0.31	[-2.39, 3.00]	0.822	0.13	[-0.73, 0.99]	0.771
Parental Involvement x MFS	−0.82	[-3.51, 1.88]	0.548	0.41	[-0.45, 1.27]	0.354
Format x Parental Involvement x MFS	−2.18	[-4.88, 1.51]	0.111	−0.22	[-1.08, 0.64]	0.621
Adjusted for age and sex:		(N = 725)			(N = 709)	
Format	−0.08	[-2.79, 2.63]	0.952	0.05	[-0.81, 0.91]	0.906
Parental Involvement	1.92	[-0.80, 4.63]	0.164	0.09	[-0.77, 0.95]	0.842
MFS	0.11	[-2.60, 2.83]	0.934	0.64	[-0.22, 1.50]	0.146
Format x Parental Involvement	−0.08	[-2.79, 2.64]	0.955	0.36	[-0.50, 1.22]	0.410
Format x MFS	0.21	[-2.50, 2.92]	0.879	0.17	[-0.69, 1.02]	0.704
Parents x MFS	−0.68	[-3.39, 2.03]	0.621	0.36	[-0.50, 1.22]	0.409
Format x Parental Involvement x MFS	−2.45	[-5.18, 0.28]	0.078	−0.14	[-1.01, 0.72]	0.742
Adjusted for age, sex and baseline scores:		(N = 646)			(N = 632)	
Format	−0.61	[-2.54, 1.32]	0.533	0.10	[-0.56, 0.76]	0.763
Par. Involvement	1.34	[-0.59, 3.27]	0.171	0.02	[-0.64, 0.67]	0.960
MFS	0.67	[-1.27, 2.60]	0.497	0.61	[-0.05, 1.27]	0.070
Format x Parental Involvement	−0.52	[-2.44, 1.41]	0.598	0.34	[-0.32, 1.00]	0.309
Format x MFS	−0.56	[-2.49, 1.37]	0.564	−0.11	[-0.77, 0.55]	0.751
Parental Involvement x MFS	0.04	[-1.89, 1.97]	0.966	0.55	[-0.11, 1.21]	0.100
Format x Parental Involvement x MFS	−0.61	[-2.55, 1.33]	0.534	−0.12	[-0.78, 0.54]	0.725

Note: Estimate denotes estimated effects (2 times the unstandardized regression coefficient (*b**2)).
Format = Group format versus blended format, Parental Involvement = High Parental Involvement versus Low Parental Involvement, MFS = MFS (Measurement Feedback System) versus no MFS.

MASC-P = Multidimensional Anxiety Scale for Children – Parent version, SMFQ-P = Short Mood and Feelings Questionnaire – Parent version.

study entry, there was a small, but statistically significant, negative main effect of MFS on SMFQ-P when adjusting for baseline (estimate = 0.82, 95% CI [0.14–1.51], $p = 0.019$). Furthermore, there was a significant main effect of MFS in the unadjusted MASC-C model, but this effect was no longer statistically significant once the model was adjusted.

3.5. User satisfaction

Children's user satisfaction with Emotion was high (Table 4), with an overall mean of 7.62 points, out of 10 possible. Five children did not complete the entire user satisfaction survey ($N = 628$). In the mixed models examining main and interaction effects for user satisfaction, the random differences between municipalities had to be deleted to achieve model stability. Schools and groups within schools could be kept in the models. There were no statistically significant main effects or interaction effects (Table 8). Thus, children's satisfaction with Emotion did not differ significantly by factor levels or by the combination of components they received.

4. Discussion

The main objective of the Echo study was to optimize the group-based intervention Emotion. Although a previous study indicated positive effects for children participating in Emotion, some users, especially the health services, found it time- and resource consuming (Rasmussen et al., 2020). To address this issue, a factorial experiment was carried out.

Of the children screened in the Echo study, 71% scored above the anxiety or depression symptom inclusion criteria for participation in the study, where the majority (67%) reported elevated symptoms of both disorders. As depression and anxiety are highly comorbid, this was unsurprising. It highlights the usefulness of transdiagnostic approaches both when assessing and addressing child emotional problems, although research also commonly shows beneficial spill-over effects for interventions targeting only one type of emotional problem. That is, interventions targeting one condition (e.g. depression), often lead to improvements in comorbid conditions (e.g. anxiety) even if they were not the primary focus of the intervention (Garber & Weersing, 2010 Dec).

Children reported higher levels of anxiety and depression than their parents observed. This finding is in line with findings from our previous study (Martinsen et al., 2019). Symptoms of both anxiety and depression declined from pre-to post-intervention, within all the study's six factor levels. Symptoms also decreased within all eight experimental conditions, with consistent statistical significance for children's self-reported symptoms. For parent-reported child symptoms, the difference was non-significant in some instances. The symptom reduction was expected, as all participants received a CBT-based intervention. Emotion has previously demonstrated positive effects on both anxious and depressive symptoms compared to a control group (Martinsen et al., 2019), though it is important to keep in mind that the present version was slightly abbreviated from the original version. Although symptom levels moved in the desired direction, effectiveness of the intervention cannot be established using the present factorial design, as the study design only allows for an investigation of effects of the factors. Without a control group, the results could be a regression towards the mean, which in indicated prevention literature is expected (Linden, 2013).

Table 8
Main and Interaction Effects on Children's User Satisfaction with the intervention (n = 628).

	Estimate	95% CI	*p*
Format	0.20	[-0.24, 0.65]	0.358
Parental Involvement	0.14	[-0.30, 0.59]	0.519
MFS	−0.23	[-0.68, 0.21]	0.299
Format x Parental Involvement	−0.14	[-0.59, 0.30]	0.518
Format x MFS	−0.08	[-0.52, 0.37]	0.729
Parents x MFS	−0.37	[-0.82, 0.07]	0.100
Format x Parental Involvement x MFS	−0.19	[-0.63, 0.26]	0.399

Estimate denotes estimated effects (2 times the unstandardized regression coefficient (b*2))Format = Group format versus blended format, Parental Involvement = High Parental Involvement versus Low Parental Involvement, MFS = MFS (Measurement Feedback System) versus no MFS.

We did not find evidence that the type of delivery format (group versus a blended group-and-digital format) had an effect on children's symptom levels. This finding is in line with the most recent meta-analytic evidence on the effects of delivery format in school-based interventions (Werner-Seidler et al., 2021). The study's conceptual model proposed that both delivery formats of the intervention would improve children's emotion regulation, problem solving and relaxation skills, increase approach behavior, and reduce negative self-talk (Ingul et al., 2021). To target these suggested mechanisms of change, both the group version and the blended version included the most common elements in well-documented interventions for anxiety and depression. While the child attendance rates for group sessions were high, the declining usage rates of the digital sessions during the intervention period call for strategies to improve usage. A mitigation strategy for future studies may be to incorporate a brief group-leader check-in to encourage the completion of digital sessions. Few previous studies have reported details of compliance and completion of internet-based sessions in guided interventions for children (Rooksby et al., 2015). Because compliance definitions also have varied across studies, the role of compliance in treatment effectiveness has been difficult to infer. In addition, both age and family support may be related to the number of digital sessions completed (Spence et al., 2019). These are interesting questions that can be examined more closely in future work.

We found no evidence for positive effects of high parental involvement in this study. A parental brochure entails a lower participation burden for all parts involved (parents, children, and health services) compared to active group-based parental involvement. It offers advantages by avoiding the extra time commitment of parental groups. There have been mixed findings in previous literature regarding the effectiveness of parental involvement in child-focused interventions (Crowe & McKay, 2017; Dardas et al., 2018), however, the present study did not examine the effect of leaving a parental component out of the intervention altogether. We also acknowledge that both levels of parental involvement entailed a degree of parental involvement in practical matters related to children's follow-through with the intervention. A start-up meeting providing parents with practical information about their child's participation was recommended as a means to follow up on child home assignments and child access to digital resources. Parental involvement in interventions can take on many formats. One could argue that, according to categorizations like those in meta-analyses by Hudson et al. (2015) and Manassis et al. (2014), both the high and low parental involvement were of low intensity and thus differed too little to reveal an effect. Notably, even these meta-analyses with different categorizations, have offered inconclusive findings about the effects of parental involvement.

In this study, we found no evidence that MFS had a positive effect on children's symptom levels. Effects for child- and parent-reported symptoms of anxiety and depression were non-significant in the primary analyses. We also conducted supplementary analyses that excluded children who were in groups that had to be postponed for a full school semester due to Covid restrictions. In this restricted group, we found an unexpected significant post-intervention effect for parent-reported child depressive symptoms. Parents reported higher symptom levels in children who received MFS compared to those who did not receive MFS. The effect was small: 0.82 SMFQ-P points on a 0 to 26 scale. Furthermore, this result was found only in the baseline-adjusted model,

where a limitation was that a considerable number of parents could not be included because of missing baseline data. It is possible that parents with children with more pervasive symptoms were more likely to complete both surveys. Alternatively, the MFS app may have led to increased parental attention to depressive symptoms displayed within the family context. Another interpretation is that MFS may have had a weak negative effect on children, possibly triggering depressive reactions through its weekly prompts. However, the finding was not supported in the primary analyses involving all parents or in models for child-reported depressive symptoms. Both the null-findings for symptoms in the primary analyses and the negative effect of MFS for parent-reported depressive symptoms in supplementary analyses suggested leaving MFS out of an optimized intervention. Several reasons may account for the seemingly limited usefulness of MFS, such as the unconventional setting for MFS usage. The application of MFS has rarely occurred in a non-clinical sample, primarily provided by first-line services, or with a structured group-based setting. This setting may be relevant for several reasons. First, the feedback loop may transpire differently in a group setting than in a one-to-one setting. It is likely that one-to-one settings allow more room for a provider and a user to reflect together about the MFS responses, thus enhancing a collaborative relation (which was proposed by Ingul et al., 2021 to be a mechanism of change in the study's conceptual model). Second, changing the course of the intervention based on feedback (i.e., tailoring) may be particularly difficult in a group setting and may also require a higher level of expertise than found in many first line services. Finally, feedback practice implementation has proven to be challenging (Bickman et al., 2016; de Jong, 2016). In the present study, both children and group leaders complied to the MFS part of the intervention only about half of the time. While all potential reasons for the suboptimal MFS implementation are not directly known, qualitative interviews with group leaders from the present trial shed light on some key barriers (Haug et al., 2024). Among these barriers, group leaders mentioned that there were too many tasks related to the intervention for both children and for themselves, and that MFS added to the workload. Additionally, group leaders expressed uncertainty, both on behalf of children and themselves, regarding the utility of MFS. The children's age may also play a role in implementation rates (Sale et al., 2020), and the present sample, being relatively young, may not have fully understood the purpose of MFS (Sale et al., 2020). Again, it is also worth considering that the group-based format is not an ideal combination with MFS, especially without adding extra time for group leaders and users to reflect together on the MFS responses. Although MFS did not significantly reduce child-symptomology, there may be dose-response relationships, sub-group differences, or practical usefulness of MFS for the group leaders which we will examine in future analyses.

The study design provided an opportunity to optimize the intervention through testing which components were necessary to produce wanted effects. There were no significant positive effects of the study factors on children's satisfaction with Emotion, nor on symptom levels. An effect size of $d = 0.25$ was used in the power analyses, therefore, the study was not powered to detect effect sizes smaller than this. In cases of non-significant findings, it is important to recognize that absence of evidence cannot be interpreted as evidence of absence. A factor may have had an effect size the study was not powered to detect, but in that case the corresponding component would be too weak to be included in the optimized intervention. From a resource management perspective, if a component is too weak to be included in the screened-in set, resources need not be devoted to detecting its effect (Collins, 2018). In the present context, the finding of no significant positive main or interaction effects on children's outcomes supported including only the low levels in the intervention. The resulting combination (blended format and a parental brochure without MFS) was most likely the simplest, least resource-intensive, and likely the most scalable of all the possible combinations tested. The effectiveness of such an optimized version of Emotion would need to be evaluated in an effectiveness trial. Based on the results from the current study, we would not expect enhanced effectiveness of this optimized intervention compared to the original Emotion, as no candidate components presumed to augment effectiveness is proposed. However, we would expect effectiveness when compared to ordinary practice. In addition, the delivery of the optimized Emotion would most likely be more efficient from a health sector perspective, as the total number of sessions would be reduced from 21 to 8 (62% decrease), thus saving health services considerable time and resources. From a societal perspective, a blended approach may require parents to allocate time to assist their children with starting and maintaining engagement with the digital sessions. From the children's perspective, the number of sessions, whether delivered face-to-face in group or digitally, remains the same (16 sessions). Consequently, the intensity of the intervention and resources required from the child would be similar across both delivery formats. However, it is worth noting that the digital sessions were briefer than group sessions, and because child group sessions typically occur during school hours when the pupils would otherwise be in the classroom, a blended delivery format also minimizes the impact of lost instruction time. Moreover, it provides the child with greater flexibility as they can choose a time suited to their schedule. Finally, more services might find it feasible to offer a less resource-intensive intervention, and the intervention could also be offered to more child groups in the same amount of time. Enhanced intervention uptake may thus be a possible implication of the optimization.

4.1. Strengths and limitations

There were several strengths of the current study. Enabled by the study design, this large, factorial trial examined factors with potential to optimize CBT interventions for symptoms of child anxiety and depression. We used both parents and children as informants of child symptoms. The study had low dropout rates, very high child response rates, and a decent sample size in which both sexes were represented. The study also had sufficient variation in anxious and depressive symptom scores, participants from urban and rural schools across the country, the use of established primary measures, and rigorous data collection methods. The study was also conducted in a real-world, school setting with first-line health services, which is important for the generalizability, interpretation of results, and for translating research into practice.

The study also had several limitations. First, the study design did not allow us to assess the effectiveness of the intervention conditions, as previously mentioned. Until a post-optimization RCT has been conducted, we cannot be certain about the effectiveness of the core intervention shared among all intervention variants. Additionally, it is possible that the study factors had an impact that the study was not powered to detect. The levels of the factors tested also had to be distinct enough to produce discernible effects. For instance, changing the parental involvement factor, like excluding the parental brochure at the low parental involvement level or increasing the number parent group sessions could have led to different outcomes. Conversely, if the constant CBT component was highly effective, it could have made it difficult to detect the effects of manipulating the factors.

Another limitation was the absence of demographic or symptom data for non-participating individuals. This information would have provided important information about the proportion of children with elevated symptoms levels that the trial reached, and about the proportion we did not reach. Additionally, it would have allowed exploration of self-selection bias among participants. The present intervention specifically targeted a particular subgroup of children, namely those at risk of developing anxiety and/or depressive disorders. It was not intended to be applicable to the entire population, and our aim was not to recruit all children who were informed about the study. Without symptom data for non-participants, we cannot determine how many individuals from our target group we reached. The participation rate among invited pupils

was 14%, and although randomized multiple gating was employed, self-selection bias among those who chose to participate cannot be ruled out. Nevertheless, conducting the study in a real-world setting and utilizing recruitment methods that mirror those used outside of the study may have increased the likelihood of participants resembling those the intervention would reach when transferred into routine practice. Among the parental responders, the educational level was higher than the general population, which may also play a role in the generalizability of the findings.

A further study limitation was suboptimal parental response rates. Although parents responded on behalf of 75% of the participating children post-intervention, two parents (when applicable) could report on a child's symptoms, and the overall parental response rate at T2 was 55%. Furthermore, as 11% of parents who responded at T2 were non-responders at T1, they could not be included in the baseline-adjusted analyses.

With respect to user satisfaction findings, a caveat of the high user satisfaction ratings was negatively skewed data with the maximum value as the sample mode, and the possible implication of ceiling effects. Further, user satisfaction was only measured among children, not parents. Parental satisfaction may have differed according to the amount of guidance received and may also have varied according to how much parental time usage and parental child-assistance a factor entailed.

The Covid-19 pandemic as a "force majeure" also affected the trial. Restrictions to reduce contagion were imposed on the country's population, with separate guidelines for primary schools and health services. This had an impact on children's micro-context and worries about the pandemic and the social restrictions may have impacted children's symptomology. It also affected the study's macro-context. Many of the group leaders were school health nurses and called upon to aid in municipal pandemic related tasks (e.g., Covid-19 contact tracing) instead of their regular duties at the school health services. The first planned wave of participants was lost to the study as child face-to-face groups could not be conducted (see "protocol deviations from study as planned"). Recommendations to keep children home from school with symptoms of respiratory illness and quarantines of close Covid-19 contacts also meant that school absenteeism (and thus child intervention group absence) may have been higher during the study.

5. Conclusion

The current study aimed to contribute to the knowledge of several components in an indicated school-based group intervention for children at risk of developing internalizing problems. As we found no significant positive main effects of format, parental involvement, or MFS, the present results were in favor of including only the lower levels of the three factors (i.e., blended format, parental brochure, no MFS) in an optimized intervention. Of the combinations of components tested in the present study, this optimized version of Emotion is most likely the least resource-demanding combination to deliver. However, the optimized intervention's effectiveness remains to be demonstrated.

Because schools may be an especially advantageous arena for reaching children, interventions that are carried out at school hold great promise for prevention and early intervention for mental health problems. While most countries have some form of school health service, they are often delivered with limited reach and scope, or they do not use evidence-based methods (World Health Organization, 2021). The implication of less resource-intensive preventive interventions may be simpler implementation and enhanced uptake, which may, in turn, contribute to closing the gap between the need for health interventions and the provision of such efforts. It is important to note that the present study used an indicated intervention to which school children were exposed, but students do not exist in a vacuum. Broader system-level approaches involving teachers and the whole school community have also been recommended to enhance individual-level effects (Werner--Seidler et al., 2021).

The current study examined only the acute post-intervention effects for the primary child outcomes and one secondary outcome. It is important to know whether these conclusions are supported by one-year follow-up data, as well as by other secondary outcomes. Further, the effectiveness of an optimized version of Emotion would need to be evaluated in a separate effectiveness trial. As optimization is a continual process, suggestions for future research include continued investigation into intervention elements in preventive and indicated interventions for child emotional problems. CBT interventions are complex interventions and the working mechanisms, as well as the contribution of individual therapeutical treatment elements within CBT interventions, remain to be disentangled.

Funding

The present study was funded by the Kavli Trust. The funder had no role in study design, data collection, analysis, writing or interpreting the report.

CRediT authorship contribution statement

Carina Lisøy: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Simon-Peter Neumer:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing – review & editing. **Frode Adolfsen:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Jo Magne Ingul:** Funding acquisition, Investigation, Project administration, Writing – review & editing. **Lene-Mari Potulski Rasmussen:** Investigation, Writing – review & editing. **Tore Wentzel-Larsen:** Formal analysis, Methodology, Writing – review & editing. **Joshua Patras:** Funding acquisition, Writing – review & editing. **Anne Mari Sund:** Investigation, Writing – review & editing. **Kristin Ytreland:** Investigation, Project administration, Writing – review & editing. **Trine Waaktaar:** Supervision, Writing – review & editing. **Solveig Holen:** Funding acquisition, Investigation, Writing – review & editing. **Anne Liv Askeland:** Investigation, Project administration, Writing – review & editing. **Ida Mari Haug:** Investigation, Writing – review & editing. **Elisabeth Valmyr Bania:** Investigation, Writing – review & editing. **Kristin Martinsen:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The following statement has been included within the submitted manuscript:

One of the authors (Kristin Martinsen) receives royalties from the sale of Emotion manuals in Norway. The rest of the authors declare that they have no competing interests.

Data availability

Data will be made available upon reasonable request when the project is completed and data is fully anonymized.

Acknowledgements

We thank all participating children and their parents, as well as collaborating health services and schools, for contributing to the study. We are grateful to the Kavli Trust for funding the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2024.104520.

References

Angold, A., Costello, E., Messer, S., Pickles, A., Winder, F., & Silver, D. (1995). The development of a questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research, 5*, 237–249.

Angold, A., Erkanli, A., Silberg, J., Eaves, L., & Costello, E. J. (2002). Depression scale scores in 8-17-year-olds: Effects of age and gender. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 43*(8), 1052–1063. https://doi.org/10.1111/1469-7610.00232

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3). https://doi.org/10.1016/j.jml.2012.11.001, 10.1016/j.jml.2012.11.001.

Bickman, L., Douglas, S. R., De Andrade, A. R. V., Tomlinson, M., Gleacher, A., Olin, S., & Hoagwood, K. (2016). Implementing a measurement feedback system: A tale of two sites. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 410–425. https://doi.org/10.1007/s10488-015-0647-8

Breinholst, S., Esbjørn, B. H., Reinholdt-Dunne, M. L., & Stallard, P. (2012). CBT for the treatment of child anxiety disorders: A review of why parental involvement has not enhanced outcomes. *Journal of Anxiety Disorders, 26*(3), 416–424. https://doi.org/10.1016/j.janxdis.2011.12.014

Chorpita, B. F., & Weisz, J. R. (2009). Modular approach to therapy for children withAnxiety, depression, trauma, or conduct problems (MATCH-ADTC). *Satellite beach, FL*. PracticeWise LLC.

Clarke, A. M., Kuosmanen, T., & Barry, M. M. (2015). A systematic review of online youth mental health promotion and prevention interventions. *Journal of Youth and Adolescence, 44*(1), 90–113. https://doi.org/10.1007/s10964-014-0165-0

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. NJ: Lawrence: Hillside.

Collins, L. (2018). Optimization of behavioral, biobehavioral and biomedical interventions. In *The Multiphase optimazation strategy (MOST), SW*. Cham: Springer International Publishing.

Collins, K. A., & Dozois, D. J. A. (2008). What are the active ingredients in preventative interventions for depression? *Clinical Psychology: Science and Practice, 15*(4), 313–330. https://doi.org/10.1111/j.1468-2850.2008.00143.x

Crowe, K., & McKay, D. (2017). Efficacy of cognitive-behavioral therapy for childhood anxiety and depression. *Journal of Anxiety Disorders, 49*, 76–87. https://doi.org/10.1016/j.janxdis.2017.04.001

Cuijpers, P., Noma, H., Karyotaki, E., Cipriani, A., & Furukawa, T. A. (2019). Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: A network meta-analysis. *JAMA Psychiatry, 76*(7), 700–707. https://doi.org/10.1001/jamapsychiatry.2019.0268

Dardas, L. A., van de Water, B., & Simmons, L. A. (2018). Parental involvement in adolescent depression interventions: A systematic review of randomized clinical trials. *International Journal of Mental Health Nursing, 27*(2), 555–570. https://doi.org/10.1111/inm.12429

de Jong, K. (2016). Deriving implementation strategies for outcome monitoring feedback from theory, research and practice. *Administration and Policy in Mental Health and Mental Health Services Research, 43*(3), 292–296. https://doi.org/10.1007/s10488-014-0589-6

Garber, J., & Weersing, V. R. (2010). Comorbidity of anxiety and depression in youth: Implications for treatment and prevention. *Clinical Psychology: Science and Practice, 17*(4), 293–306. https://doi.org/10.1111/j.1468-2850.2010.01221.x

Haug, I. M., Saus, M., Rye, M., Rasmussen, L. P., Bania, E. V., Douglas, M., … Patras, J. (2024). *Exploring group leaders experiences implementing a measurement feedback system in municipality child and family services* [Manuscript submitted for publication].

Hemming, K., Eldridge, S., Forbes, G., Weijer, C., & Taljaard. (2017). How to design efficient cluster randomised trials. *BMJ, 358*, j3064. https://doi.org/10.1136/bmj.j3064

Higa-Mcmillan, C. K., Francis, S. E., Rith-Najarian, L., & Chorpita, B. F. (2016). Evidence base update: 50 Years of research on treatment for child and adolescent anxiety. *Journal of Clinical Child and Adolescent Psychology, 45*(2), 91–113. https://doi.org/10.1080/15374416.2015.1046177

Hollis, C., Falconer, C. J., Martin, J. L., Whittington, C., Stockton, S., Glazebrook, C., & Davies, E. B. (2017). Annual Research Review: Digital health interventions for children and young people with mental health problems - a systematic and meta-review. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 58*(4), 474–503. https://doi.org/10.1111/jcpp.12663

Hudson, J. L., Keers, R., Roberts, S., Coleman, J. R., Breen, G., Arendt, K., Bögels, S., Cooper, P., Creswell, C., Hartman, C., Heiervang, E. R., Hötzel, K., In-Albon, T., Lavallee, K., Lyneham, H. J., Marin, C. E., McKinnon, A., Meiser-Stedman, R., Morris, T., Nauta, M., … Eley, T. C. (2015). Clinical predictors of response to cognitive-behavioral therapy in pediatric anxiety disorders: The genes for treatment (GxT) study. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(6), 454–463. https://doi.org/10.1016/j.jaac.2015.03.018

Ingul, J. M., Martinsen, K., Adolfsen, F., Sund, A. M., Ytreland, K., Bania, E. V., Lisøy, C., Rasmussen, L.-M., Haug, I. M., Patras, J., Collins, L. M., Kendall, P. C., & Neumer, S.-P. (2021). Inside the clockwork of the echo factorial trial: A conceptual model with proposed mediators for prevention of emotional problems in children. *Frontiers in Psychology, 12*(2365). https://doi.org/10.3389/fpsyg.2021.703224

Kendall, P. C., O'Neil Rodriguez, K. A., Villab, ø. M. A., Martinsen, K. D., Stark, K., & Banneyer, K. (2014). Cognitive-behavioral therapy with children and adolescents. In J. Ehrenreich, & B. C. Chu (Eds.), *Transdiagnostic treatments for children and adolescents principles and practice* (pp. 161–182). New York, NY: The Guilford Press.

Kern, L., Mathur, S. R., Albrecht, S. F., Poland, S., Rozalski, M., & Skiba, R. J. (2017). The need for school-based mental health services and recommendations for implementation. *School Mental Health, 9*(3), 205–217. https://doi.org/10.1007/s12310-017-9216-5

Kugler, K. C., Dziak, J. J., & Trail, J. (2018). Coding and interpretation of effects in analysis of data from a factorial experiment. In L. Collins, & K. Kugler (Eds.), *Optimization of behavioral, biobehavioral, and biomedical interventions*. Cham: Springer. https://doi.org/10.1007/978-3-319-91776-4_6. Advanced Topics.

Linden, A. (2013). Assessing regression to the mean effects in health care initiatives. *BMC Medical Research Methodollogy, 13*(119). https://doi.org/10.1186/1471-2288-13-119

Lundervold, A., Breivik, K., Posserud, M.-B., Stormark, K. M., & Hysing, M. (2013). Symptoms of depression as reported by Norwegian adolescents on the short Mood and feelings questionnaire. *Frontiers in Psychology, 4*(613). https://doi.org/10.3389/fpsyg.2013.00613

Manassis, K., Lee, T. C., Bennett, K., Zhao, X. Y., Mendlowitz, S., Duda, S., Saini, M., Wilansky, P., Baer, S., Barrett, P., Bodden, D., Cobham, V. E., Dadds, M. R., Flannery-Schroeder, E., Ginsburg, G., Heyne, D., Hudson, J. L., Kendall, P. C., Liber, J., Masia-Warner, C., … Wood, J. J. (2014). Types of parental involvement in CBT with anxious youth: A preliminary meta-analysis. *Journal of Consulting and Clinical Psychology, 82*(6), 1163–1172. https://doi.org/10.1037/a0036969

March, J., Parker, J., Sullivan, K., Stallings, P., & Conners, K. (1997). The multidimensional anxiety scale for children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(4), 554–565. https://doi.org/10.1097/00004583-199704000-00019

Martinsen, K. D., Holen, S., Neumer, S.-P., Waaktaar, T., Rasmussen, L. M. P., Kendall, P. C., & Patras, J. (2017a). The factor structure of MASC youth report in Norwegian school children. *Journal of Child and Family Studies, 26*(7), 1808–1816. https://doi.org/10.1007/s10826-017-0705-9

Martinsen, K. D., Kendall, P. C., Stark, K. D., Rodriguez, K. O., & Arora, P. (2017b). *Mestrende barn, gruppeledermanual barn*. Oslo, Norway: Gyldendal Norsk Forlag.

Martinsen, K. D., Kendall, P. C., Stark, K. D., Rodriguez, K. O., & Arora, P. (2017c). *Mestrende barn, Arbeidsbok foreldre*. Oslo, Norway: Gyldendal Norsk Forlag.

Martinsen, K. D., Rasmussen, L. M. P., Wentzel-Larsen, T., Holen, S., Sund, A. M., Lovaas, M. E. S., Patras, J., Kendall, P. C., Waaktar, T., & Neumer, S. P. (2019). Prevention of anxiety and depression in school children: Effectiveness of the transdiagnostic EMOTION program. *Journal of Consulting and Clinical Psychology, 87*(2), 212–219. https://doi.org/10.1037/ccp0000360

McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: A review of current efforts. *American Psychologist, 65*(2), 73–84. https://doi.org/10.1037/a0018121

Messer, S. C., Angold, A., Costello, E. J., Loeber, R., van Kammen, W., & Stouthammer-Loeber, M. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. *International Journal of Methods in Psychiatric Research, 5*(4), 251–262.

Mulraney, M., Coghill, D., Bishop, C., Mehmed, Y., Sciberras, E., Sawyer, M., Efron, D., & Hiscock, H. (2021). A systematic review of the persistence of childhood mental health problems into adulthood. *Neuroscience & Biobehavioral Reviews, 129*, 182–205. https://doi.org/10.1016/j.neubiorev.2021.07.030

Nahum-Shani, I., & Dziak, J. J. (2018). In L. Collins, & K. C. Kugler (Eds.), *Optimization of behavioral, biobehavioral and biomedical interventions. Advanced topics*. Cham, SW: Springer International Publishing.

Neumer, S.-P., Patras, J., Holen, S., Lisøy, C., Askeland, A. L., Haug, I. M., Jeneson, A., Wentzel-Larsen, T., Adolfsen, F., Rasmussen, L.-M. P., Ingul, J. M., Ytreland, K., Bania, E. V., Sund, A. M., & Martinsen, K. (2021). Study protocol of a factorial trial ECHO: Optimizing a group-based school intervention for children with emotional problems. *BMC Psychology, 9*(1). https://doi.org/10.1186/s40359-021-00581-y

Olason, D. T., Sighvatsson, M. B., & Smami, J. (2004). Psychometric properties of the multidimensional anxiety scale for children (MASC) among Icelandic schoolchildren. *Scandinavian Journal of Psychology, 45*(5), 429–436. https://doi.org/10.1111/j.1467-9450.2004.00424.x

Pinheiro, J. C., & Bates, D. M. (2000). Mixed effects models in S and S PLUS. *Statistics and computing*. New York: Springer.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry, 56*(3), 345–365.

Rapee, R. M., Wignall, A., Sheffield, J., Kowalenko, N., Davis, A., Mcloone, J., & Spence, S. H. (2006). Adolescents' reactions to universal and indicated prevention programs for depression: Perceived Stigma and consumer satisfaction. *Prevention Science, 7*(2), 167–177. https://doi.org/10.1007/s11121-006-0035-4

Rasing, S. P. A. (2021). Blended treatment for depressive disorders in youth: A narrative review. *International Journal of Cognitive Therapy, 14*(1), 47–58. https://doi.org/10.1007/s41811-020-00088-1

Rasmussen, L.-M. P., Patras, J., Neumer, S.-P., Adolfsen, F., Martinsen, K. D., Holen, S., Sund, A. M., & Martinussen, M. (2020). Facilitators and barriers to the implementation of EMOTION: An indicated intervention for young schoolchildren. *Scandinavian Journal of Educational Research, 64*(5), 728–743. https://doi.org/10.1080/00313831.2019.1596976

Raudenbush, S. W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods, 5*(2), 199–213. https://doi.org/10.1037/1082-989X.5.2.199

Rhew, I. C., Simpson, K., Tracy, M., Lymp, J., McCauley, E., Tsuang, D., & Stoep, A. V. (2010). Criterion validity of the Short Mood and Feelings Questionnaire and one- and two-item depression screens in young adolescents. *Child and Adolescent Psychiatry and Mental Health, 4*(1), 8. https://doi.org/10.1186/1753-2000-4-8

Rognstad, K., Wentzel-Larsen, T., Neumer, S.-P., & Kjøbli, J. (2022). A systematic review and meta-analysis of measurement feedback systems in treatment for common mental health disorders. *Administration and Policy in Mental Health and Mental Health Services Research*. https://doi.org/10.1007/s10488-022-01236-9

Rooksby, M., Elouafkaoui, P., Humphris, G., Clarkson, J., & Freeman, R. (2015). Internet-assisted delivery of cognitive behavioural therapy (CBT) for childhood anxiety: Systematic review and meta-analysis. *Journal of Anxiety Disorders, 29*, 83–92. https://doi.org/10.1016/j.janxdis.2014.11.006

Rynn, M. A., Barber, J. P., Khalid-Khan, S., Siqueland, L., Dembiski, M., McCarthy, K. S., & Gallop, R. (2006). The psychometric properties of the MASC in a pediatric psychiatric sample. *Journal of Anxiety Disorders, 20*(2), 139–157. https://doi.org/10.1016/j.janxdis.2005.01.004

Sale, R., Bearman, S. K., Woo, R., & Baker, N. (2020). Introducing a measurement feedback system for youth mental health: Predictors and impact of implementation in a community agency. *Administration and Policy in Mental Health and Mental Health Services Research, 48*, 327–342. https://doi.org/10.1007/s10488-020-01076-5

Schuster, R., Fichtenbauer, I., Sparr, V. M., Berger, T., & Laireiter, A.-R. (2018). Feasibility of a blended group treatment (bGT) for major depression: Uncontrolled interventional study in a university setting. *BMJ Open, 8*(3), Article e018412. https://doi.org/10.1136/bmjopen-2017-018412

Silverman, W. K., & Ollendick, T. H. (2005). Evidence-based assessment of anxiety and its disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 380–411. https://doi.org/10.1207/s15374424jccp3403_2

Spence, S. H., March, S., & Donovan, C. L. (2019). Social support as a predictor of treatment adherence and response in an open-access, self-help, internet-delivered cognitive behavior therapy program for child and adolescent anxiety. *Internet Interventions, 18*, Article 100268. https://doi.org/10.1016/j.invent.2019.100268StatisticsNorway. https://www.ssb.no/en, 10 january 2023.

Thabrew, H., Stasiak, K., Bavin, L. M., Frampton, C., & Merry, S. (2018). Validation of the Mood and feelings questionnaire (MFQ) and short Mood and feelings questionnaire (SMFQ) in New Zealand help-seeking adolescents. *International Journal of Methods in Psychiatric Research, 27*(3), Article e1610. https://doi.org/10.1002/mpr.1610

Thulin, U., Svirsky, L., Serlachius, E., Andersson, G., & Ost, L. G. (2014). The effect of parent involvement in the treatment of anxiety disorders in children: A meta-analysis. *Cognitive Behaviour Therapy, 43*(3), 185–200. https://doi.org/10.1080/16506073.2014.923928

Tollefsen, T. K., Darrow, S. M., Lohne, V., & Berg-Nielsen, T. S. (2020). Experiences with using an idiographic assessment procedure in primary mental health care services for adolescents. *International Journal of Qualitative Studies on Health and Well-Being, 15*(1). https://doi.org/10.1016/j.adolescence.2019.11.006

Villabø, M., Gere, M., Torgersen, S., March, J. S., & Kendall, P. C. (2012). Diagnostic efficiency of the child and parent versions of the multidimensional anxiety scale for children. *Journal of Clinical Child and Adolescent Psychology, 41*(1), 75–85. https://doi.org/10.1080/15374416.2012.632350

Wei, C., & Kendall, P. C. (2014). Parental involvement: Contribution to childhood anxiety and its treatment. *Clinical Child and Family Psychology Review, 17*(4), 319–339. https://doi.org/10.1007/s10567-014-0170-6

Werner-Seidler, A., Spanos, S., Calear, A. L., Perry, Y., Torok, M., O'Dea, B., Christensen, H., & Newby, J. M. (2021). School-based depression and anxiety prevention programs: An updated systematic review and meta-analysis. *Clinical Psychology Review, 89*, Article 102079. https://doi.org/10.1016/j.cpr.2021.102079

World Health Organization. (2021). *WHO Guideline on school health services*. Geneva: World Health Organisation. Retrieved from https://www.who.int/publications/i/item/9789240029392.