

UiT

THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Health Sciences, Regional Centre for Child and Youth Mental Health & Child Welfare

Providing an indicated preventive intervention for children with symptoms of anxiety and depression: A study of effectiveness, implementation factors, and program fidelity

—

Lene-Mari Potulski Rasmussen

A dissertation for the degree of Philosophiae Doctor – 2019



Photo: Leo Eliassen



Providing an indicated preventive intervention for children with symptoms of anxiety and depression: A study of effectiveness, implementation factors, and program fidelity

UIT – The arctic university of Norway

Faculty of Health Sciences

Regional Centre for Child and Youth Mental Health & Child Welfare (RKBU North)

Lene-Mari Potulski Rasmussen

A dissertation for the degree of Philosophiae Doctor

Tromsø, 2019

Table of Contents

Acknowledgements	3
Abstract	5
Abbreviations	7
List of articles.....	8
Introduction	9
Anxiety and depression in children and adolescents.....	9
Prevention of anxiety and depression	11
Context of delivery	12
Effectiveness evaluation	14
Implementation	15
Implementation theories, frameworks and models	16
Treatment fidelity.....	21
Measuring fidelity.....	21
Implementation of CBT programs for emotional problems.....	23
The TIM study	26
Data.....	27
The EMOTION program	28
Implementation of EMOTION	28
Thesis aims	30
Summary of the articles.....	31
Summary of article 1.....	31
Objectives	31
Sample and data collection	31
Measures.....	32
Analyses.....	32
Results	32
Conclusion	33
Summary of article 2.....	34
Objectives	34
Sample and data collection	34
Measures.....	34
Analyses.....	36
Results	36
Conclusion.....	37

Summary of article 3.....	38
Objectives	38
Sample and data collection	38
Measures.....	38
Analyses.....	39
Results	39
Conclusion.....	40
Research ethics.....	41
Discussion	43
Discussion of the main findings in Paper 1.....	43
Discussion of the main findings in Paper 2.....	44
Discussion of the main findings in Paper 3.....	46
General discussion	47
Methodological considerations	49
Future research.....	53
Conclusion.....	56
References	58
Paper 1-3.....	69

Acknowledgements

This work is the results of my doctoral project, which was carried out at the Regional Centre for Child and Youth Mental Health and Child Welfare (RKBU North) at the Faculty of Health Sciences, UIT – The Arctic University of Norway, which together with the Research Council of Norway financed this doctoral project.

First and foremost, I would like to thank all children, families and teachers throughout the country who have participated in this relatively large research project. This could not have been done without you sharing your time and effort to be a part of this study. I especially want to thank all the group leaders and supervisors devoting their time and competence during completion of the groups, and the different services willingly sharing the competence of their employees.

Also, my sincerest gratitude to my supervisors throughout this ph.d-journey; Joshua Patras, Monica Martinussen and Simon-Peter Neumer. From the first day, I have considered myself very lucky to be surrounded by your knowledge and experience in the different fields I found myself diving into. Thank you for *always* answering my questions no matter how trivial they were.

Joshua, your expertise in implementation science have guided me forward in this complicated field. I appreciate all our interesting discussions, and your kindness and constant cheering through closedown of journals and every round of review has meant a lot – thank you!

Monica, you have the ability to make difficult things appear so much easier after a quick talk – and despite your very busy schedule, you are always available to give advice. Thank you for your support and for generously sharing your wisdom.

Simon-Peter, your capacity is tremendous! You have scored numerous videos for me, read everything I sent you, and through your constructive feedback and wise questions, you continuously push me one-step forward. I am possibly even more amazed of your skills as a project leader managing the whole TIM-study with a firm hand, steadiness, and a lot of humour – which have made it so much fun to be a part of. The TIM project group has been an incredible team, and it has been a blast working with all of you!

I also want to thank my scoring team; Ane Sætrum, Reidun Mikkelsen, Kirsten B. Rustad and Helle D. Brochmann for their work in scoring the videos. You have all made a great contribution to this study (and helped me finish in time).

To my fellow ph.d-students for endless discussions in our corner at Gimleveien, and other great colleagues at RKBU North – it would not have been as much fun to go to work without you.

Family and friends - my dearest and nearest! I will forever be grateful for all the babysitting, meals and help during this journey, but mostly – thank you for your genuine support and belief in me.

Lastly, I want to express my deepest gratitude to my lovely boys at home – Henrik Mathias, Niklas and Truls. Special thanks to you Truls, for sharing this experience with me, and always having endless faith in me. You are the best!

Abstract

Anxiety and depression are among the most prevalent difficulties experienced by children and adolescents, and if they are left untreated, the consequences could be detrimental. Preventive efforts to reduce these issues have been shown to be effective; however, implementing psychosocial interventions with good results in real-life settings is a challenging and complex procedure.

The main goals of this thesis were to 1) investigate the effectiveness of the EMOTION: “*Coping Kids*” *Managing Anxiety and Depression* program and 2) examine the factors closely related to the implementation and treatment fidelity of this new transdiagnostic intervention. This national multi-site randomized controlled trial (RCT) included 36 schools from three regions in Norway, where professionals employed in different municipal services (e.g., school mental health services) delivered the group-based EMOTION intervention to children with elevated symptoms of anxiety and depression.

The first study examined the effectiveness of EMOTION, delivered to $n = 266$ children, compared to a control condition ($n = 443$). The results indicated that EMOTION reduced symptoms of self-reported anxiety and depression compared with those reported in the control group. Parent reports showed a significant decrease in depression in the EMOTION group but did not indicate a significant reduction in anxiety symptoms. The results of the evaluation of EMOTION in real-life settings provide evidence of the intervention’s effectiveness in reducing symptoms of anxiety and depression in school-aged children.

The second study used a mixed methods design to investigate facilitators of and barriers to the implementation of the EMOTION program. Healthcare and childcare professionals from different municipal services trained in the intervention completed one survey prior to conducting the new groups ($N = 63$) and a separate survey after completion of the groups ($N = 66$). Additionally, $n = 12$ group leaders were interviewed to further elaborate their experience of implementing the intervention in the organization they were affiliated with. The results indicated that group leaders showed positive attitudes towards the program and felt the need for such an intervention. Issues related to time constraints, lack of support from leaders, and limited participation from the schools were potential barriers to implementation.

The third study investigated the factor structure of the 11-item Competence and Adherence Scale for Cognitive Behavioural Therapy (CAS-CBT; Bjaastad et al., 2016), which was used

to measure adherence and competence during the delivery of manual-based interventions. In this study, six raters assessed a total of 239 individual videos (sessions), evaluating the group leaders' adherence and competence during completion of the EMOTION program. We were not able to replicate the original two-factor structure reported by Bjaastad and colleagues (2016) using a confirmatory factor analysis (CFA) with our data. The model fit was inadequate, particularly when items assessing the main goals of the session were included. Further investigations, including modifications and removal of the main goal items, yielded an alternative factor structure with acceptable model fit.

In conclusion, although the EMOTION program seems to have promising results regarding symptom reduction for anxiousness and sadness, focusing on key implementation factors is important for the continued use of the program. For instance, group leaders' positive attitudes towards EMOTION are insufficient to implement the intervention; organizational factors associated with resources (e.g., time) and support from organizational leaders warrant some attention. If the services want to continue using the school setting for delivering EMOTION interventions, schools should be included more extensively in the planning and execution of the program. Additionally, there should be an increased emphasis on developing valid measurements for use in implementation research, including evaluations assessing adherence and competence regarding quality of delivery.

Abbreviations

APA: American Psychological Association

CAS-CBT: Competence and Adherence Scale for Cognitive Behavioral Therapy

CBT: Cognitive Behavioral Therapy

CC: Control condition

EBP: Evidence-Based Practice

IG: Intervention group

MASC: Multidimensional Anxiety Scale for Children

MFQ-S: Mood and Feelings Questionnaire – short version

ORC: Organizational Readiness for Change

RCT: Randomized Controlled Trial

ROLE: Readiness for Learning and Evaluation

TAU: Treatment as Usual

TIM: Tidlig Intervensjon –Mestrende barn [Early Intervention – Coping Kids]

List of articles

- 1) Martinsen, K. D., **Rasmussen, L-M. P.**, Wentzel-Larsen, T., Holen, S., Sund, A. M., Loevaas, M. E., Patras, J., Kendall, P., Waaktaar, T., & Neumer, S.-P. (2018). Prevention of anxiety and depression in school-aged children: Effectiveness of the transdiagnostic EMOTION program. *Journal of Consulting and Clinical Psychology*, 87. doi:10.1037/ccp0000360
- 2) **Rasmussen, L-M. P.**, Patras, J., Neumer, S-P., Adolfsen, F., Martinsen, K. D., Holen, S., Sund, A. M., & Martinussen, M. (2019). Facilitators and barriers to the implementation of EMOTION: An indicated intervention for young schoolchildren. *Scandinavian Journal of Educational Research*, doi: 10.1080/00313831.2019.1596976.
- 3) **Rasmussen, L-M. P.**, Patras, J., Handegård, B. H., Neumer, S-P., Martinsen, K. D., Adolfsen, F., Sund, A. M., & Martinussen, M. (*In review*). A validation of the adapted version of the Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT).

Introduction

Using a mixed- and multi-method design, this thesis investigates the effectiveness of the indicated preventive intervention EMOTION: “*Coping Kids*” *Managing Anxiety and Depression* (Martinsen, Stark, Rodriguez, & Kendall, 2014) and factors related to the implementation and fidelity of the program. EMOTION is a transdiagnostic group-based intervention for children aged 8 – 12 years with symptoms of anxiety and/or depression and is implemented in primary health care services. Employees from different municipal and mental health services conducted the interventions, but the schools were used as the delivery context. Hence, this thesis provides information and new knowledge regarding the effectiveness, implementation, and fidelity of the program, which is being introduced as a new intervention within this complex service setting.

Anxiety and depression in children and adolescents

Anxious and sad feelings are normal for everyone, including children. It is common, and natural, for children to experience anxiety related to their development (e.g., anxiety when separated from a parent, fear of strangers, or fear of the dark). When symptoms persist and become age inappropriate, however, the child might be at risk of developing a clinical anxiety disorder (Muris, Merckelbach, Mayer, & Prins, 2000). In general, anxious symptoms often include excessive fear, worry, perfectionistic behavior, and/or obsessiveness (Ludwig, Lyon, & Ryan, 2015). Typically, these symptoms are displayed through avoidance of age-appropriate activities, e.g., sleepovers, socializing with peers, and many complaints of somatic pain, such as stomach aches or headaches (Ludwig et al., 2015). As with anxiety, symptoms of sadness related to specific incidents, such as the loss of close friends or family members, chronic disease, bullying, or other worrying occurrences, are normal and expected. When such symptoms persist and are present for a period of time, the child is at risk of developing depression (American Psychiatric Association, 2013). Depressive children often experience loneliness, hopelessness, low self-esteem and a general feeling of fatigue (American Psychiatric Association, 2013), and symptoms of depression seem to predict later onset of a depressive disorder (Ialongo, Edelsohn, & Kellam, 2001; Keenan et al., 2008).

In fact, studies show that anxiety and depression are among the most prominent disorders in children and adolescents (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015). According to Polanczyk and colleagues (2015), the worldwide pooled prevalence among youths (6-18 years) was 6.5% for anxiety and 2.6% for depression. Research also indicates that 2-17% have an anxiety disorder during childhood,

while 1-13% fulfil the criteria for clinical depression during childhood and adolescence (Angold, Erkanli, Silberg, Eaves, & Costello, 2002; Costello et al., 2003). In Norway, according to a report published by the Norwegian Institute of Public Health (NIPH), as many as one out of five children will experience an anxiety disorder at some point during youth and adolescence (Mykletun, Knudsen, & Mathiesen, 2009). Furthermore, a large epidemiological study showed that 3% of children 8-10 years old fulfilled the criteria for an anxiety disorder (Heiervang et al., 2007). The same study by Heiervang and colleagues (2007) also demonstrated that approximately 1% of children in the same age group had a depressive disorder.

Research also shows that many children experience anxious and depressive symptoms but are still below the threshold for a clinical disorder. Balázs and colleagues (2013) indicate that up to 32% of children have anxious symptoms, and 29% of adolescents show subthreshold depression. Having subthreshold symptoms or non-identified problems can have a large impact on children's everyday life (e.g., more reluctant to participate in activities, sleepovers). Several studies have shown that subthreshold anxiety and depressive symptoms predict later onset of a clinical disorder (Polanczyk et al., 2015). The number of children and adolescents with internalizing difficulties along with how these problems occur could be explained by different potential risk factors.

Longitudinal studies generally support the etiological model, which emphasizes child, parent/family and environmental factors as predictors of anxiety and depression (Kroes et al., 2002; Shaw, Keenan, Vondra, Delliquardi, & Giovannelli, 1997; Spence, Najman, Bor, O'Callaghan, & Williams, 2002). Research indicates that the child's temperament is a predictor of later psychopathology. For example, having a biological predisposition (e.g., having a sensitivity towards anxiety or showing behavioral inhibition) has been shown to be an important risk factor in the development of later anxiety (Allan et al., 2015; Milrod et al., 2014). Likewise, gender is considered a predictor of depression, with girls reporting more symptoms over time than boys (Garber, Keiley, & Martin, 2002). Furthermore, both parental/familial (e.g., psychopathology, family environment, parenting behaviors) and environmental factors (e.g., poverty, traumatic experiences) have been shown to affect children's mental health (Stark, Humphrey, Crook, & Lewis, 1990; Tiet et al., 2001; Wood, McLeod, Sigman, Hwang, & Chu, 2003). For instance, certain parenting styles, such as being overprotecting or modelling anxious behavior, have previously been linked to child anxiety (Wood et al., 2003). Several studies have also shown that parental depression increases the

risk of the child developing a depressive disorder (Beardslee, Gladstone, Wright, & Cooper, 2003; Beardslee, Versage, & Gladstone, 1998; Weissman, Warner, Wickramaratne, Moreau, & Olfson, 1997).

The consequences of leaving these disorders untreated may lead to lifelong issues such as social withdrawal, school dropout, drug abuse, and, in the worst-case scenario, suicide (Birmaher et al., 1996; Costello et al., 2003; Kendall, Safford, Flannery-Schroeder, & Webb, 2004; Woodward & Fergusson, 2001). Children and adolescents with internalizing issues, however, often fail to receive sufficient help from local mental health and municipal services due to the quiet nature of their suffering (Chavira, Stein, Bailey, & Stein, 2004; Heiervang et al., 2007). Hence, anxiety and depression may have a large impact on children's daily life and everyday functioning, and the potential impairments indicate a need for interventions that can alter a negative trajectory.

As anxiety and depression share many common factors (i.e., overlapping symptoms, underdeveloped emotion regulation, lack of coping skills), these disorders often co-occur, or one of the disorders may be an antecedent for the other (Avenevoli, Knight, Kessler, & Merikangas, 2008; Chavira et al., 2004). Compared to having just one of the disorders, having both anxiety and depression also decreases the chance of improvement (Birmaher et al., 1996). Hence, targeting both problems simultaneously using one protocol (i.e., a transdiagnostic approach) may promote the overall effectiveness of a treatment, with the added benefit of reducing the need for training in different interventions. Identifying the children with internalizing problems at an early stage to avoid later development of a clinical disorder is therefore of utmost importance (Kovacs & Lopez-Duran, 2010; Luby, 2010)

Prevention of anxiety and depression

Prevention and early intervention have the potential to reduce risk factors and strengthen protective factors, thereby decreasing the likelihood of developing a mental disorder and psychopathology (Mifsud & Rapee, 2005; National Research Council Institute of Medicine, 2009; Webster-Stratton, Reid, & Hammond, 2004). There are different approaches to prevention depending on the risk status of the individual (Mrazek & Haggerty, 1994). Universal prevention targets the whole population, with the idea of providing some benefit to all. Selective prevention targets at-risk populations with increased probability of developing a disorder (e.g., children from low socio-economic backgrounds or from disadvantaged neighbourhoods). Indicated prevention targets individuals identified as having a known

vulnerability or symptoms of a disorder and aims to intervene before treatment for a disorder is required.

Cognitive behavioral therapy (CBT) is considered one of the most effective treatments for internalizing disorders, given its explicit focus on using cognitive and behavioral processes to influence and potentially alter a negative behavioral pattern (Kendall, 2012; Silverman, Pina, & Viswesvaran, 2008). Extensive research also shows that CBT as a preventive effort shows results with small to large effect sizes on internalizing problems (e.g., Mychailyszyn, Brodman, Read, & Kendall, 2012; Stockings et al., 2016; Werner-Seidler, Perry, Calcar, Newby, & Christensen, 2017). There are, however, inconclusive results regarding the effectiveness of universal interventions compared to selective and indicated preventive efforts, where the latter approach seems to yield better results, particularly for depressive children (Calcar & Christensen, 2010; Mychailyszyn et al., 2012; Stice, Shaw, Bohon, Marti, & Rohde, 2009; Teubert & Pinquart, 2011). Furthermore, group-based CBT has the advantage of targeting multiple children at the same time, thereby reducing the number of therapists needed and ultimately diminishing the resources required to handle these difficulties (Flannery-Schroeder, Choudhury, & Kendall, 2005; Wergeland et al., 2014).

Access to adequate CBT efforts and an acceptable therapeutic dose is limited, however, often due to lack of available mental health providers and tools to implement treatments effectively (Farmer, Burns, Phillips, Angold, & Costello, 2003; Ginsburg, Becker, Drazdowski, & Tein, 2012; Weist, Rubin, Moore, Adelsheim, & Wrobel, 2007). Therefore, the focus of investigations in this field has turned from effectiveness studies only to include implementation research. By studying the relevant factors when transferring interventions into new contexts, implementation science serves as the link between research and practice (Fixsen, Naoom, Blase, & Wallace, 2007; Proctor et al., 2009a; Tabak, Khoong, Chambers, & Brownson, 2012). Hence, the context of delivery becomes an essential part of the therapeutic offer because interventions need to be implemented in settings where children can be easily reached and the professionals have the knowledge and skills required to conduct the intervention adequately.

Context of delivery

Researchers and therapists recognize the need to identify children who are struggling and reach them where they are. Therefore, preventive interventions for children and adolescents are becoming an important part of children's primary mental health and school mental health

services (Greenberg, Domitrovich, & Bumbarger, 2001; Skogen, Smith, Aarø, Siqveland, & Øverland, 2018; Weisz, Sandler, Durlak, & Anton, 2005). In Norway, the municipalities manage the first-line mental health services for children and adolescents, which are governed and financed by political mandates, rules and regulations (Stamsø, 2017). Because these first-line services (e.g., school mental health) serve as a link between prevention and treatment, the introduction of preventive efforts in this context has increased during the last decade. It is, however, important to find suitable arenas for conducting interventions to make them available to children.

Schools have been considered an ideal location to deliver and implement preventive interventions because children spend a lot of time there, and children's access to the intervention is therefore easier (Ginsburg, Becker, Newman, & Nichols, 2008; Lee & Gortmaker, 2017). Children with internalizing problems may be easier to identify in school settings, as emotional issues are often displayed within these settings (e.g., speaking or reading aloud in the classroom, socializing with other students) and are therefore more easily recognized by teachers and service providers. Furthermore, these children are less likely to seek help and are generally not inclined to receive mental health services (Chavira et al., 2004; Heiervang et al., 2007), which makes it even more important to identify them at an early stage. Previous studies have also demonstrated that preventive CBT interventions delivered in school settings show positive results (Mychailyszyn et al., 2012; Werner-Seidler et al., 2017).

Conducting interventions in the context of schools also introduces some challenges, and researchers acknowledge that various factors can affect implementation quality (Domitrovich et al., 2008; Payne & Eckert, 2010). For example, Domitrovich and colleagues (2008) proposed a multi-level framework in which the individual level (e.g., attitudes, characteristics), school level (e.g., school culture and climate, resources) and macro-level (e.g., policies and financing) are all important for successful implementation within a school setting. Furthermore, many interventions require collaboration between mental health providers working in municipal services and schools, which increases the complexity of delivering interventions. In general, there is a large gap between existing effective school-based health interventions and the programs, policies, and services offered to children (Lee & Gortmaker, 2017). Therefore, preventive efforts require extensive focus on adequate effectiveness evaluations and implementation to manage the complexities related to real-life settings such as schools.

Effectiveness evaluation

Effectiveness evaluations are important to determine whether an intervention demonstrates good results in real-world settings (Glasgow, Lichtenstein, & Marcus, 2003; Marchand, Stice, Rohde, & Becker, 2011). When moving interventions into real-world settings, researchers have much less control over the surroundings than they do in efficacy trials delivered under strict, highly controlled and optimal conditions (Marchand et al., 2011). The non-optimal conditions in real-world settings generally mean that the service providers might not have enough time to receive full training and supervision, the problems displayed are more complex (e.g., difficult family situations), or the intervention might not be conducted as intensively as needed (Weisz, McCarty, & Valeri, 2006). Hence, it is important that interventions can prove their effect even under such conditions; therefore, results from effectiveness studies have strong external validity.

Researchers generally consider randomized controlled trials (RCTs) the gold standard when evaluating health care interventions (Schulz, Altman, Moher, & CONSORT Group, 2010). Due to the randomization of individuals to control and intervention conditions, the chance of systematic bias between groups, which may occur in other designs (e.g., cohort designs), is greatly diminished. Thus, a positive effect observed in the intervention group can be attributed to the intervention tested.

Another issue associated with effectiveness trials in schools is the possibility of spillover effects between the intervention and control groups. When implementing interventions in school settings, the risk of contamination between individuals in the intervention group and the control group increases within the same school. Because of this contamination effect and due to practical issues, it is common to use the school as the unit of randomization, clustering the children within the same school (Ukoumunne et al., 1999).

Different approaches may be taken to assess preventive interventions. It is most common to evaluate the effect of the intervention by testing the participants at different time points (typically before and after the intervention) (Shadish, Cook, & Campbell, 2002), and the only way to establish whether an intervention is actually working or not is by conducting an effect evaluation. Different standards have been developed to describe the evidence and evaluate quality in research on interventions (Flay et al., 2005; Gottfredson et al., 2015), including measuring user satisfaction, conducting cost-benefit analyses, and/or performing process evaluations. User satisfaction is valuable for determining whether the participants like the

intervention and find it helpful, and cost-benefit analyses assess profitability. Process evaluations concentrate more on the implementation and fidelity of the intervention. By understanding the implementation process and the contextual influences, the ultimate goal is to translate interventions into the practice field without diminishing their effect (Marchand et al., 2011). Implementation is therefore an important element in effectiveness evaluation, as results indicate that effectiveness trials alone without process evaluations and implementation research could falsely promote or discredit an intervention (Durlak & DuPre, 2008; Glasgow et al., 2003).

Implementation

In the early 1980s, the impact of implementation and its relevance for outcome research began to emerge, and studies including implementation research began to evolve.

Implementation research has since become an important part of studies within areas such as education; health science; mental health treatment, prevention and promotion; and program evaluation (Meyers, Durlak, & Wandersman, 2012).

Within behavioral health, implementation is defined as “[a] specified set of activities designed to put into practice an activity or program of known dimensions“ (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005, p. 5). In other words, when implementing an intervention or program within a particular setting (e.g., schools), implementation often refers to the program’s content and what is being delivered. It is well known, however, that implementation is a longitudinal and recursive process (Fixsen, Blase, Naoom, & Wallace, 2009; Metz & Bartley, 2012) that comprises several activities to consider (e.g., making decisions, preparing the organizations, and managing change) (Metz & Bartley, 2012). Other factors, such as transferring successfully to new settings or organizations, ensuring quality of delivery and maintaining implementation over time, are also important parts of the implementation process (Fixsen et al., 2005; Rogers, 2003).

Implementing evidence-based interventions in municipal services is important for improving the mental health services offered to children and adolescents. Transferring and implementing interventions in municipal services is a challenging process, compounded by issues related to both the implementing organization/service and the intervention. This includes factors such as uptake by the services, limited control by the researchers, and restricted resources (Mendel, Meredith, Schoenbaum, Sherbourne, & Wells, 2008), as well as factors linked to the acceptability and feasibility of the intervention (Proctor et al., 2011). By improving the

transfer from research to practice, the effect and quality of the interventions delivered will also increase, and research has shown that focusing on the implementation process clearly produces better results for the youths involved (Durlak & DuPre, 2008). Generally, studies conducted outside highly controlled research settings with effective interventions produce weaker results than more controlled studies, suggesting lower treatment quality (Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005; Gottfredson & Gottfredson, 2002).

Primarily driven by empirical studies, implementation science has been criticised for its lack of a theoretical foundation (Eccles, Grimshaw, Walker, Johnston, & Pitts, 2005; Sales, Smith, Curran, & Kochevar, 2006). Recently, however, there has been an increased focus on providing the field with a theoretical underpinning (Nilsen, 2015). This has led to the development of different theories, models, and frameworks for grappling with the multi-faceted nature of implementation and understanding which factors lead to success or failure.

Implementation theories, frameworks and models

Given the assumption that implementation requires both behavioral (individual) and organizational (collective) change, implementation research has borrowed theories from psychology, sociology and organizational studies (e.g., theory of planned behavior; Ajzen, 1985, 1991) (Eccles et al., 2005; Nilsen, 2015). Within municipal services, there are several factors that influence the implementation process (e.g., policies, stakeholders, agency leaders, staff) (Novins, Green, Legha, & Aarons, 2013). Several theoretical frameworks or models have attempted to explain the different components of the implementation process (Tabak et al., 2012) and the complexity associated with bringing systematic change to mental health care practice. The models have considerable overlap and encompass many of the same key concepts, including the (a) characteristics of the intervention being implemented; (b) organizational characteristics; (c) characteristics of individual practitioners, and (d) implementation process or stages (e.g., Aarons, 2005; Aarons, Hurlburt, & Horwitz, 2011; Fixsen et al., 2005; Glisson & Hemmelgarn, 1998; Glisson & James, 2002; Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004; Klein & Sorra, 1996; Rogers, 2003).

The theoretical framework most referred to within the context of this thesis is Aarons and colleagues' (2011) conceptual model for implementation – *The Exploration, Preparation, Implementation and Sustainment (EPIS) model*. The model is a comprehensive multi-level framework that derives from the public service setting for children and families. It consists of four key phases, framing implementation factors across different levels within each phase

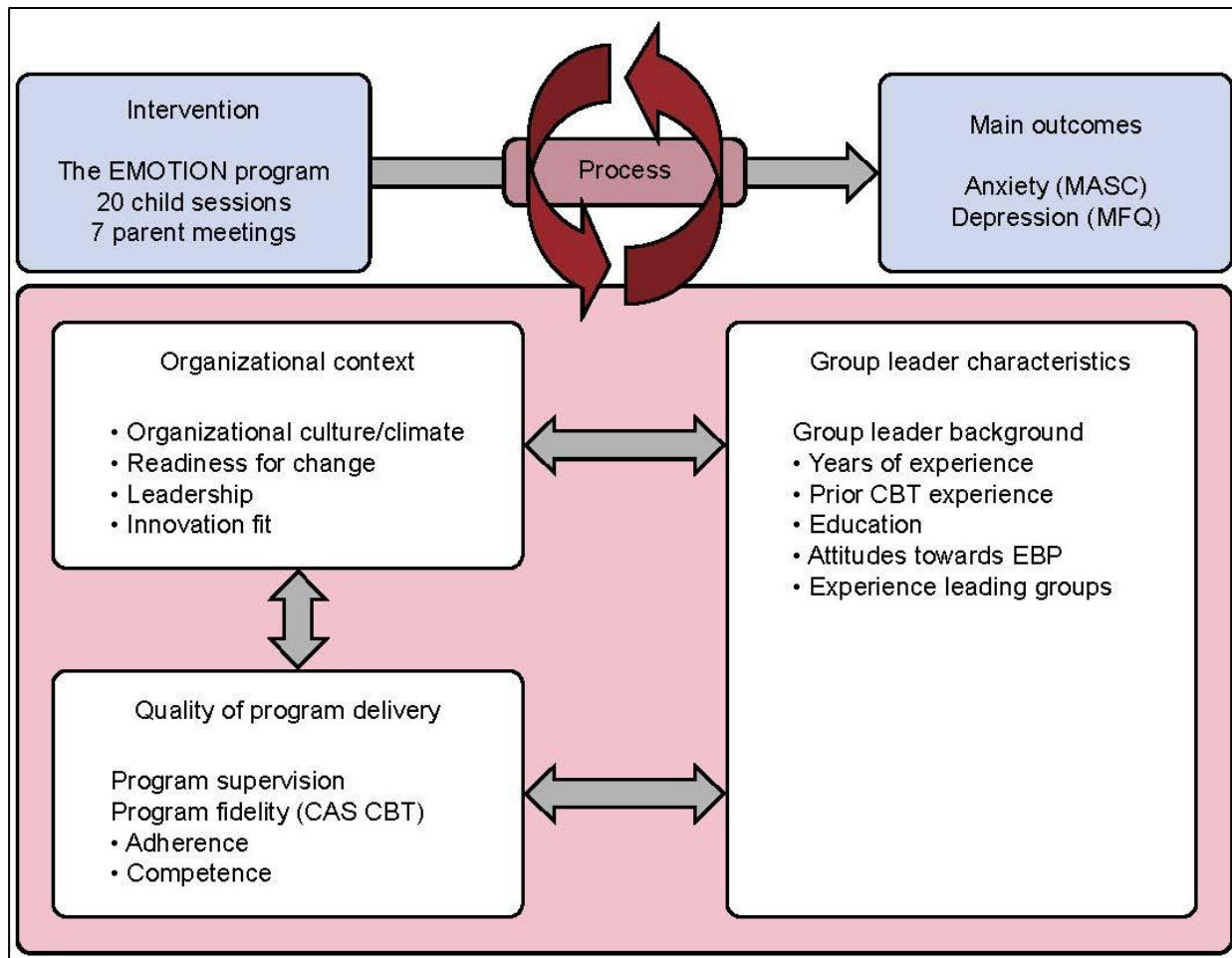
(Aarons, Hurlburt, et al., 2011), and was developed to address important issues particularly relevant within this setting. Therefore, because providers employed in mental health services outside the schools delivered the EMOTION intervention, this model was relevant for this study, as it emphasizes that implementation is shaped by the service context.

The EPIS model (Aarons, Hurlburt, et al., 2011) organizes implementation into the following phases: exploration, adoption decision/preparation, active implementation, and sustainment. Within each phase is a list of factors inside the adopting organization (inner context) and external factors that influence the organization (outer context). These factors are to some extent present in all phases but are more distinct in some and comprise different issues depending on the phase. The exploration phase involves directing attention to issues in the service field or raising awareness regarding challenges in the organization that are not met, whereas the preparation/adoption phase broadly explores the different decisions that must be made before active implementation (Aarons, Hurlburt, et al., 2011). The active implementation phase refers to the specific factors relevant during actual implementation, while the sustainment phase involves factors related to the continued use of the innovation (Aarons, Hurlburt, et al., 2011). The model encompasses a multitude of variables within each phase; however, within the context of this thesis, the *active implementation* phase, especially the inner context, is the most relevant. This is because we are investigating pertinent issues during the ongoing implementation of a new intervention that we could readily collect data from. The other stages that are described in the EPIS model, though instructive and helpful in certain situations, were not included in the present study, as we were focusing on implementation within the context of the effectiveness study.

Based on the active implementation phase in the EPIS model (Aarons, Hurlburt, et al., 2011), we developed an implementation framework for our study reflecting relevant issues that were important to address (see Figure 1). The implementation framework for the Early Intervention – Coping Kids study [Tidlig Intervensjon – Mestrende barn] (the TIM study) highlights some of the processes and different mechanisms involved in the implementation of the EMOTION intervention – a transdiagnostic intervention combining 20 child sessions and seven parental meetings. During this active implementation phase, characteristics within the organization, such as organizational culture and organizational climate, are important issues to consider. Together with readiness for change, innovation fit and adopter characteristics, these are essential factors in the implementation process. Leadership and how the different service

providers experience support from leaders are also considered central to the implementation.

Figure 1. The implementation framework for the Early Intervention – Coping Kids study [Tidlig Intervensjon – Mestrende barn] (the TIM study).



Illustrated by: Joshua Patras.

Organizational culture and climate.

Related but still distinct concepts, organizational culture and organizational climate influence the work environment in organizations (Glisson, Dukes, & Green, 2006). Glisson and James (2002) define organizational culture as the “normative beliefs and shared behavioral expectations” within the organization, whereas organizational climate denotes how the employees perceive the general work environment (Glisson & James, 2002, p. 769-770). Organizational culture is ultimately what makes the organization unique and reflects the history and values of the organization and how individuals communicate with each other (Aarons, Moullin, & Ehrhart, 2017). Organizational climate reflects individuals’ perceptions of how the work environment affects “well-being” at work (e.g., management practices and

procedures). Recently, researchers have further divided organizational climate into molar/generic and strategic climate (Aarons et al., 2017); the latter approach is most relevant for implementation research. Strategic climate includes specific elements inside the organization (e.g., attitudes towards EBPs) and further how staff perceive the management's emphasis on these particular issues.

Organizational readiness for change

Organizational readiness for change generally concerns the ability to change and depends on multiple factors within the organization (e.g., structure, process, equipment and technology, and staff skills and attitudes). Hence, the organization's financial, material, and human resources; context; and supportive processes indicate its innovation-specific capacity (Scaccia et al., 2015). However, one key aspect is staff members' motivation and willingness to change (Aarons et al., 2017). Often, motivation is separated into collective commitment (a shared intention to implement) and collective efficacy (the shared belief that implementation is manageable within the service setting) (Aarons et al., 2017). Hence, settings that are open and support new ideas and interventions are viewed as receptive contexts for implementation.

Leadership

The organizational climate and culture needed for the adoption of new interventions are largely affected by leadership (Aarons, Hurlburt, et al., 2011). According to leadership theories in general, transformational leadership is associated with positive outcomes and is considered more effective than other leadership styles (e.g., transactional or laissez-faire) (Judge, Piccolo, & Ilies, 2004). A transformational leadership style includes a vision where the leader works together with the staff to inspire, encourage, motivate, and serve as a role model to achieve organizational goals (Bass, 1985; Woods & West, 2010). Being more task-oriented, a transactional leadership style uses rewards and punishment to motivate employees and improve their performance, whereas laissez-fair largely represents a lack of leadership (Woods & West, 2010). Leadership is particularly important during the implementation of new interventions. Such processes potentially lead to substantial changes in the organization, and having supportive leaders may serve to buffer a negative organizational climate, reduce frictions, and decrease staff turnover (Aarons, Sommerfeld, & Willging, 2011). Leaders are also in charge of making decisions related to how individuals work and how resources are spent. Thus, having supportive leaders that are capable of ensuring staff members' motivation and creating an environment for change means that the implementation is more likely to succeed (Aarons, 2006; Flodgren et al., 2007). Aarons, Sommerfeld, and Willging (2011)

further demonstrated that leadership is associated with turnover and turnover intention in personnel through its impact on the organizational climate. This indicates the importance of the relationship between organizational factors and leadership, particularly during organizational change; however, more research linking leadership to implementation factors is necessary (Ogden & Fixsen, 2014).

Innovation fit

Innovation fit, also highlighted by Aarons, Hurlburt, and Horwitz (2011) in the EPIS model, is an important aspect of an organization's inner context. Innovation fit is organizations' and individuals' understanding of how the innovation incorporates the organization's values and purpose and service providers' tasks and responsibilities. Proctor et al. (2011) refer to this as the *appropriateness* of an innovation, in which perceived fit enhances implementation efforts. Appropriateness refers to the intervention's relevance and applicability within the service setting. *Acceptability* is the perceived satisfaction with an intervention and reflects whether the content and complexity of the intervention are acceptable given the service setting (Proctor et al., 2011). The difference between appropriateness and acceptability lies in the structure of the intervention: the intervention may be appropriate and compatible with the service setting but unacceptable to conduct (due to resource demands, an extensive manual, etc.). *Feasibility* refers to how well an intervention can be carried out in a given service setting. This often relates to issues such as training, supervision or other requirements, which could have an impact on the completion of the intervention. Other intervention-related issues addressed in the literature are *adaptability*, which refers to the suitability of new interventions and how they fit into the service setting, as well as topics such as cost and treatment fidelity (Proctor et al., 2011).

Group leader characteristics (individual adopter characteristics)

The individual characteristics of the adopters must also be considered in the implementation process in general, but particularly within the active implementation phase. Demographic variables such as education and clinical experience, personal values and goals, and adaptability and attitudes toward interventions are all factors that potentially affect the future utilization of a new intervention. Previous research has shown a relationship between educational level, professional experience, and openness toward adopting EBPs (Aarons, 2004, 2005; Aarons et al., 2010). Furthermore, results vary when evaluating the relation of clinical experience to an implementation outcome such as training (e.g., Carpenter et al., 2012; Garner, Hunter, Godley, & Godley, 2012; Beidas et al., 2014). However, according to

Damschroder et al. (2009), the dynamic relation between individuals and their organization has received limited attention, particularly regarding how that interaction influences behavior change. Nevertheless, the individuals on the front line executing and implementing an intervention clearly indicate that demographic factors and individual characteristics have an impact on adoption and how the intervention is conducted.

Treatment fidelity

Fidelity commonly refers to the program providers' ability to follow the core components in an intervention as designed by the program developers and avoid drift or systematic use of other elements that the program developers have not assigned to the program (Perepletchikova, Treat, & Kazdin, 2007). Treatment outcome is often linked to a high degree of fidelity towards an effective program, and therefore, fidelity is considered highly relevant for implementation quality (Durlak & DuPre, 2008).

In the evaluation of treatment outcome, researchers identify several important aspects of implementation quality and recognize that there are different approaches to measure it. Adherence or the structural dimension of fidelity reflects whether the main elements or key components of the program were delivered and to what degree the program manual was followed (O'Donnell, 2008; Odom, 2008). Process and relational skills (or the procedural dimension) addresses the quality of the relationship between the program providers and those receiving the program (O'Donnell, 2008). However, there is considerable overlap between these dimensions.

Other aspects of fidelity, such as dosage, are also important. Dosage refers to the amount or frequency of the intervention received (Dusenbury, Brannigan, Falco, & Hansen, 2003; Perepletchikova et al., 2007). Dane and Schneider (1998) also focused on participant responsiveness, which reflects how respondents receive the intervention and the degree of engagement displayed. It is argued that all features of fidelity should be measured (Dane & Schneider, 1998); however, other researchers believe that including specific measures will allow for a relevant contribution to the assessment of fidelity (Carroll et al., 2007; Mihalic, 2004). Nevertheless, it is important to have sufficient measures to evaluate fidelity.

Measuring fidelity

Program fidelity is often assessed indirectly by self-reports (e.g., intervention logs, diaries, or checklists) or directly via observations (e.g., in-person observation, "shadowing" or

audio/video recordings), with the latter approach being considered the gold standard (Allen, Shelton, Emmons, & Linnan, 2017). In direct observation, the data are considered to be more accurate, whereas self-reports are more inclined to exhibit reporter bias (Bellg et al., 2004; Lillehoj, Griffin, & Spoth, 2004). Direct observations are more costly and less feasible, however, than self-reports, which are relatively inexpensive and less time consuming (Allen et al., 2017).

Measuring fidelity to ensure that the providers of the intervention follow the program manual and core components of the intervention requires an adequate tool. Monitoring fidelity could also be helpful to determine which elements of the intervention are most beneficial for the children and thereby guide the future development and implementation of the program (Allen et al., 2017). It is therefore necessary to have specific, predefined core elements to measure intervention fidelity accurately. Having psychometrically appropriate measures is also of utmost importance, although there are few assessment tools focusing on fidelity (and implementation in general) that have been evaluated sufficiently (Allen et al., 2017; Martinez, Lewis, & Weiner, 2014).

Investigating the psychometric properties of an instrument is a central element in research to ensure that the tool measures what it is designed to measure and that it can be applied to other contexts (American Educational Research Association (AERA), 2014; EFPA, 2013). The field considers reliability and validity the most important aspects when assessing the psychometric properties of an instrument. Reliability reflects the consistency of the obtained scores and may be estimated in different ways, (e.g., test-retest reliability or internal consistency). When assessing instruments where different observers evaluate a specific behavior (e.g., CAS-CBT; Bjaastad et al., 2016), interrater reliability is the most applicable method to test instrument reliability.

Test validity has many aspects, but the most important one is construct validity. Construct validity commonly refers to whether the intended construct is actually being measured by the instrument (Cook & Beckman, 2006; Streiner & Norman, 2003). One way of examining construct validity is by using confirmatory factor analysis (CFA) (EFPA, 2013; Floyd & Widaman, 1995). According to Martinez, Lewis and Weiner (2014), one should strive to assess structural validity to investigate whether the data represent a unidimensional structure or multiple latent factors according to the theory. Thus, instrument evaluation is important in

all aspects of research, including fidelity assessment and during the implementation of intervention programs in general.

Implementation of CBT programs for emotional problems

In the literature regarding the implementation of CBT interventions for children with emotional issues, several promoters and inhibitors are identified. Previous studies from both the community setting and the school setting support the implementation of CBT interventions for clinical anxiety disorders (Beidas, Mychailyszyn, et al., 2012; Ginsburg et al., 2008; Ringle et al., 2015). Ringle and colleagues (2015) examined CBT in a community setting and identified factors related to the clients (e.g., motivation, complex issues), the intervention itself (e.g., structure), and the organization (e.g., support), which all seemed to influence future use of the intervention. They also found that factors that were facilitators for some of the providers could be regarded as barriers for others. For example, the CBT structure was useful for some, whereas others felt constrained (Ringle et al., 2015). In a similar setting, Beidas et al. (2014) examined related issues regarding the treatment of anxiety. They found an association between variables linked to the inner context (e.g., individual adopter characteristics) and implementation outcomes (e.g., adherence and skills reflecting therapist fidelity).

Within the school context, a recent study by Beidas and colleagues (2012) investigated provider and organizational factors related to training in and implementation of CBT for children with anxiety. They reported that provider attitudes regarding EBPs affected implementation, whereas other provider-level factors or organizational factors did not display any significant associations (Beidas, Mychailyszyn, et al., 2012).

With respect to the implementation of CBT for youth depression, Lewis and Simons (2011) explored these issues in a preliminary report within the community setting. The results indicated that therapist variables, such as attitudes towards empirically supported treatments (ESTs) and readiness for change, correlated positively both before and after training in CBT (Lewis & Simons, 2011). However, this study also found that the therapists' attitudes correlated negatively with perceived client barriers to the implementation of CBT and that factors related to the work setting and clients were negatively associated with implementation, as reported by the therapists (Lewis & Simons, 2011).

Implementation research regarding CBT interventions for children with depression in school settings is limited, however, focusing mostly on the feasibility of the interventions and how best to transport them into school settings (Phillips, Corcoran, & Grossman, 2003; Ruffolo & Fischer, 2009). One study by Langley and colleagues (2010) explored potential barriers and facilitators in the implementation of the Cognitive Behavioral Intervention for Trauma in Schools (CBITS). CBITS is a school-based group intervention targeting youths (ages 11-15 years) with symptoms of PTSD and depression who have been exposed to traumatic events. In their study, program providers were interviewed, reporting several barriers (i.e., competing responsibilities, logistics, parental consent, and administrator/teacher support) and facilitating factors (i.e., professional networks and financial resources) (Langley et al., 2010).

Targeting both anxiety and depression, Lyon and colleagues (2011) examined the feasibility and implementation of modular psychotherapy in a school-based setting. They focused on the therapists from the school-based health centers (SBHC) and found that with full training and a support system, the selection of children, administration of measures to monitor symptom change, and ability to follow the children's use of treatment modules were adequate. However, these results were preliminary and on a small scale. Further, since the results reflect CBT interventions aimed at clinical disorders, different outcomes may be found within a preventive setting.

Kösters and colleagues (2017) investigated program integrity in the context of the implementation of the FRIENDS for Life program (Barrett, 2004a, 2004b), used as an indicated program in a Dutch natural school setting. The results showed lower adherence to the program protocol, which has primarily been used as a preventive intervention for emotional problems, than in previous studies (e.g., Barrett, Sonderegger, & Xenos, 2003; Rodgers & Dunsmuir, 2015). Overall, preventive school-based CBT interventions investigating the effects on both anxiousness and sadness have shown positive results and a reduction in symptoms in the intervention condition (Mychailyszyn et al., 2012; Stockings et al., 2016; Werner-Seidler et al., 2017). Looking closer at the implementation factors, Werner-Seidler and colleagues' (2017) systematic review and meta-analysis of school-based depression and anxiety prevention found that 58% of the investigated studies (of 81 studies total) reported information on fidelity to varying degrees. Furthermore, program dosage was infrequently reported but included information on attendance (e.g., mean number of sessions attended).

Studies focusing on the implementation process during the assessment of preventive interventions have been sparse (Durlak & DuPre, 2008; Durlak & Wells, 1997). The results also indicate that these interventions fail to implement with full fidelity and high quality (Dusenbury et al., 2005; Gottfredson & Gottfredson, 2002; Ringwalt et al., 2003), and little is known regarding continued use and future sustainability. Further, investigations of school health care systems and the factors related to implementation within this context are limited (Forman et al., 2013; Lyon et al., 2011). Despite the challenges encountered in delivering interventions in the school setting, the advantages in terms of accessibility and the ability to reach children at an early stage justify the continual development and implementation of interventions in these settings (Lyon et al., 2011).

Novins and colleagues (2013) conducted a systematic review investigating the existing knowledge regarding the dissemination and implementation of EBPs in child and adolescent mental health service settings, including both prevention and treatment interventions. Organized in accordance with the EPIS model, the results showed that of 73 articles, 23 papers were from the prevention field and addressed issues from the active implementation phase (which is most applicable to this study). Similar to the present study, these papers focused solely on internal contextual factors, including training/fidelity, monitoring and support, and individual characteristics (Novins et al., 2013). The results from this review showed that adherence/fidelity to the intervention increases with ongoing supervision, fidelity monitoring, and support to providers, which ultimately have an impact on the intervention outcome for children and adolescents.

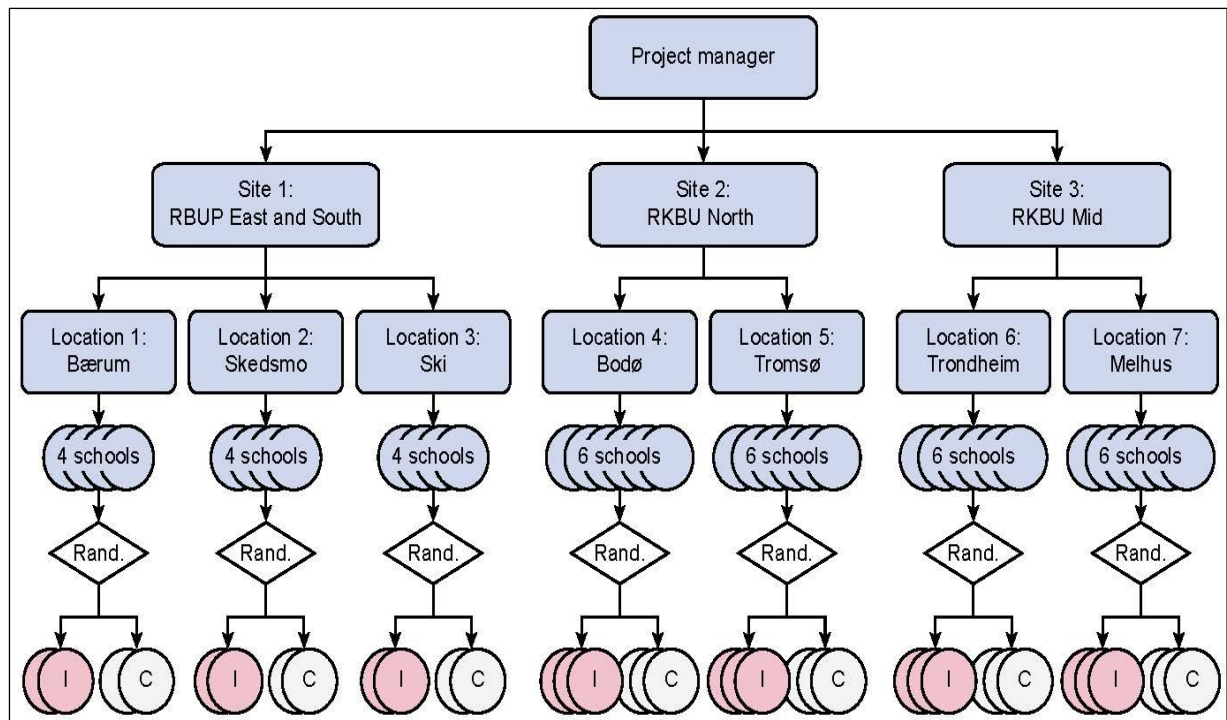
In summary, the research regarding the implementation of CBT-based programs for emotional problems often diverges depending on whether it is community-based or school-based and whether it studies therapy or prevention. Many of the studies exploring implementation factors were preliminary studies, presenting results with limited data. Further, existing research focuses solely on evidence-based interventions with established effective results but does not study how to transfer them into real-world practice in the same extent. In the study presented in this thesis, the implementation research took place during the effectiveness trial, and the main goal was to investigate implementation issues related to further use of a new program. New implementation studies are underway; however, further investigations are needed to strengthen the empirical evidence, particularly regarding preventive interventions delivered in schools by employees from different municipal services.

The TIM study

The data presented and discussed in this thesis are part of a large, multi-site study with a clustered randomized design following the extended CONSORT guidelines (Campbell, Elbourne, & Altman, 2004). For a full description of the protocol of the TIM study [Tidlig Intervensjon – Mestrende barn], see Patras et al. (2016).

The study used a clustered randomized design with restricted randomization, which requires controlling the allocations of the study conditions to some extent to ensure balance across the locations (e.g., blocking, stratification, or minimization) (Higham, Tharmanathan, & Birks, 2015). In this study, there were three participating sites: North, Mid, and South East Norway. At each site, there were at least two geographical locations, including four or more schools, and each school represented a cluster. Hence, because of theoretical and practical considerations (Ukoumunne et al., 1999), the schools were the unit of randomization and were recruited through meetings with the school staff and school leaders. Allocation of the schools to the (a) intervention or (b) control group involved pairing schools based on geographical area, school size and demography and then randomly assigning schools to one condition (See Figure 2). The school remained in the randomly chosen condition during the entire project period, which meant that each semester, the group leaders conducted the intervention with a new group at the intervention schools (i.e., up to two groups per school per year). Beyond supporting the children during screening and being the context of the intervention delivery, the schools had limited responsibility for delivering the intervention. The control schools followed usual care (health nurse, family physician, etc.) and in both the control and intervention schools, we offered teachers and other school personnel a general one-day information workshop on how to detect and help children with symptoms of anxiety and depression. The schools did not receive any compensation for participating.

Figure 2. Illustration of sites, locations and clusters



Data

For the articles presented in this thesis, data collection took place during the active delivery and implementation of the EMOTION program (Martinsen et al., 2014), which lasted from spring 2014 until spring 2016 (five cohorts). The data collection was conducted before the groups were run (T1) and after the groups had ended (T2) for both the children and the group leaders participating. The data were collected electronically using the Conformat software system managed from RBUP East and South.

Paper 1 contains pre-post data for the children and their parents. Paper 2 includes the baseline data for the group leaders (first completion of the surveys). In addition, the qualitative material included in paper 2 was gathered from August 2015 to February 2016. Paper 3 includes video recordings of randomly selected sessions during the group interventions to assess adherence to and competence in using the program.

In 2011, Martinsen and colleagues (2016) tested the feasibility and acceptability of the program manual in a pilot study. The results from this study led to some revisions of the manual, including one less parent session and increased focus on the session structure cohering with the content. These revisions were included in the effectiveness study (paper 1).

The EMOTION program

EMOTION: “*Coping Kids*” *Managing Anxiety and Depression* (Martinsen et al., 2014) is an indicated prevention program targeting children aged 8-12 years with elevated symptoms of anxiety and depression. The program is a newly developed, group-based intervention built upon general CBT principles (Beck, 1976; Martinsen et al., 2014). The trans-diagnostic approach was introduced by combining the well-established programs Coping Cat (Kendall & Hedtke, 2006), aimed at anxiety disorders, and Action (Stark et al., 2007), directed towards depression. Two group leaders delivered the program in one-hour sessions twice a week for 10 weeks during regular school hours or immediately after. The first ten sessions include psychoeducation as well as learning different coping skills and strategies to overcome difficulties related to anxiousness and sadness. The last ten sessions focus more on cognitive restructuring, exposure training/behavioral activation and building a positive self-schema. Additionally, the parents receive seven group sessions, and the children attend four of these. The parent meetings include elements such as positive time with the child, positive reinforcement, instruction on how to deal with punishment and consequences, and ultimately guidance on how to support and help the child handle negative feelings. During the intervention, children (and parents) actively participate through games, role-play, exposure training/behavioral activation, and a variety of tasks, carefully chosen to enhance knowledge of and ability to cope with anxiousness and sadness.

Implementation of EMOTION

Given that members of the research staff developed the intervention and that the research group is investigating it, the implementation of the EMOTION program reflects a top-down implementation strategy (Ogden & Fixsen, 2014). Employees from different municipal services perform the intervention but deliver the program in local schools, and the following implementation steps were undertaken during this study:

Recruitment.

Professionals from different municipal and regional health services were recruited as group leaders, mainly through meetings with leaders of the respective services and/or leaders of the local municipalities. The group leaders were qualified professionals working in different municipal services as health care and childcare providers (e.g., health care nurses, educational and psychological counsellors [EPCs], psychologists). In the Norwegian system, the educational and psychological service (EPS) is an advisory facility within all municipalities and counties. Its main responsibility is supporting preschools, schools and families regarding

education and related issues. A few of the group leaders were employed by the Child and Adolescent Psychiatric Clinic (BUP).

Training and supervision.

Group leader training consisted of a three-day training in the intervention. The first day was a general introduction to CBT principals, and a two-day workshop followed, going through each session. The program developer and the project manager of the study conducted the training, and to avoid bias, all sites received training from both trainers. In addition to lecture-style presentations, experiential learning was emphasized that involved role-play of several of the strategies in the program and discussions.

After most semesters, the research staff held a one-day booster session to discuss some of the challenges encountered during the execution of the intervention. The main topics covered in these booster sessions were reviews of the basic principles of exposure and behavioral activation, how to motivate and collaborate with parents, how to handle restless children in a group setting, flexible use of the manual, and the principles of running groups in general.

The different study locations also had CBT supervisors who were trained in the program to supervise the group leaders running the EMOTION groups. The supervisors met with the group leaders one session prior to startup and then every week during the ten-week program period (two on-site meetings, the remainder via Skype/telephone or face-to-face meetings). Additionally, the supervisors had regular Skype meetings with the trainers to discuss important issues during the intervention period and to ensure similar execution of the intervention across sites.

Quality assurance.

Fidelity. To measure fidelity to the program, the research staff gathered video tapes from 17% of the total number of sessions completed. Using the Competence and Adherence Scale for Cognitive Behavioral Therapy (Bjaastad et al., 2016), the total adherence to the program (rated from 0 = *None* to 6 = *Thorough*) was $M = 3.53$ ($SD = 1.25$), and the mean competence score (rated from 0 = *Poor skills* to 6 = *Excellent skills*) was $M = 3.59$ ($SD = 1.26$). Thus, the results in this study showed a lower mean score than, for example, the scores of the therapists in Bjaastad et al. (2016), who had a mean adherence of $M = 4.57$ ($SD = 0.91$) and mean competence of $M = 4.30$ ($SD = 0.91$).

Dosage. Over five semesters, 17 schools ran 53 EMOTION groups. Unfortunately, due to technical problems in the first semester, attendance (dosage) registration is not complete for all semesters. Therefore, there is a discrepancy between the registered attendance ($n = 241$) and the number of children who completed the intervention ($n = 266$). Furthermore, in some of the groups, parents were only registered as present, with no indication of whether it was the mother, father or both who attended. Nevertheless, the registered mean dosage for children was 89.8% (18 of 20 sessions). Over seven sessions, parents (mother, father or both) showed a mean attendance rate of 5.63 (80%).

Thesis aims

The overall goal of this thesis was to investigate the different aspects involved in providing a high-quality preventive intervention for children with symptoms of anxiety and depression. The different articles give a general idea of the three most important efforts to consider when implementing a new effectiveness intervention. The aim of the thesis was to 1) investigate the effectiveness of EMOTION: “*Coping Kids*” *Managing Anxiety and Depression*; 2) evaluate the facilitators of and barriers to implementation of the EMOTION program, delivered in municipal services; and 3) investigate the instrument assessing group leaders’ adherence and competence during delivery of the intervention. To provide a better overview of the objectives, participants and findings of the different studies, a summary of each paper is hereby presented.

Summary of the articles

Summary of article 1

Martinsen, K. D., **Rasmussen, L-M. P.**, Wentzel-Larsen, T., Holen, S., Sund, A. M., Loevaas, M. E., Patras, J., Waaktaar, T., Neumer, S-P., & Kendall, P. (2018). Prevention of anxiety and depression in school-aged children: Effectiveness of the transdiagnostic EMOTION program. *Journal of Consulting and Clinical Psychology*, 37, 212-219. doi: 10.1037/ccp0000360

Objectives

The objective of the first paper was to investigate the effectiveness of the EMOTION program using a randomized controlled trial (RCT) measuring symptom levels before and after the intervention period (pre-post control group design). EMOTION is an indicated preventive program for children from third to 6th grade (9-12 years) with symptoms of anxiety and depression.

Sample and data collection

Of the total number of children ($N = 7322$ from 36 participating schools) receiving information about the study, $N = 1692$ (23.1%) completed the initial screening. Of these, 837 children scored above the predefined cut-off and were invited to participate in the study (43.1% boys); $N = 430$ were enrolled in the intervention group (IG) and $N = 443$ were part of the control condition (CC). After withdrawal, exclusion criteria (e.g., mental retardation, pervasive developmental disorder) and exclusion due to lack of resources (e.g., not enough group leaders to conduct more than one group) were accounted for, the IG consisted of 266 children, and the control group included $N = 428$. Age was calculated using grade levels, for a mean age of 9.64 years ($SD = 0.93$), and over 95% of the children were Norwegian, Nordic or of Western European origin. In the IG, $N = 268$ parents completed the pre-assessment, and $N = 193$ completed the post-assessment. The parents in the CC completed $N = 301$ pre-assessments and $N = 228$ post-assessments.

Recruitment of children to the effectiveness study followed a stepwise procedure: first, research staff attended school or parent meetings and handed out information regarding the study to the students and parents. Second, with knowledge that this was an effectiveness study for children with symptoms of anxiety and depression, parents were required to give their consent if the children wanted to participate. Third, the children then completed an online

questionnaire at school that screened for symptoms of anxiety and depression. Finally, all children who scored one SD above a predefined cut-off (based on the population mean) on anxiety, depression or both received an invitation to participate in the study. We did not have an upper limit for inclusion. We applied gender-specific cut-off scores for anxiety due to the discrepancy in the mean scores.

Measures

MASC

The Multidimensional Anxiety Scale for Children (MASC-C; March, 1997) is a 39-item multidimensional self-report instrument assessing anxiety in children 8-19 years old over the two last weeks. The instrument consists of four scales: (1) Physical symptoms, (2) Harm avoidance, (3) Social anxiety and (4) Separation anxiety/panic; three of these have additional subscales (March, 1997). The responses are rated from 0 (*never true about me*) to 3 (*often true about me*). A similar version exists for parents (MASC-P). In this study, the internal consistency was $\alpha = 0.91$ for the MASC-C and $\alpha = 0.90$ for the MASC-P.

MFQ-S

The Mood and Feelings Questionnaire-short version (SMFQ; Angold et al., 1995) is a 13-item screening tool for identifying symptoms of depression in children aged 8-18 years within the last two weeks. In addition, we added one item about suicidality. The SMFQ-S was rated on a scale of 0 = *Not true*, 1 = *Sometimes* and 2 = *True*. Similarly, a parent version (SMFQ-P) exists for this instrument as well. In this study, $\alpha = 0.94$ for the child version and $\alpha = 0.88$ for the parent version.

Analyses

Mixed model analyses were used in this study, where the fixed effect was a time by randomization group interaction, including analyses adjusting for gender and age group (3rd and 4th grade = younger; 5th and 6th = older). The R package nlme (The R Foundation for Statistical Computing, Vienna, Austria) was used to estimate the models and included intent-to-treat (ITT) analysis.

Results

The results showed a significant Time \times Condition interaction for anxious symptoms ($p < .001$) and depressive symptoms ($p = .040$) indicating that the intervention group had a significantly larger symptom reduction compared to the control group, as reported by

children. At post-test, there was a significant difference between IG and CC, where the CC scored higher (5.35 points). In the IG, the anxious symptoms decreased 11.83 points overall, $p < .001$, which demonstrated a reduction between 17.4% and 19.7%, depending on age (older and younger children) and gender. For comparison, the CC had an overall reduction of 4.63 points, $p < .001$, which yielded a decrease between 7.0% and 8.0%. The subgroup analysis of anxious symptoms showed a significant symptom reduction in all groups, where boys and older children decreased the most. Further, for depressive symptoms, there was a reduction of 2.31 points ($p < .001$) in the IG, ranging between 21.0% and 25.0% for the different groups examined, compared to the CC with a 1.50-point ($p < .001$) decrease, which indicated a reduction between 14.6% and 17.6%. The subgroup analysis showed a significant symptom reduction in older children (3.30 points).

Similar results were obtained for parent-reported symptoms regarding depression but not anxiousness. The pre-intervention results were significantly higher in the IG for both anxiety and depression, but at post-intervention for the depressive symptoms, the results were not. This indicates a significant reduction in parent-reported depressive symptoms ($p < .001$) compared to the CC ($p < .133$). Subgroup analyses for gender and age showed a decrease in anxious symptoms in both, whereas for depression, only older children reported a significant reduction. Overall, parents reported about 1 *SD* lower on both anxious and depressive symptoms compared to the child reports.

Conclusion

The transdiagnostic indicated prevention program resulted in a decrease in both anxious and depressive symptoms as reported by children. Parents also reported symptom reduction for depression. Hence, the EMOTION program has the potential to reduce internalizing problems in youths and thereby prevent the onset of emotional disorders.

Summary of article 2

Rasmussen, L-M., Patras, J., Neumer, S-P., Adolfsen, F., Martinsen, K. D., Holen, S., Sund, A. M., & Martinussen, M. (2019). Facilitators and barriers to the implementation of EMOTION: An indicated intervention for young schoolchildren. *Scandinavian Journal of Educational Research*, 1-16. doi: 10.1080/00313831.2019.1596976

Objectives

The objectives of the second article were to identify characteristics that might facilitate or inhibit the implementation process of the EMOTION program. By using a mixed methods design, we investigated organizational and individual factors promoting or inhibiting implementation. We also examined group leaders' satisfaction and intention to continue with the program. In addition, we explored the group leaders' experiences with implementing the intervention in the municipal services by conducting qualitative interviews.

Sample and data collection

Of 68 group leaders trained in the intervention, 63 completed the group leader questionnaire prior to running the groups (93% participation rate), and 97% ($N = 66$) completed the post-intervention questionnaire after the groups were finished. Almost 95% of the sample was female, and the mean age of the total sample was 39.6 years ($SD = 9.7$). The group leaders were recruited from seven municipalities within the three participating regions (North, Mid, and South East) in Norway.

Additionally, eleven qualitative interviews of $n = 12$ group leaders (two informants were present at the same time in one of the interviews) were conducted. The interviews took place at the practitioners' workplace or in other suitable settings (e.g., a nearby café) and lasted approximately 1-1.5 hours. All interviews were audio recorded and verbatim transcription of the interviews then followed. The participants were four health care nurses, four psychologists and four educators. All the informants were women, and all except one worked in the local municipal health services (e.g., school health service).

Measures

Group leader questionnaire (T1)

Demographics. Fourteen questions regarding gender, age, work place and municipality, profession, percentage of full-time employment (e.g., 50%, 100%), clinical or other

specialties and experience related to work (e.g., years in the field, anxiety and/or depression, treatment methods) were included.

Work environment and innovation fit. To address issues relevant within this context, ten questions concerning work environment and intervention fit were developed for this study. An exploratory factor analysis yielded three subscales: “Innovation fit” ($\alpha = .95$), “Organizational support” ($\alpha = .69$) and “Attitudes towards evidence-based programs” ($\alpha = .83$). The item “I have such a large workload that it will be difficult to find time to run the EMOTION program” did not fit any of the subscales and was therefore treated separately. All items were rated on a scale from 1 (Strongly disagree) to 7 (Strongly agree).

Organizational Readiness for Change (ORC). Inspired by the Employee Problem Scale from the Organizational Readiness for Change questionnaire (ORC; Lehman, Greener, & Simpson, 2002), a subset of 32 questions was developed for this study. The items were rated from 1 (Strongly disagree) to 7 (Strongly agree) and comprised six subscales: Adaptability ($\alpha = .62$), Program goals ($\alpha = .59$), Cohesion ($\alpha = .77$), Efficacy ($\alpha = .80$), Autonomy ($\alpha = .66$) and Communication ($\alpha = .80$). Six questions were included from the following subscales because of their relevance to this study: Adaptability (one item), Efficacy (four items) and Autonomy (one item). Reliability analyses showed that including these items maintained or increased the alpha coefficients. In addition, four items were removed because they did not fit within this context (e.g., “You have the skills needed to conduct individual counseling”). Of course, the wording of the items was changed from second person (you) to first person (I) to resemble the other questions in this study.

Readiness for Organizational Learning and Evaluation Instrument (ROLE). From the Readiness for Organizational Learning and Evaluation Instrument (ROLE; Preskill & Torres, 1999), we included items reflecting work culture and leadership. The subscale Work Culture ($\alpha = .91$) comprised 17 items, whereas the Leadership scale consisted of nine items ($\alpha = .89$). These items were rated on a five-point Likert scale ranging from 1 (Strongly disagree) to 5 (Strongly agree).

Post-intervention questionnaire (T2)

After finishing the intervention, the group leaders completed the post-intervention questionnaire (T2). This survey consisted of 14 questions asking the group leaders to report on the number of sessions led and their satisfaction with being a group leader and intention to

continue with the EMOTION program. The responses were reported on a scale from 1 (Very dissatisfied) to 5 (Very satisfied). The last question regarding the intention to continue with the program was rated from 1 (Very unlikely) to 5 (Very likely).

Qualitative interview

For this study, we developed a semi-structured interview guide focusing on the group leaders' organizational setting as well as the strengths and weaknesses within the organization and with the program regarding implementation. Building upon Aarons and colleagues' (2011) conceptual model of implementation, the main goal of the interview was to explore the group leaders' experience working with EMOTION within their organizational context.

Analyses

To conduct the quantitative descriptive analysis, we used Pearson's r (two-tailed), and to assess the Cronbach's alpha for testing internal consistency in the different subscales, we used the statistical package IBM SPSS (24).

The qualitative analyses were conducted using Lacey and Luff's (2001) analytical framework, which is theoretically driven and consists of five main stages: familiarization (initial reading and familiarization with the text), identify a thematic framework (initial coding of the data), indexing (searching for themes), charting (organize the data), and finally interpretation (searching for patterns relevant to this study). To validate the qualitative data, one of the co-authors with experience regarding qualitative analyses read and discussed the data with me until we agreed on the main results.

Results

This study showed that several factors are important in the active phase of implementation and affect the group leaders' inclination to continue with the EMOTION program. Although the program received positive responses and was considered meaningful in the service context (e.g., 90% of the group leaders slightly agreed, agreed or strongly agreed that the intervention was needed in the service setting), some issues warrant more attention. Particularly, organizational factors such as time constraints, a heavy workload and lack of support from leaders seem to influence future use of the program. For instance, the leadership scale evaluating the management and leadership support in the organization was endorsed (group leaders agreed or strongly agreed with the statements) by less than 30% of respondents. We

also found a negative correlation between heavy workload and satisfaction with being a group leader ($r = -.36, p < .01$) and intention to continue ($r = -.29, p < .05$).

The qualitative material supported these results, as the three main findings from the interviews – *organizational factors*, *interventional aspects* and *school investment* – included both facilitators of and barriers to the implementation of the intervention. The group leaders found their workplace highly suitable for delivering the program. However, the main barriers were closely linked to time, resources and general support from leaders. Interventional aspects reflected an overall need for the intervention but indicated that it could be further adjusted to better fit the services. In particular, our last finding importantly highlighted the fact that increased collaboration between the schools involved and the services during delivery of the intervention would improve the offer to the children.

Conclusion

The group leaders highlighted several important aspects of the facilitators and barriers in the implementation of a new intervention targeting anxiety and depression in municipal services. Although the program provides the services with extra tools to handle these issues, there are still some barriers (e.g., lack of time and resources), which could affect future utilization of the EMOTION program. Supportive leaders and a cooperative school are also important for continued use of the intervention.

Summary of article 3

Rasmussen, L-M., Patras, J., Handegård, B-H., Neumer, S-P., Martinsen, K. D., Adolfsen, F., Sund, A. M., & Martinussen, M. (*In review*). A validation of an adapted version of the Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT).

Objectives

The objective of the third article was to investigate the factor structure and reliability of an adapted version of the Cognitive and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT; Bjaastad et al., 2016) by conducting a confirmatory factor analysis (CFA) in our sample and testing the instrument in a group format, which has not previously been done.

Sample and data collection

The group leaders received video cameras and a list of sessions (a block of 4 child and 2 parent sessions) prior to starting new groups. Randomly selected videos were collected in blocks to ease the data collection, meaning that if the group leaders in one group started with session 4, then sessions 5, 6, and 7 followed. After completion of the groups, the group leaders handed in the video cameras, and the video files were stored on a secure server. A total of $N = 239$ sessions (17% of all sessions) were recorded and scored for $N = 52$ groups led by the group leaders. During the project period, a total of $N = 266$ children participated in the intervention (mean age of 9.64 years, $SD = 0.93$).

Measures

CAS-CBT

The Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT; Bjaastad et al., 2016) is an instrument used to evaluate adherence and competence in cognitive behavioral therapy (CBT) with children and adolescents and was originally developed for anxiety disorders. The 11-item instrument comprises three main sections covering the key domains in CBT for children with anxiety (Bjaastad et al., 2016): cognitive therapy structure (e.g., homework, session structure and progress), process and relational skills (e.g., reinforcement, collaboration, flexibility) and goals (specific goals for the session from the treatment protocol). Then, a competence score is assessed globally for each of the three main sections (e.g., competence score for cognitive therapy structure). In addition, the observers make an overall evaluation of the group leaders' adherence and competence in the session. The adherence score ranges from 0 (*None*) to 6 (*Thorough*). The competence score

also ranges from 0 (*Poor skills*) to 6 (*Excellent skills*), with an explanation attached to the ratings indicating the different qualities that must be exhibited. There are also three general questions about the video quality and challenges with the session.

In this study, we made a few adaptations to fit the EMOTION program under consultation with the CAS-CBT developer. On the CAS-CBT, the parents are included with one item called “parental involvement” (Bjaastad et al., 2016). In EMOTION, the parents participated in seven individual sessions, and therefore, the item was removed. The parent sessions were rated separately with the same instrument. Additionally, in the original version, there are two program goals to be rated, but in our version, we had up to three goals, so one item assessing the third goal was added. The instrument developer approved the modifications. The CAS-CBT has previously shown good internal consistency ($\alpha = .87$), good to excellent interrater reliability and high rater stability (Bjaastad et al., 2016).

Analyses

To examine the interrater reliability between the expert scorer and the student scorer during the fidelity checks, as well as the internal consistency and Pearson’s r , SPSS statistical packages (24.0) were used. In addition, using Mplus 7.0 statistical software with the weighted least squares estimator (WLSMV; Muthèn & Muthèn, 1998-2010) with ordered categorical (ordinal) indicators, we performed a confirmatory factor analysis (CFA). We further investigated the structure of the CAS-CBT by testing different models. Several fit indices were used to assess how well the model fit the data, including chi square, the root mean square error of approximation (RMSEA; Steiger and Lind 1980), Bentler’s comparative fit index (CFI; Bentler, 1990) and the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973). A significant chi square result ($p < .05$) indicates misfit (Kline, 2011), whereas an RMSEA $< .08$ and a CFI and TLI $> .90$ indicate adequate model fit. Preferably, an RMSEA $< .05$ and CFI and TLI $> .95$ indicate a good model fit. In addition, a p-value is given for the RMSEA and is interpreted as the probability that the RMSEA is $< .05$.

Results

The findings from this study indicated fair to good interrater reliability, ranging from $\alpha = .40$ to $.74$. When conducting the CFA, we first tested a unidimensional model, which showed poor model fit ($\chi^2 = 497.076$, $p < .05$, $df = 44$, RMSEA = 0.208, $p < .05$, CFI = .953, and TLI = .941). Then, we investigated a second model based on the structure and scoring of the CAS-

CBT. The results implied a model misspecification, indicating that the model was not trustworthy. We then tried to replicate Bjaastad et al. (2016), but the results did not show adequate fit on all indices ($\chi^2 = 183.69, p < .05, df = 43$; RMSEA = 0.117 $p < .05$; CFI = .985; and TLI = .981). A modified version of the previous model (correlating the residuals of two items based on theory) yielded the following results: $\chi^2 = 162.10, p < .05, df = 42$; RMSEA = 0.109, $p < .05$; CFI = .987; and TLI = .984. In a further modification of this model, we correlated the residuals between two items with a strong association, but model fit did not improve ($\chi^2 = 163.37, p < .05, df = 41$; RMSEA = 0.112, $p < .05$; CFI = .987; and TLI = .983). Finally, we tested a model in which all the items evaluating the session goals were removed. This improved model fit to some extent ($\chi^2 = 23.26, p < .05, df = 11$; RMSEA = 0.068, $p = .19$; CFI = .998; and TLI = .997). Internal consistency estimates for the two subscales derived from the last model showed good reliability for the subscale “CBT structure” ($\alpha = .85$) and an excellent alpha value for the subscale “Process and relational skills” ($\alpha = .93$).

Conclusion

This study showed that when conducting a CFA in our sample, we were unable to estimate a good model fit, especially when the items evaluating the session goals were included. When we removed these items, however, we were still not able to achieve an adequate model fit, although it improved. This implies that the CAS-CBT might benefit from further development to effectively evaluate CBT group interventions for children with anxiousness and sadness.

Research ethics

The Regional Committee for Health and Medical Research Ethics (REK) approved the study (2013/1909/REK Sør-Øst). Hence, all procedures performed in this study were in accordance with the ethical standards of the regional research committee and with the 1964 Helsinki declaration (World Medical Association, 2013) and its later amendments or comparable ethical standards. We obtained informed consent from all individual participants included in the study, and parents consented on behalf of their children. The raters who scored the video recordings of the sessions signed a declaration of confidentiality.

However, when conducting research on humans, and children in particular, there will always be some ethical considerations requiring extra attention. According to the Health Research Act [Helseforskningsloven] and national ethical guidelines [De nasjonale forskningsetiske komiteene] (NEM National Committee for Medical and Health Research Ethics; NESH The National Committee for Research Ethics in the Social Sciences and Humanities, 2016), all participants need to understand all aspects of a research project, including the purpose and consequences of participation. In this project, we include children aged 8-12 years. They do not have the competence to give informed consent and are thus, by definition, labelled a “vulnerable group” (NESH The National Committee for Research Ethics in the Social Sciences and Humanities, 2016). Hence, we as researchers have a responsibility to provide appropriate information to ensure that the participants understand what they are participating in.

In this study, this was prudently taken into consideration when the information letters to all participants were produced and later approved by the REK. During project presentations at the schools, the children received age-appropriate information and had the opportunity to ask questions. If possible, the research group held parent meetings to inform parents about the project and respond to any questions they might have. Furthermore, in the information letter to the parents, we emphasized that participation was voluntary and that they could withdraw from the study at any time.

As this study targeted symptomatic children, the possibility that some children would experience stigma was present. First, one could be exposed to stigma by handing in the informed consent. However, the research group instructed all the children to hand in the consent form and (after discussion with the parents about participation) to check the box that

was most suitable for that particular child (we want/do not want to participate in the study). Second, the children could experience potential stigma by participating in the groups, as they were delivered during school hours or immediately after. We wanted to reduce this issue as much as possible and therefore carefully selected how we presented and communicated the project to the participants. Further, the issue of stigma was also investigated in the pilot study (Martinsen et al., 2016), which indicated that the potential benefits of participating in an indicative intervention outweighed the potential negative effects.

In this project, we also included two different qualitative methods – observation (video recordings of the group sessions) and interviews. In both of these methods, it is important to follow national ethical guidelines and strive to ensure the integrity of the participants (NESH The National Committee for Research Ethics in the Social Sciences and Humanities, 2016). We informed the group leaders thoroughly about the study and the option to withdraw their participation at any time. During the interviews and in the interpretation of the material afterwards, the intention was solely to ensure an accurate presentation of the participants. Regarding the video observations, the group leaders received instructions to arrange the recordings of the sessions, thereby leaving the choice to participate up to them. Furthermore, video observations require researchers to be especially careful because observing the participants and interpreting their behavior might be considered degrading (NESH The National Committee for Research Ethics in the Social Sciences and Humanities, 2016). This care was ensured during the training of the raters and when they signed the declaration of confidentiality.

Discussion

The major goal of this thesis was to investigate both the effectiveness and implementation of the EMOTION program. The EMOTION program seems to show a potential benefit in reducing anxiousness and sadness. For further use of the program, however, implementation efforts must be considered. The results from the studies in this thesis indicate that different factors affect implementation, including factors closely linked to the organization where the employees delivering the intervention worked and other collaborators (e.g., schools). Additionally, aspects of the intervention and issues regarding the measurements all seem to have an impact on the continued use of the intervention.

Discussion of the main findings in Paper 1

The effectiveness evaluation of the intervention revealed a decrease in both anxious and depressive symptoms, as reported by the children. This clearly shows that the indicated prevention program EMOTION has the potential to reduce internalizing problems in young school children. Although both the intervention group and the control condition experienced symptom reduction, this reduction was significantly greater in the intervention group. These results are well aligned with those of other studies investigating CBT interventions delivered in school settings, although interventions for depression have shown less distinct results (Calear & Christensen, 2010; Mychailyszyn et al., 2012; Werner-Seidler et al., 2017). Furthermore, anxiousness often precedes depressive symptoms (e.g., social phobia leads to loneliness and sad feelings; Kovacs & Lopez-Duran, 2010), which highlights the importance of targeting both issues. One reason for the lower symptom reduction in sad children could stem from the more abstract features of depression, such as negative thoughts and a general feeling of sadness.

Subgroup analyses for gender and age (older = 5th and 6th grade, younger = 3rd and 4th grade) supported this notion to some extent. The results showed a significant decrease in anxious symptoms in the intervention condition for both subgroups. Previous studies have presented similar results, demonstrating symptom reduction in different subgroups (Gillham, Hamilton, Freres, Patton, & Gallop, 2006). Among the more depressed youths, however, only the older children showed a significant reduction in symptoms. Depression typically has a later onset than anxiety (Stice et al., 2009), which could explain why older children gained more from the intervention than younger children. Furthermore, low self-esteem, solitude and general fatigue generally reflect depressive symptoms. This may imply that compared to treating

more specific anxious symptoms (such as being afraid of the dark or speaking aloud in the classroom), when treating depressive symptoms, the children need to reach a certain age to understand the therapeutic mechanisms (e.g., cognitive restructuring) and abstract reasoning. Thus, given the nature of anxiety and depression, it is easier to identify specific goals for anxious children than for depressed children and thereby develop a plan to reach these goals.

Parents reported similar results as the children; however, the decrease in anxious symptoms was not significant. The results related to parent reports of child anxiousness have been diverse, which could indicate that parents are unaware of the children's difficulties. In general, the parents reported lower symptoms than the children. This reflects a disagreement between parents and children, which has been shown in previous studies as well (De Los Reyes et al., 2015; Wei, Cummings, Villabø, & Kendall, 2014). This illustrates the importance of including self-reports, because parents might not be aware of the child's situation, particularly with relation to internalizing problems (Wei, Hoff, et al., 2014). Additionally, as previously mentioned, anxious or depressive symptoms may be more relevant in other settings (e.g., schools) than in the home context, causing parents to underreport a child's difficulties (De Los Reyes et al., 2015). Thus, it is important to gather information from children and other informants, including parents and teachers.

Discussion of the main findings in Paper 2

In the second paper, we sought to identify the facilitating and hindering factors closely linked to the implementation and future utilization of the EMOTION program. The results from this study show the complexity of implementation in primary care using group leaders employed in mental health and municipal services outside of the schools.

Although the mental health professionals enjoyed the intervention and found it necessary within the services, they still had some concerns regarding future utilization of the program. In Aarons et al.'s (2011) EPIS model, innovation-values fit is considered an essential aspect of implementation. These authors state that interventions that fit the goals, tasks and duties of organizations and individuals, as well as other administrative and practical tasks, will most likely promote implementation. In this study, the results implied that while the group leaders appreciated developing their competence within this area, it was also clear that some features of the program were unacceptable to carry out (e.g., the extensive manual). Other studies have also highlighted the importance of satisfaction with and the feasibility of an intervention for continued implementation (Forman, Fagley, Chu, & Walkup, 2012; Proctor et al., 2011).

However, positive attitudes towards the intervention are not enough to change practice, proving that other factors, such as the implementing organization, have an impact on the active implementation process. The results from this study further showed that the lack of time to conduct the EMOTION program and the heavy workload were some of the most prominent hindering factors of implementation. In fact, approximately 50% of the participants were unsure if they would continue as group leaders after the initial project period. Furthermore, 73% of the group leaders indicated that the heavy workload interfered with the continuation of EMOTION. Additionally, during the interviews, the group leaders highlighted the limited time and resources assigned to run the intervention as important factors for the future utilization of the program. These findings reflected a huge barrier to continued use of the program that unfortunately is not unique to this particular study. Recent studies (Beidas et al., 2016; Bond et al., 2014) support these results, identifying time issues and the limited resources allocated (e.g., time, money, tangible support) as major threats to the implementation of interventions. Furthermore, Langley et al. (2010) reported that competing responsibilities were the strongest barrier during implementation and were also highlighted frequently by the successful implementers as among the main barriers. Similar results have been shown previously as well (Forman, Olin, Hoagwood, Crowe, & Saka, 2009) and clearly emphasize the issues related to limited resources when implementing interventions in schools and municipal services, which should be considered in future studies.

Related to the organizational factors, another important issue for the group leaders was the experience of autonomy and support from leaders in the organizations. Particularly in the interviews, the importance of autonomy and being able to manage the time spent on the EMOTION groups was considered a promoting factor for continued use. Additionally, supportive and positive leaders were deemed important, implying that without the direct backing of the leader, further implementation of the program was impossible. Other studies have shown the significant contribution of positive and effective leadership in organizations (e.g., Corrigan & Garman, 1999) and how leadership affects attitudes towards EBPs (Aarons & Sommerfeld, 2012), thereby influencing future use and continued implementation of interventions (Aarons, 2006; Rodriguez, Lau, Wright, Regan, & Brookman-Frazer, 2018). Recently, however, researchers have started to investigate leadership and implementation, particularly looking closer at strategic leadership (e.g., implementation leadership), and which leadership behaviors are important during organizational change (Aarons, Ehrhart, & Farahnak, 2014; Gifford, Graham, Ehrhart, Davies, & Aarons, 2017). This has led to the

development of the Implementation Leadership Scale (ILS; Aarons et al., 2014), which assesses the strategic climate in organizations. Additionally, there are approaches looking specifically at the connection between transformational leadership and ways to integrate it in implementation leadership and implementation effectiveness (A. Richter et al., 2016).

Another important finding from this study was related to the schools' investment and engagement in the implementation of the intervention. Having support from the teachers and school staff in general facilitated the completion of the EMOTION groups, both practically and by raising awareness of internalizing issues in this population. The importance of teachers and school staff in the successful implementation of school-based interventions has been demonstrated in prior studies as well (Domitrovich et al., 2008; Forman et al., 2009; Langley et al., 2010). This highlights the need to include schools to an even greater degree in future studies, especially if the school is still the context of delivery.

Discussion of the main findings in Paper 3

The third article of this thesis investigated the psychometric properties of the observation measure – the CAS-CBT (Bjaastad et al., 2016), which was used to evaluate the group leaders' quality of delivery (fidelity). We assessed the group leaders' adherence and competence during completion of the program manual by scoring video-recorded sessions of the EMOTION groups.

We performed CFA in Mplus, and the results showed that we were not able to replicate the original factor structure proposed by Bjaastad et al. (2016), particularly when we included the items assessing the goals for the session. Internal consistency estimates were, however, good to excellent for the two structures (excluding the session goal items) tested in this study. We further investigated the structure of the instrument and found that when we correlated the residuals of specific items (which theoretically were strongly associated), model fit improved slightly. Finally, we removed the session goals from the scale, which yielded an acceptable model fit, indicating that these items did not fit the scale adequately.

In our model, the theoretical justification for removing the session goal items was based on the structure of the instruments. The observers rated the goals for the sessions independently, and as these items varied from session to session, they accordingly were difficult to adjust to the instrument structure. Furthermore, given the transdiagnostic and comprehensive nature of the program, choosing the goals for the different sessions was not an easy task for the

program developer(s). Furthermore, the items assessing adherence to the session goals correlated better with the items reflecting relational skill than with the items within the same subscale (CBT structure). This could therefore explain why these items did not fit the model originally tested. Model modifications, however, are subject to some skepticism (Schreiber, Nora, Stage, Barlow, & King, 2006). Modifications generally include altering the model parameters, which could remove the researcher from the initial model, and therefore require a theoretical defense (Schreiber et al., 2006).

In the wake of treatment manual development, which subsequently facilitated the monitoring of treatment protocols, the importance of measuring treatment fidelity has begun to be explored. In a systematic review, Prowse, Nagel, Meadows, and Enticott (2015) found that in general, measuring both adherence and competence provided better results on treatment fidelity quality. This indicates that building solid evidence of effectiveness requires adequate measures of treatment quality (Prowse et al., 2015). However, there is limited use of fidelity measures in the field, which creates a valuable opening for future research, proposing the inclusion of such procedures to promote better treatment fidelity.

General discussion

This mixed- and multi-method study has provided insightful knowledge regarding the effectiveness and implementation of an indicated preventive intervention for children with symptoms of anxiety and depression. The transdiagnostic EMOTION program seems to promote a significant reduction in anxious and depressive symptoms in children at risk, confirming the initial hypotheses; the intervention group (IG) showed a larger decrease in child-reported depressive and anxious symptoms than the control condition (CC). The hypothesis was also confirmed regarding parent-reported symptoms of depression, indicating that the IG displayed a greater reduction than the CC. Anxious symptoms as reported by parents in the IG, however, were not significantly different from those in the control group.

Implementing an indicated intervention for children with anxious and depressive symptoms in municipal services requires considerable effort. First, having an effective intervention that is appropriate, accepted, and feasible for the participants and within the service context is important (Aarons, Hurlburt, et al., 2011; Proctor et al., 2011). During the main study, attendance rates were 94% for the children and 75% for the parents. Additionally, after the EMOTION groups began, the dropout rate was low. This generally reflects a high degree of

satisfaction with the program, especially considering its intensity, as 20 child sessions and 7 parent sessions were conducted over a 10-week period.

The group leaders also found the intervention acceptable and highly important; however, there was some apprehension about its feasibility. Conducting the EMOTION program in addition to regular work, as mentioned previously, was one of the major barriers to implementation. Awareness of the complexity of implementing new interventions in municipal services, particularly when schools provide the context of delivery, should be exercised in future studies. Implementation within this context generally involves individuals employed in different organizations who are part of complicated intra- and interorganizational settings, including being subject to external policies (Lewis, Proctor, & Brownson, 2017; Novins et al., 2013). The experienced barriers may therefore be a result of municipal services and organizational leaders not being aware of or capable of handling the efforts needed to implement new interventions, which could indicate poor “readiness for change”. This situation could be accommodated by introducing some requirements for the organizations (e.g., assessing capacity) and preparing the organizations more extensively before starting a new intervention (e.g., increasing knowledge about implementation, especially as it involves organizational leaders).

According to Durlak and DuPre (2008), the main goal is to find the right balance between adapting the intervention to fit local needs and fidelity to the program to ensure that program core components are delivered as intended and hence produce outcomes as anticipated. Although the literature highlights the importance of ongoing fidelity assessments (Botvin, 2004; Novins et al., 2013), fidelity during the implementation of adopted school-based programs is achieved 50% or less of the time (Gottfredson & Gottfredson, 2002). The results of the fidelity scoring (fidelity was rated on a scale from 0 = *(None/Poor skills)* to 6 = *(Thorough/Excellent skills)*) of the sessions in this project indicated a large variation during completion of the groups. The mean adherence score was $M = 3.53$ ($SD = 1.25$), with a range from 0.43 to 6.00 (mean of seven items). The mean competence was $M = 3.59$ ($SD = 1.26$), with a range from 0.25 to 6.00 (mean of four items). The variation between groups was extensive, ranging from 0.25 to 6.00. Compared to those of therapists conducting individual therapy for anxious youths (Bjaastad et al., 2016), the results in this study were somewhat lower. The reasons for this discrepancy could be the different approaches, as some of the tasks are easier to conduct during individual therapy than in indicative prevention in groups. Furthermore, clinicians often have more experience with anxious and depressive children and

how to use CBT, whereas many of those working in the prevention field do not have formal CBT training. In a systematic review conducted by Rapley and Loades (2018), they found few studies, with mixed and inconclusive results, regarding therapists' adherence and competence when treating children during individual CBT, indicating that more research is warranted.

In the TIM study, the EPIS model was chosen to frame the structure of the implementation research. However, as with all such models and frameworks, it is difficult to cover all aspects. Although personal characteristics are mentioned, the EPIS model may not embrace them within an organizational setting, as it is more focused on the inner (and outer) settings and does not sufficiently acknowledge individual contributions. This could be particularly relevant within the Norwegian context, where professionals are often permanently employed and given a high degree of autonomy to conduct their work. Hence, to counter this, we could also have chosen Damschroder and colleagues' (2009) Consolidated Framework for Implementation Research (CFIR), which focuses more on individual characteristics; however, we were mostly oriented towards the organizational settings in this thesis. Furthermore, because the EPIS model focuses on the public service sector, where most children and adolescents receive mental health, we found it most applicable to our study. In the future, emphasizing individual characteristics in the EPIS model more extensively in addition to the organizational factors should be undertaken. Alternatively, integrating the individual characteristics identified in the CFIR model as a framework for the research could be applied.

Methodological considerations

Overall, the effectiveness evaluation of the EMOTION program was conducted with good empirical and methodological quality. However, due to practical considerations, which are a significant part of conducting research in real-life situations, there will always be some issues affecting internal validity. Cluster randomization, for instance, was chosen to address potential spill-over effects. Hence, randomization took place at the school level, and the schools continued in the same condition (intervention or control) throughout the project period. Therefore, the children and families were aware of which condition they were assigned after the first semester their school participated. This could lead to a contamination effect between different participants at the same school. Additionally, bringing attention to internalizing issues might have contributed to a positive effect in both conditions, even though the children in the control schools did not receive the intervention. This could reduce or diminish the difference between the two experimental conditions and lead to more

conservative results. Further, the “blinding” of which condition the children were recruited to was impossible. Thus, the elevated symptoms consistently reported by the children in the intervention condition could be a response to *self-selection bias*. Another bias to consider is *performance bias*, in which the children and parents anticipate an effect because they know they are part of the intervention; in addition, those in the control condition could seek other forms of care knowing they are in the control group (Porta, 2016).

Using only child self-reports for the recruitment of such young children also merits some attention. Querying multiple informants (e.g., parents, teachers) is generally considered the best approach to ascertaining a child’s mental health status. However, previous research has shown a divergence between child and parent reports (De Los Reyes et al., 2015; Villabø, Gere, Torgersen, March, & Kendall, 2012). In this study, the parents consistently rated the children’s symptom level lower than the children themselves did, which clearly could have yielded fewer children to be included. Furthermore, relying on the children’s understanding and interpretation of the instrument questions could be insufficient. However, the measures used were tested prior to this study, including on other Norwegian samples (J. Richter & Sund, 2013; Villabø et al., 2012), and showed good psychometric properties. The psychometrics of MASC screening for anxiety symptoms were assessed during this study, supporting previous results (Martinsen et al., 2017). Furthermore, using diagnostic interviews, which could present a more accurate picture of the children’s health status, is not considered applicable within a school setting. Thus, according to Dierker et al. (2001), rating scales provide acceptable results when screening for symptoms in a preventive setting.

Although ours was a preventive setting, we were recruiting “at-risk” children with increased symptoms of anxiety and depression. Therefore, it was important to identify a cut-off for inclusion to indicate which children experienced symptoms affecting normal function and could benefit from the program. However, having a high cut-off might have excluded many symptomatic children and would have made this a treatment program rather than a prevention. Based on Norwegian, Nordic and international norm studies, as well as discussions with Norwegian experts, the inclusion cut-off in our sample was one SD above the population mean (no upper limit). Not having an upper limit might have led to the recruitment of children fulfilling the criteria for a clinical diagnosis and hence qualifying for treatment. Therefore, children in this study might display higher symptom levels than expected for participants in a purely preventive intervention from which participants fulfilling diagnostic criteria should be excluded. This may have resulted in an overestimation of the treatment effect in this study.

Regression to the mean is also an issue worth discussing, bearing in mind the elevated symptom level of the indicated sample of children in this project. Regression towards the mean implies that at post-assessment, the children who initially reported high symptom levels moved closer to the population mean (Ostermann, Willich, & Lüdtkke, 2008). Having a control group not receiving the intervention typically resolves this issue, thus providing more assurance that the intervention caused the effect. In our study, both conditions potentially regressed towards the mean; however, the intervention group displayed greater symptom reduction than the control group, indicating that the intervention had an effect. Furthermore, because we were recruiting children with symptoms of anxiety and depression, the sample inherently experienced more internalizing problems.

Ultimately, it would have been interesting to investigate the relation between implementation factors and child outcomes. However, as the overall study primarily focused on the recruitment of the children, this resulted in a small sample size of group leaders, which makes it harder to detect any differences. Ogden and Fixsen (2014), among others, have also addressed this issue, emphasizing the need to focus on the research design of implementation studies to facilitate testing of the associations between the experimental variables and outcomes. Additionally, in our study, we did not assign the group leaders a primary or secondary role, and therefore, we had to merge the results from the group leaders in the analyses, creating a group mean score based on the group leader variables. This might have cancelled out any differences and further made it impossible for us to study growth. Additionally, because the group leader pairs often changed from one semester to the next, we were not able to conduct comparisons with the results from previous groups.

Further, this raises a general issue in the implementation field – the methodological challenges, particularly those related to measures. We observed relatively high mean scores and small standard deviations on the group leader questionnaires, indicating low variation in the response categories, at least for some of the questions. This is a general problem in implementation research, as there are few psychometrically validated measures, thereby increasing the use of self-made questionnaires (Lewis et al., 2015; Lewis et al., 2017). According to Lewis and colleagues (2017), the measurement issues encountered in implementation science result in a smaller pool of information that could provide field-targeted implementation strategies to overcome barriers. One reason for this could be the considerable number of theories, models and frameworks applied in the field, which also

produces a taxonomical and linguistic lack of clarity. Hence, there is a need to operationalize and standardize implementation constructs and then develop instruments that measure these accordingly (Lewis et al., 2017; Martinez et al., 2014; Ogden & Fixsen, 2014). In a systematic review conducted by Chaudoir, Dugan, and Barr (2013), they identified 62 available measures, mostly at the organizational, provider, and innovation levels. Few of these were associated with implementation outcomes. They also concluded that specifying and refining the constructs and measurements must be undertaken to improve implementation success. These issues should be addressed in future studies with the goal of investigating the relationship between implementation and outcome empirically.

However, another explanation for the lack of variation between the group leaders could be that in general the implementation was robust and of high quality. Although the group leaders conducting this intervention were primarily working in the municipal services and had varied experience with training in CBT and manualized interventions, all the group leaders received the same follow-up during the project. We maintained high standards for training and supervision, including supervision of supervisors (which is not even common in outpatient clinics). The group leaders attended a three-day training in the program, had weekly supervision by a trained CBT therapist and were offered a booster session at the end of the semesters. Furthermore, we conducted quality assessments of the group leaders (e.g., adherence and competence ratings) upon completion of the groups. Additionally, the group leaders volunteered to participate, as they had a general interest in the topic, which may have resulted in little variation across individuals.

By using a mixed methods design (paper 2), we did address some of the measurement issues we were challenged with. Using qualitative methods in implementation research can be a suitable approach for triangulation (to verify the results) and can provide new information (Palinkas et al., 2011). Qualitative methods are an important information source in the implementation field, particularly when the sample size is limited (Novins et al., 2013; Palinkas et al., 2011). However, we interviewed only the group leaders in our study, and to strengthen the results, it would have been beneficial to interview organizational leaders and school personnel as well.

Finally, validation of instruments is an important asset of psychometric testing, particularly regarding implementation and fidelity. This is because, as previously mentioned, well-established instruments with adequate psychometric properties seem to be missing in this

research area (Lewis et al., 2017; Martinez et al., 2014). Although construct validity represents an important validation approach, a potential limitation in paper 3 could be the lack of other types of validation methods. Assessing convergent/divergent validity against similar instruments could have been beneficial. However, this requires other validated observational tools assessing adherence and competence in CBT for internalizing youths to rate the CAS-CBT against, which seem to be lacking. Additionally, it would require extensive resources, training of raters, and ethical approval to test for convergence and divergence with other observation instruments. Furthermore, one important yet highly understudied issue in general is assessing criterion-related validity regarding the relationship between instruments and the theoretical frameworks defining the constructs included (e.g., organizational culture, leadership) (Chaudoir et al., 2013; Martinez et al., 2014). This is important because we are trying to predict implementation efforts based on the instruments, which are built upon the theoretical frameworks. Hence, information on instruments' concurrent and predictive validity may result in a revision of theoretical constructs and thereby affect the direction of implementation science in the future (Proctor et al., 2009b).

Optimally, a multilevel analysis to assess the between-level (groups) and within-level (group leaders) data would have been the best approach. However, the two group leaders were treated as the unit of analysis at the between level because the composition of the pairs differed at different measurement points. Furthermore, estimating such models requires a substantial amount of data, which was not possible to obtain in this study due to the relatively small sample size.

Future research

Based on the feedback from the participants in the study, a revised version of EMOTION has already been introduced to the services (16 child sessions and 5 parent sessions). This more flexible version of the program should be tested and further developed, with the goal of finding the right balance between session number, adherence and adaptation by the services. Another interesting approach is testing a version where some of the sessions are web-based (completed at home) to minimize the burden on the group leaders. Using an innovative research design, such as a factorial design (Collins, Dziak, Kugler, & Trail, 2014), allows testing of two or more independent variables simultaneously. Hence, investigating different versions of the intervention (e.g., brief vs. long, web-based vs. regular) could provide an

indicated program that is feasible within the prevention field but still provides the amount of treatment that is needed by these children.

Identifying the facilitators of and barriers to future utilization of the program in this context was essential. However, the next step is to gain in-depth knowledge on which factors are critical for obtaining change and maintaining it over time. Hence, it is important to increase the knowledge and follow-up of implementation within the services. Previous research has shown the importance of evaluating organizations' readiness to implement a new intervention and allocating resources to provide an adequate support system for the providers of the intervention (Armenakis, Harris, & Mossholder, 1993; Lehman et al., 2002; Wanless & Domitrovich, 2015). To advance this, an increased focus on the organizational implementation context (OIC), as suggested by Lyon and colleagues (2018), is relevant for further development of the implementation process in an organization. The OIC reflects specific factors that are important during implementation within a specific setting (e.g., schools) that are closely linked to the implementers' behavior. One of the constructs presented as part of the OIC is strategic implementation leadership (ILS), comprising specific behaviors facilitating or impeding implementation (e.g., supporting and preserving during the implementation process). To address some of the main barriers from the TIM study, focusing on leadership training and developing implementation leadership is important for future use of interventions (Aarons et al., 2017; Lyon et al., 2018), particularly in the municipal services.

Additionally, specifying the group leader tasks, especially the main goals of each session, and assigning the group leaders a primary and a secondary role could promote the feasibility of the intervention. Both considering the research and emphasizing the structural and relational processes of program implementation could enhance some of the core elements of the program and help identify the most important change mechanisms. Through continued feedback to the providers using a measurement feedback system (MFS; Bickman, 2008; Bickman, Kelley, & Athay, 2012), for instance, the group leaders could tailor the intervention to the children's development during the completion of the groups. Similarly, an implementation and/or research team could monitor the implementation process more closely, thereby capturing unanticipated influences and actuating efforts if necessary.

Options to address the issues regarding the group leader sample size and statistical power encountered in this study could include a continued focus on data collection (e.g., estimating statistical power for group leader data) or pooling datasets. Another possibility could be to

explore new research designs (e.g., rollout designs). Rollout designs consist of several cohorts of providers or organizations randomized in sequence, and the cohorts then function as the control group for the previous cohort (Landsverk, Brown, Rolls Reutz, Palinkas, & Horwitz, 2011; Novins et al., 2013). Using established and validated measures with good psychometric properties will also continue to remedy the methodological issues in the field. This also includes revising the CAS-CBT (Bjaastad et al., 2016) to fit the group format, and tailoring the assessment of the group leaders to this context is a necessary, important effort.

Furthermore, in this study, we did not test or manipulate any of the implementation factors or strategies (e.g., recruitment, training, supervision, administrative support; Fixsen et al., 2009). Research shows, however, that by focusing on implementation outcomes, implementation success can be modeled and tested (Proctor et al., 2011). The implementation outcome is different from the outcomes related to the services or individuals and can be viewed as a way to measure the implementation process. Proctor and colleagues (2011, p. 65) define it as “the effects of deliberate and purposive actions to implement new treatments, practices, and services”, and implementation effectiveness is therefore a measure of how well the intervention was implemented. This is important because we want to distinguish whether an intervention was ineffective within a new setting or whether the intervention was implemented incorrectly (Proctor et al., 2011).

One possibility could be to test different strategies related to training or supervision (e.g., brief vs. intensive, in person vs. video lectures, or high supervision vs. no-supervision/co-supervision). Previous research has not been able to reveal any significant discrepancies between different training approaches (Beidas, Edmunds, Marcus, & Kendall, 2012; Herschell et al., 2009; Rohrbach, Graham, & Hansen, 1993; Vismara, Young, Stahmer, Griffith, & Rogers, 2009). However, including ongoing support or consultation seems to be a distinct implementation strategy that should be included in the training process (Edmunds, Beidas, & Kendall, 2013; Nadeem, Gleacher, & Beidas, 2013). In the future, testing different combinations of training and supervision could therefore be the next step on the agenda.

The testing of different implementation strategies seems to be increasing in the field, and multiple studies are in development (Eiraldi et al., 2016; Kilbourne et al., 2014) but have yet to publish any results. However, as the implementation field is in continuous movement and consistently needs to accommodate complex settings, it is important to recognize that some implementation research questions are not that easily solved with experimental designs (e.g.,

changes in policy); thus, rigorous experimental designs encompassing all of the possible influential variables are impossible to execute.

Conclusion

The main goals of this thesis were to examine the effectiveness of the EMOTION program and to simultaneously investigate the factors associated with the implementation of the program. We also closely examined the quality of delivery, evaluated by assessing the group leaders' adherence and competence. To our knowledge, this is one of the first studies using a transdiagnostic approach targeting anxious and depressive schoolchildren, identifying potential implementation issues at the same time. Although there has been an increased emphasis on implementation and its influence, implementation is notoriously under-studied in large trials.

The results from this thesis indicate that the EMOTION program shows positive results in reducing symptoms of anxiety and depression in children aged 8-12 years, as reported by self-reports. Parents reported a significant symptom reduction for depression but not for anxiousness compared to the levels in the control group. Hence, the initial effects indicate that the children profited from the intervention. Further, targeting both anxiety and depression in a group format was found to be applicable, which implies the possibility of widening the scope of children reached. Thus, delivering effective programs in primary care settings is critical for increasing their potential public health impact.

Implementation is, however, a tedious and laborious process, indicating that support and guidance are needed when adopting innovations in new contexts and populations (Meyers et al., 2012). The findings derived from the current study using both quantitative and qualitative methods indicated that overall, the group leaders found the EMOTION program to be highly relevant within the municipal services. Barriers such as a lack of organizational support, time issues and insufficient resources to conduct the groups did, however, have an impact on intention to continue with the program. Collaboration with the participating schools was also an issue that needs attention in future application of the program. Additionally, an increased focus on further development of appropriate measures to assess implementation factors, including fidelity, was an important learning outcome from this study.

Furthermore, regardless of the growing body of theoretical frameworks and models applied in the field, it is important to preserve an understanding of the implementation process and strive

to carry out each step of the process with high quality. This generally entails an increased focus on high-quality implementation conducted with proper training, ongoing supervision, and the use of well-established instruments to assess the implementation process and keep track of the different implementation strategies. In practice, this will necessitate an increased focus of the service organizations, in collaboration with the program developers, on establishing procedures to accommodate these issues and maintain them over time. When the organizational and other implementation issues are identified, it is imperative to develop strategies to overcome them and thereby improve the implementation process (Aarons et al., 2017). This leads to more effectively implemented interventions, which improves care and, ultimately, benefits the children and families in need.

References

- Aarons, G. A. (2004). Mental health provider attitudes toward adoption of evidence-based practice: The Evidence-Based Practice Attitude Scale (EBPAS). *Mental Health Services Research, 6*. doi:10.1023/b:mhsr.0000024351.12294.65
- Aarons, G. A. (2005). Measuring provider attitudes toward evidence-based practice: Consideration of organizational context and individual differences. *Child and Adolescent Psychiatric Clinics of North America, 14*, 255-271. doi:10.1016/j.chc.2004.04.008
- Aarons, G. A. (2006). Transformational and transactional leadership: Association with attitudes toward evidence-based practice. *Psychiatric Services, 57*. doi:10.1176/ps.2006.57.8.1162
- Aarons, G. A., Ehrhart, M. G., & Fahrenak, L. R. (2014). The implementation leadership scale (ILS): development of a brief measure of unit level implementation leadership. *Implementation Science, 9*, 1-10. doi:10.1186/1748-5908-9-45
- Aarons, G. A., Glisson, C., Hoagwood, K., Kelleher, K., Landsverk, J., & Cafri, G. (2010). Psychometric properties and United States national norms of the Evidence-Based Practice Attitude Scale (EBPAS). *Psychological Assessment, 22*. doi:10.1037/a0019188
- Aarons, G. A., Hurlburt, M., & Horwitz, S. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research, 38*, 4-23. doi:10.1007/s10488-010-0327-7
- Aarons, G. A., Moullin, J. C., & Ehrhart, M. G. (2017). The Role of Organizational Processes in Dissemination and Implementation Research. In R. Brownson, G. Colditz, & E. Proctor (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice* (2 ed.). New York: Oxford University Press.
- Aarons, G. A., & Sommerfeld, D. H. (2012). Leadership, innovation climate, and attitudes toward evidence-based practice during a statewide implementation. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*. doi:10.1016/j.jaac.2012.01.018
- Aarons, G. A., Sommerfeld, D. H., & Willging, C. E. (2011). The soft underbelly of system change: The role of leadership and organizational climate in turnover during statewide behavioral health reform. *Psychological Services, 8*, 269-281. doi:10.1037/a0026196
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11-39). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior Human Decision Processes, 50*, 179 - 211.
- Allan, N. P., Norr, A. M., Capron, D. W., Raines, A. M., Zvolensky, M. J., & Schmidt, N. B. (2015). Specific associations between anxiety sensitivity dimensions and fear and distress Dimensions of emotional distress disorders. *Journal of Psychopathology and Behavioral Assessment, 37*, 67-78. doi:10.1007/s10862-014-9437-y
- Allen, J. D., Shelton, R. C., Emmons, K. M., & Linnan, L. A. (2017). Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination. In R. Brownson, G. Colditz, & E. Proctor (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice* (2 ed.). New York: Oxford University Press.
- American Educational Research Association (AERA). (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association (AERA).
- American Psychiatric Association. (2013). *Diagnostic And Statistical Manual Of Mental Disorders (DSM-5)* (Fifth ed.). Washington, DC: American Psychiatric Association.
- Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. (1995). The development of a short 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. *Journal of Child Psychology and Psychiatry, 43*, 1052-1063. doi:10.1111/1469-7610.00232

- Armenakis, A. A., Harris, S. G., & Mossholder, K. W. (1993). Creating readiness for organizational change. *Human Relations*, *46*, 681-703. doi:10.1177/001872679304600601
- Avenevoli, S., Knight, E., Kessler, R. C., & Merikangas, K. R. (2008). Epidemiology of depression in children and adolescents. In *Handbook of Depression in Children and Adolescents*. (pp. 6-32). New York, NY, US: Guilford Press.
- Balázs, J., Miklósi, M., Keresztény, Á., Hoven, C. W., Carli, V., Wasserman, C., . . . Wasserman, D. (2013). Adolescent subthreshold-depression and anxiety: Psychopathology, functional impairment and increased suicide risk. *Journal of Child Psychology and Psychiatry*, *54*, 670-677. doi:doi:10.1111/jcpp.12016
- Barrett, P. M. (2004a). *FRIENDS for Life: Group leader's Manual for Children*. Brisbane, Australia: Australian Academic Press.
- Barrett, P. M. (2004b). *FRIENDS for Life: Workbook for Children*. Brisbane, Australia: Australian Academic Press.
- Barrett, P. M., Sonderegger, R., & Xenos, S. (2003). Using FRIENDS to combat anxiety and adjustment problems among young migrants to Australia: A national trial. *Clinical Child Psychology and Psychiatry*, *8*, 241-260. doi:10.1177/1359104503008002008
- Bass, B. M. (1985). *Leadership and Performance Beyond Expectations*. New York: Free Press.
- Beardslee, W. R., Gladstone, T. R. G., Wright, E. J., & Cooper, A. B. (2003). A family-based approach to the prevention of depressive symptoms in children at risk: Evidence of parental and child change. *Pediatrics*, *112*, e119-e131. doi:10.1542/peds.112.2.e119
- Beardslee, W. R., Versage, E. M., & Gladstone, T. R. G. (1998). Children of affectively ill parents: A review of the past 10 years. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*, 1134-1141. doi:10.1097/00004583-199811000-00012
- Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. Oxford, England: International Universities Press.
- Beidas, R. S., Edmunds, J., Ditty, M., Watkins, J., Walsh, L., Marcus, S., & Kendall, P. (2014). Are inner context factors related to implementation outcomes in cognitive-behavioral therapy for youth anxiety? *Administration and Policy in Mental Health and Mental Health Services Research*, *41*, 788-799. doi:10.1007/s10488-013-0529-x
- Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: A randomized trial. *Psychiatric Services*, *63*, 660-665. doi:10.1176/appi.ps.201100401
- Beidas, R. S., Mychailyszyn, M. P., Edmunds, J. M., Khanna, M. S., Downey, M. M., & Kendall, P. C. (2012). Training school mental health providers to deliver cognitive-behavioral therapy. *School Mental Health*, *4*, 197-206. doi:10.1007/s12310-012-9074-0
- Beidas, R. S., Stewart, R. E., Adams, D. R., Fernandez, T., Lustbader, S., Powell, B. J., . . . Barg, F. K. (2016). A multi-level examination of stakeholder perspectives of implementation of evidence-based practices in a large urban publicly-funded mental health system. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*, 893-908. doi:10.1007/s10488-015-0705-2
- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, *23*. doi:10.1037/0278-6133.23.5.443
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246. doi:10.1037/0033-2909.107.2.238
- Bickman, L. (2008). A measurement feedback system (MFS) is necessary to improve mental health outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 1114-1119. doi:10.1097/CHI.0b013e3181825af8
- Bickman, L., Kelley, S. D., & Athay, M. (2012). The technology of measurement feedback systems. *Couple & Family Psychology*, *1*, 274-284. doi:10.1037/a0031022
- Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., Kaufman, J., Dahl, R. E., . . . Nelson, B. (1996). Childhood and adolescent depression: A review of the past 10 years. Part I. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*, 1427-1439. doi:10.1097/00004583-199611000-00011

- Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W., Torsheim, T., Havik, O. E., Heiervang, E. R., & Öst, L.-G. (2016). Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. *Psychological Assessment, 28*, 908-916. doi:10.1037/pas0000230
- Bond, G. R., Drake, R. E., McHugo, G. J., Peterson, A. E., Jones, A. M., & Williams, J. (2014). Long-term sustainability of evidence-based practices in community mental health agencies. *Administration and Policy in Mental Health and Mental Health Services Research, 41*, 228-236. doi:10.1007/s10488-012-0461-5
- Botvin, G. J. (2004). Advancing prevention science and practice: Challenges, critical issues, and future directions. *Prevention Science, 5*, 69-72. doi:10.1023/B:PREV.0000013984.83251.8b
- Calear, A. L., & Christensen, H. (2010). Systematic review of school-based prevention and early intervention programs for depression. *Journal of Adolescence, 33*, 429-438. doi:10.1016/j.adolescence.2009.07.004
- Campbell, M. K., Elbourne, D. R., & Altman, D. G. (2004). CONSORT statement: Extension to cluster randomised trials. *BMJ (Clinical research ed.), 328*, 702-708. doi:10.1136/bmj.328.7441.702
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*, 40.
- Chaudoir, S. R., Dugan, A. G., & Barr, C. H. (2013). Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. *Implementation Science, 8*, 22. doi:10.1186/1748-5908-8-22
- Chavira, D. A., Stein, M. B., Bailey, K., & Stein, M. T. (2004). Child anxiety in primary care: Prevalent but untreated. *Depression and Anxiety, 20*, 155-164. doi:10.1002/da.20039
- Collins, L. M., Dziak, J. J., Kugler, K. C., & Trail, J. B. (2014). Factorial experiments: Efficient tools for evaluation of intervention components. *American Journal of Preventive Medicine, 47*, 498-504. doi:10.1016/j.amepre.2014.06.021
- Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *The American Journal of Medicine, 119*, 166.e167-166.e116. doi:10.1016/j.amjmed.2005.10.036
- Corrigan, P. W., & Garman, A. N. (1999). Transformational and transactional leadership skills for mental health teams. *Community Mental Health Journal, 35*, 301-312. doi:10.1023/a:1018757706316
- Costello, E., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60*, 837-844. doi:10.1001/archpsyc.60.8.837
- Damschroder, L., Aron, D., Keith, R., Kirsh, S., Alexander, J., & Lowery, J. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science, 4*, 50.
- Dane, A. V., & Schneider, B. H. (1998). Program Integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review, 18*, 23-45. doi:10.1016/S0272-7358(97)00043-3
- De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*, 858-900. doi:10.1037/a0038498
- Dierker, L. C., Albano, A. M., Clarke, G. N., Heimberg, R. G., Kendall, P. C., Merikangas, K. R., . . . Kupfer, D. J. (2001). Screening for anxiety and depression in early adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*, 929-936. doi:10.1097/00004583-200108000-00015
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., . . . Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion, 1*, 6-28. doi:10.1080/1754730X.2008.9715730
- Durlak, J. A., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350. doi:10.1007/s10464-008-9165-0

- Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25*, 115-152. doi:10.1023/A:1024654026646
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237-256. doi:10.1093/her/18.2.237
- Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research, 20*, 308-313. doi:10.1093/her/cyg134
- Eccles, M., Grimshaw, J., Walker, A., Johnston, M., & Pitts, N. (2005). Changing the behavior of healthcare professionals: the use of theory in promoting the uptake of research findings. *Journal of Clinical Epidemiology, 58*, 107-112. doi:10.1016/j.jclinepi.2004.09.002
- Edmunds, J. M., Beidas, R. S., & Kendall, P. C. (2013). Dissemination and implementation of evidence-based practices: Training and consultation as implementation strategies. *Clinical Psychology, 20*, 152-165. doi:10.1111/cpsp.12031
- EFPA. (2013). *EFPA Review model for the description and evaluation of psychological tests: Test review form and notes for reviewers version 4.2.6*. In.
- Eiraldi, R., Khanna, M. S., Jawad, A. F., Fishman, J., Glick, H. A., Schwartz, B. S., . . . Beidas, R. (2016). A hybrid effectiveness-implementation cluster randomized trial of group CBT for anxiety in urban schools: rationale, design, and methods. *Implementation Science, 11*, 92. doi:10.1186/s13012-016-0453-z
- Farmer, E. M., Burns, B. J., Phillips, S. D., Angold, A., & Costello, E. (2003). Pathways into and thorough mental health services for children and adolescents. *Psychiatric Services, 54*, 60-66. doi:10.1176/appi.ps.54.1.60
- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*. doi:10.1177/1049731509335549
- Fixsen, D. L., Naoom, S., Blase, K., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, the National Implementation Research Network.
- Fixsen, D. L., Naoom, S., Blase, K., & Wallace, F. (2007). Implementation: The missing link between research and practice. *APSAC Advisor, 19*, 4-11.
- Flannery-Schroeder, E., Choudhury, M. S., & Kendall, P. C. (2005). Group and individual cognitive-behavioral treatments for youth with anxiety disorders: 1-Year Follow-Up. *Journal of Cognitive Therapy Research, 29*, 253-259. doi:10.1007/s10608-005-3168-z
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., . . . Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*, 151-175.
- Flodgren, G., Parmelli, E., Doumit, G., Gattellari, M., O'Brien, M. A., Grimshaw, J., & Eccles, M. P. (2007). Local opinion leaders: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews*. doi:10.1002/14651858.CD000125.pub3
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286-299. doi:10.1037/1040-3590.7.3.286
- Forman, S. G., Fagley, N. S., Chu, B. C., & Walkup, J. T. (2012). Factors influencing school psychologists' "Willingness to Implement" evidence-based interventions. *School Mental Health, 4*, 207-218. doi:10.1007/s12310-012-9083-z
- Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health, 1*, 26-36. doi:10.1007/s12310-008-9002-5
- Forman, S. G., Shapiro, E. S., Coddling, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., . . . Stoiber, K. C. (2013). Implementation science and school psychology. *School Psychology Quarterly, 28*, 77-100. doi:10.1037/spq0000019
- Garber, J., Keiley, M. K., & Martin, N. C. (2002). Developmental trajectories of adolescents' depressive symptoms: Predictors of change. *Journal of Consulting and Clinical Psychology, 70*, 79-95. doi:10.1037/0022-006X.70.1.79

- Gifford, W., Graham, I. D., Ehrhart, M. G., Davies, B. L., & Aarons, G. A. (2017). Ottawa Model of Implementation Leadership and Implementation Leadership Scale: Mapping concepts for developing and evaluating theory-based leadership interventions. *Journal of Healthcare Leadership, 9*, 15-23. doi:10.2147/JHL.S125558
- Gillham, J. E., Hamilton, J., Freres, D. R., Patton, K., & Gallop, R. (2006). Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology, 34*, 195-211. doi:10.1007/s10802-005-9014-7
- Ginsburg, G. S., Becker, K. D., Drazdowski, T. K., & Tein, J.-Y. (2012). Treating anxiety disorders in inner city schools: Results from a pilot randomized controlled trial comparing CBT and usual care. *Child & Youth Care Forum, 41*, 1-19. doi:10.1007/s10566-011-9156-4
- Ginsburg, G. S., Becker, K. D., Newman, J., & Nichols, T. (2008). Transporting CBT for childhood anxiety disorders into inner-city school-based mental health clinics. *Cognitive and Behavioral Practice, 15*, 148-158.
- Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health, 93*, 1261-1267.
- Glisson, C., Dukes, D., & Green, P. (2006). The effects of the ARC organizational intervention on caseworker turnover, climate, and culture in children's service systems. *Child Abuse & Neglect, 30*. doi:10.1016/j.chiabu.2005.12.010
- Glisson, C., & Hemmelgarn, A. (1998). The effects of organizational climate and interorganizational coordination on the quality and outcomes of children's service systems. *Child Abuse & Neglect, 22*. doi:10.1016/s0145-2134(98)00005-2
- Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior, 23*, 767-794. doi:10.1002/job.162
- Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., & Zafft, K. M. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science, 16*, 893-926. doi:10.1007/s11121-015-0555-x
- Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency, 39*, 3-35. doi:10.1177/002242780203900101
- Greenberg, M., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention and Treatment, 4*, 1-62. doi:10.1037/1522-3736.4.1.41a
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly, 82*, 581 - 629.
- Heiervang, E., Stormark, K. M., Lundervold, A. J., Heimann, M., Goodman, R., Posserud, M.-B., . . . Gillberg, C. (2007). Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors, and service use. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*, 438-447. doi:10.1097/chi.0b013e31803062bf
- Herschell, A. D., McNeil, C. B., Urquiza, A. J., McGrath, J. M., Zebell, N. M., Timmer, S. G., & Porter, A. (2009). Evaluation of a treatment manual and workshops for disseminating, parent-child interaction therapy. *Administration and Policy in Mental Health and Mental Health Services Research, 36*, 63-81. doi:10.1007/s10488-008-0194-7
- Higham, R., Tharmanathan, P., & Birks, Y. (2015). Use and reporting of restricted randomization: A review. *Journal of Evaluation in Clinical Practice, 21*, 1205-1211. doi:doi:10.1111/jep.12408
- Ialongo, N. S., Edelsohn, G., & Kellam, S. G. (2001). A further look at the prognostic power of young children's reports of depressed mood and feelings. *Child Development, 72*, 736-747. doi:doi:10.1111/1467-8624.00312
- Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology, 89*, 36-51. doi:10.1037/0021-9010.89.1.36

- Keenan, K., Hipwell, A., Feng, X., Babinski, D., Hinze, A., Rischall, M., & Henneberger, A. (2008). Subthreshold symptoms of depression in preadolescent girls are stable and predictive of depressive disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *47*, 1433-1442. doi:10.1097/CHI.0b013e3181886eab
- Kendall, P. C. (2012). Anxiety disorders in youth. In *Child and Adolescent Therapy: Cognitive-Behavioral Procedures* (4th ed., pp. 143-189). New York, NY, US: Guilford Press.
- Kendall, P. C., & Hedtke, K. (2006). *Cognitive-Behavioral Therapy for Anxious Children: Therapist manual* (3 ed.). Ardmore: PA: Workbook publishing.
- Kendall, P. C., Safford, S., Flannery-Schroeder, E., & Webb, A. (2004). Child anxiety treatment: Outcomes in adolescence and impact on substance use and depression at 7.4-year follow-up. *Journal of Consulting and Clinical Psychology*, *72*, 276-287. doi:10.1037/0022-006X.72.2.276
- Kilbourne, A. M., Almirall, D., Eisenberg, D., Waxmonsky, J., Goodrich, D. E., Fortney, J. C., . . . Thomas, M. R. (2014). Protocol: Adaptive Implementation of Effective Programs Trial (ADEPT): Cluster randomized SMART trial comparing a standard versus enhanced implementation strategy to improve outcomes of a mood disorders program. *Implementation Science*, *9*, 132. doi:10.1186/s13012-014-0132-x
- Klein, K., & Sorra, J. (1996). The challenge of innovation implementation. *Academy of Management Review*, *21*.
- Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3 ed.). New York: Guilford Publications.
- Kovacs, M., & Lopez-Duran, N. (2010). Prodromal symptoms and atypical affectivity as predictors of major depression in juveniles: Implications for prevention. *Journal of Child Psychology and Psychiatry*, *51*, 472-496. doi:doi:10.1111/j.1469-7610.2010.02230.x
- Kroes, M., Kalff, A. C., Steyaert, J., Kessels, A. G. H., Feron, F. J. M., Hendriksen, J. G. M., . . . Vles, J. S. H. (2002). A longitudinal community study: Do psychosocial risk factors and child behavior checklist scores at 5 years of age predict psychiatric diagnoses at a later age? *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 955-963. doi:10.1097/00004583-200208000-00014
- Kösters, M. P., Chinapaw, M. J. M., Zwaanswijk, M., van der Wal, M. F., Utens, E. M. W. J., & Koot, H. M. (2017). FRIENDS for Life: Implementation of an indicated prevention program targeting childhood anxiety and depression in a naturalistic setting. *Mental Health & Prevention*, *6*, 44-50. doi:10.1016/j.mhp.2017.03.003
- Lacey, A., & Luff, D. (2001). Qualitative data analysis. *Qualitative Research*, *1*, 303-323.
- Landsverk, J., Brown, C. H., Rolls Reutz, J., Palinkas, L., & Horwitz, S. M. (2011). Design elements in implementation research: A structured review of child welfare and child mental health studies. *Administration and Policy in Mental Health*, *38*, 54-63. doi:10.1007/s10488-010-0315-y
- Langley, A. K., Nadeem, E., Kataoka, S. H., Stein, B. D., & Jaycox, L. H. (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Mental Health*, *2*, 105-113. doi:10.1007/s12310-010-9038-1
- Lee, R. M., & Gortmaker, S. L. (2017). Health Dissemination and Implementation Within Schools. In R. Brownson, G. Colditz, & E. Proctor (Eds.), *Dissemination and Implementation Research in Health: Translating Science to Practice*. UK: Oxford University Press.
- Lehman, W. E. K., Greener, J. M., & Simpson, D. D. (2002). Assessing organizational readiness for change. *Journal of Substance Abuse Treatment*, *22*, 197-209. doi:10.1016/S0740-5472(02)00233-7
- Lewis, C. C., Fischer, S., Weiner, B. J., Stanick, C., Kim, M., & Martinez, R. G. (2015). Outcomes for implementation science: An enhanced systematic review of instruments using evidence-based rating criteria. *Implementation Science*, *10*, 155. doi:10.1186/s13012-015-0342-x
- Lewis, C. C., Proctor, E., & Brownson, R. C. (2017). Measurement Issues in Dissemination and Implementation Research. In R. Brownson, G. Colditz, & E. Proctor (Eds.), *Dissemination and Implementation Research in Health* (2 ed.). New York: Oxford University Press.
- Lewis, C. C., & Simons, A. D. (2011). A pilot study disseminating cognitive behavioral therapy for depression: Therapist factors and perceptions of barriers to implementation. *Administration*

- and Policy in Mental Health and Mental Health Services Research, 38, 324-334.
doi:10.1007/s10488-011-0348-x
- Lillehoj, C. J., Griffin, K. W., & Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and relation to youth outcomes. *Health Education and Behavior, 31*, 242-257. doi:10.1177/1090198103260514
- Luby, J. L. (2010). Preschool depression: The importance of identification of depression early in development. *Current Directions in Psychological Science, 19*, 91-95.
doi:10.1177/0963721410364493
- Ludwig, K., Lyon, A. R., & Ryan, J. L. (Eds.). (2015). *Anxiety in Youth: Assessment, Treatment, and School-Based Service Delivery*. NY: Springer.
- Lyon, A. R., Charlesworth-Attie, S., Vander Stoep, A., & McCauley, E. (2011). Modular psychotherapy for youth with internalizing problems: Implementation with therapists in school-based health centers. *School Psychology Review, 40*, 569-581.
- Lyon, A. R., Cook, C. R., Brown, E. C., Locke, J., Davis, C., Ehrhart, M., & Aarons, G. A. (2018). Assessing organizational implementation context in the education sector: Confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. *Implementation Science, 13*, 5. doi:10.1186/s13012-017-0705-6
- March, J. S. (1997). *Multidimensional Anxiety Scale for Children: Technical manual*. Tonawanda, NY: Multi-Health Systems Inc.
- Marchand, E., Stice, E., Rohde, P., & Becker, C. B. (2011). Moving from efficacy to effectiveness trials in prevention research. *Behaviour Research and Therapy, 49*, 32-41.
doi:10.1016/j.brat.2010.10.008
- Martinez, R. G., Lewis, C. C., & Weiner, B. J. (2014). Instrumentation issues in implementation science. *Implementation Science, 9*, 118. doi:10.1186/s13012-014-0118-8
- Martinsen, K. D., Holen, S., Neumer, S.-P., Waaktaar, T., Rasmussen, L. M. P., Kendall, P. C., & Patras, J. (2017). The Factor Structure of MASC Youth Report in Norwegian School Children. *Journal of Child and Family Studies, 1-9*. doi:10.1007/s10826-017-0705-9
- Martinsen, K. D., Kendall, P. C., Stark, K., & Neumer, S.-P. (2016). Prevention of anxiety and depression in children: Acceptability and feasibility of the transdiagnostic EMOTION program. *Cognitive and Behavioral Practice, 23*, 1-13. doi:10.1016/j.cbpra.2014.06.005
- Martinsen, K. D., Stark, K., Rodriguez, K. O., & Kendall, P. C. (2014). *Mestrende barn Manual*. Oslo: Gyldendal Norsk Forlag.
- Mendel, P., Meredith, L., Schoenbaum, M., Sherbourne, C., & Wells, K. (2008). Interventions in organizational and community context: A framework for building evidence on dissemination and implementation in health services research. *Administration and Policy in Mental Health and Mental Health Services Research, 35*, 21 - 37.
- Metz, A., & Bartley, L. (2012). Active implementation frameworks for program success: How to use implementation science to improve outcomes for children. *Zero to Three Journal, 32*, 11-18.
- Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*, 462-480. doi:10.1007/s10464-012-9522-x
- Mifsud, C., & Rapee, R. M. (2005). Early intervention for childhood anxiety in a school setting: outcomes for an economically disadvantaged population. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 996-1004. doi:10.1097/01.chi.0000173294.13441.87
- Mihalic, S. (2004). The importance of implementation fidelity. *Emotional & Behavioral Disorders in Youth, 4*.
- Milrod, B., Markowitz, J. C., Gerber, A. J., Cyranowski, J., Altemus, M., Shapiro, T., . . . Glatt, C. (2014). Childhood separation anxiety and the pathogenesis and treatment of adult anxiety. *The American Journal of Psychiatry, 171*, 34-43. doi:10.1176/appi.ajp.2013.13060781
- Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing Risks for Mental Disorders; Frontiers for Preventive Intervention Research*. Washington, DC: National Academy Press.
- Muris, P., Merckelbach, H., Mayer, B., & Prins, E. (2000). How serious are common childhood fears? *Behaviour Research and Therapy, 38*, 217-228. doi:10.1016/S0005-7967(98)00204-6
- Muthèn, L., & Muthèn, B. (1998-2010). *Mplus : Statistical analysis with latent variables : User's guide*. Los Angeles, CA: Muthen & Muthen.

- Mychailyszyn, M. P., Brodman, D. M., Read, K. L., & Kendall, P. C. (2012). Cognitive-behavioral school-based interventions for anxious and depressed youth: A meta-analysis of outcomes. *Clinical Psychology: Science and Practice, 19*, 129-153. doi:10.1111/j.1468-2850.2012.01279.x
- Mykletun, A., Knudsen, A. K., & Mathiesen, K. S. (2009). *Psykkiske lidelser i Norge: Et folkehelseperspektiv*. Oslo; Nydalen:
- Nadeem, E., Gleacher, A., & Beidas, R. S. (2013). Consultation as an implementation strategy for evidence-based practices across multiple contexts: Unpacking the black box. *Administration Policy in Mental Health Mental Health Services Research, 40*, 439-450. doi:10.1007/s10488-013-0502-8
- National Research Council Institute of Medicine. (2009). Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities. Committee on the prevention of mental disorders and substance abuse among children, youth, and young adults: Research advances and promising interventions. In M. E. O'Connell, T. Boat, & K. E. Warner (Eds.), *Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education*. Washington (DC): National Academies Press
- NEM National Committee for Medical and Health Research Ethics. *Guidelines by The Norwegian National Research Ethics Committee for medical and health research (NEM)*. www.etikkom.no: The Norwegian National Research Ethics Committees.
- NESH The National Committee for Research Ethics in the Social Sciences and Humanities. (2016). *Guidelines for Research Ethics in the Social Sciences, Humanities, Law and Theology*. www.etikkom.no: The Norwegian National Research Ethics Committees.
- Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science, 10*. doi:10.1186/s13012-015-0242-0
- Novins, D. K., Green, A. E., Legha, R. K., & Aarons, G. A. (2013). Dissemination and implementation of evidence-based practices for child and adolescent mental health: A systematic review. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*, 1009-1025. doi:10.1016/j.jaac.2013.07.012
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*, 33-84. doi:10.3102/0034654307313793
- Odom, S. L. (2008). The Tie That Binds: Evidence-Based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*, 53-61. doi:10.1177/0271121408329171
- Ogden, T., & Fixsen, D. L. (2014). Implementation science: A brief overview and a look ahead. *Zeitschrift für Psychologie, 222*, 4-11. doi:10.1027/2151-2604/a000160
- Ostermann, T., Willich, S. N., & Lüdtke, R. (2008). Regression toward the mean – a detection method for unknown population mean based on Mee and Chua's algorithm. *BMC Medical Research Methodology, 8*, 52. doi:10.1186/1471-2288-8-52
- Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlain, P., Hurlburt, M., & Landsverk, J. (2011). Mixed method designs in implementation research. *Administration Policy in Mental Health Mental Health Services Research, 38*, 44-53. doi:10.1007/s10488-010-0314-z
- Patras, J., Martinsen, K. D., Holen, S., Sund, A. M., Adolfsen, F., Rasmussen, L.-M. P., & Neumer, S.-P. (2016). Study protocol of an RCT of EMOTION: An indicated intervention for children with symptoms of anxiety and depression. *BMC Psychology Vol 4 2016, ArtID 48, 4*.
- Payne, A. A., & Eckert, R. (2010). The relative importance of provider, program, school, and community predictors of the implementation quality of school-based prevention programs. *Prevention Science, 11*, 126-141. doi:10.1007/s11121-009-0157-6
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*, 829-841. doi:10.1037/0022-006X.75.6.829
- Phillips, J. H., Corcoran, J., & Grossman, C. (2003). Implementing a cognitive-behavioral curriculum for adolescents with depression in the school setting. *Children & Schools, 25*, 147-158.

- 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*, 345-365. doi:10.1111/jcpp.12381
- Porta, M. (Ed.) (2016). *Performance Bias* (6 ed.). UK: Oxford University Press.
- Preskill, H., & Torres, R. T. (1999). *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, CA: Sage.
- Proctor, E., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009a). Implementation Research in Mental Health Services: an Emerging Science with Conceptual, Methodological, and Training challenges. *Administration Policy in Mental Health and Mental Health Services Research*, *36*, 24-34. doi:10.1007/s10488-008-0197-4
- Proctor, E., Landsverk, J., Aarons, G. A., Chambers, D., Glisson, C., & Mittman, B. (2009b). Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research*, *36*, 24-34. doi:10.1007/s10488-008-0197-4
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G. A., Bunger, A., . . . Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 65-76. doi:10.1007/s10488-010-0319-7
- Prowse, P., Nagel, T., Meadows, G., & Enticott, J. (2015). Treatment fidelity over the last decade in psychosocial clinical trials outcome studies: A systematic review. *Journal of Psychiatry*. doi:10.4172/Psychiatry.1000258
- Rapley, H., & Loades, M. E. (2018). A systematic review exploring therapist competence, adherence, and therapy outcomes in individual CBT for children and young people *Psychotherapy Research*, 1-10. doi:10.1080/10503307.2018.1464681
- Richter, A., von Thiele Schwarz, U., Lornudd, C., Lundmark, R., Mosson, R., & Hasson, H. (2016). iLead—a transformational leadership intervention to train healthcare managers’ implementation leadership. *Implementation Science*, *11*, 108. doi:10.1186/s13012-016-0475-6
- Richter, J., & Sund, A. M. (2013). Måleegenskaper ved den norske versjonen av Mood and Feelings Questionnaire - Humøret Ditt (MFQ). *PsykTestBarn*, *1*.
- Ringle, V. A., Read, K. L., Edmunds, J. M., Brodman, D. M., Kendall, P. C., Barg, F., & Beidas, R. S. (2015). Barriers to and facilitators in the implementation of cognitive-behavioral therapy for youth anxiety in the community. *Psychiatric Services*, *66*, 938-945. doi:10.1176/appi.ps.201400134
- Ringwalt, C. L., Ennett, S., Johnson, R., Rohrbach, L. A., Simons-Rudolph, A., Vincus, A., & Thorne, J. (2003). Factors associated with fidelity to substance use prevention curriculum guides in the nation’s middle schools. *Health Education Behavior*, *30*. doi:10.1177/1090198103030003010
- Rodgers, A., & Dunsmuir, S. (2015). A controlled evaluation of the ‘FRIENDS for Life’ emotional resiliency programme on overall anxiety levels, anxiety subtype levels and school adjustment. *Child and Adolescent Mental Health*, *20*, 13-19. doi:10.1111/camh.12030
- Rodriguez, A., Lau, A. S., Wright, B., Regan, J., & Brookman-Frazee, L. (2018). Mixed-method analysis of program leader perspectives on the sustainment of multiple child evidence-based practices in a system-driven implementation. *Implementation Science*, *13*, 44. doi:10.1186/s13012-018-0737-6
- Rogers, E. M. (2003). *Diffusion of Innovation* (5 ed.). New York: Free Press.
- Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine*, *22*, 237-260. doi:10.1006/pmed.1993.1020
- Ruffolo, M. C., & Fischer, D. (2009). Using an evidence-based CBT group intervention model for adolescents with depressive symptoms: Lessons learned from a school-based adaptation. *Child & Family Social Work*, *14*, 189-197. doi:10.1111/j.1365-2206.2009.00623.x
- Sales, A., Smith, J., Curran, G., & Kochevar, L. (2006). Models, strategies, and tools. Theory in implementing evidence-based findings into health care practice. *Journal of General Internal Medicine*, *21* S43-S49. doi:10.1111/j.1525-1497.2006.00362.x

- Scaccia, J. P., Cook, B. S., Lamont, A., Wandersman, A., Castellow, J., Katz, J., & Beidas, R. S. (2015). A practical implementation science heuristic for organizational readiness: R = MC(2). *Journal of Community Psychology, 43*, 484-501. doi:10.1002/jcop.21698
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*, 323-338. doi:10.3200/JOER.99.6.323-338
- Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Annals of Internal Medicine, 152*, 726-732. doi:10.7326/0003-4819-152-11-201006010-00232 %J Annals of Internal Medicine
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA, US: Houghton, Mifflin and Company.
- Shaw, D. S., Keenan, K., Vondra, J. I., Delliquardi, E., & Giovannelli, J. (1997). Antecedents of preschool children's internalizing problems: A longitudinal study of low-income families. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 1760-1767. doi:10.1097/00004583-199712000-00025
- Silverman, W. K., Pina, A. A., & Viswesvaran, C. (2008). Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 37*, 105-130. doi:10.1080/15374410701817907
- Skogen, J. C., Smith, O. R. F., Aarø, L. E., Siqveland, J., & Øverland, S. (2018). *Mental health among children and adolescents. Health-promoting and preventive public health interventions. A summary of evidence about effects. [Barn og unges psykiske helse: Forebyggende og helsefremmende folkehelse tiltak. En kunnskapsoversikt]*. Oslo: Folkehelseinstituttet
- Spence, S. H., Najman, J. M., Bor, W., O'Callaghan, M., & Williams, G. M. (2002). Maternal anxiety and depression, poverty and marital relationship factors during early childhood as predictors of anxiety and depressive symptoms in adolescence. *Journal of Child Psychology and Psychiatry, 43*, 457-470. doi:10.1111/1469-7610.00037
- Stamsø, M. A. (Ed.) (2017). *Organisering og finansiering av velferdstjenester*. Oslo: Gyldendal akademisk.
- Stark, K. D., Humphrey, L. L., Crook, K., & Lewis, K. (1990). Perceived family environments of depressed and anxious children: Child's and maternal figure's perspectives. *Journal of Abnormal Child Psychology, 18*, 527-547. doi:10.1007/bf00911106
- Stark, K. D., Schonebelen, S., Simpson, J., Hargrave, J., Molnar, J., & Glen, R. (2007). *Treating Depressed Youth: Therapist manual for 'ACTION'*. Ardmore, PA: Workbook Publishing.
- Stice, E., Shaw, H., Bohon, C., Marti, C. N., & Rohde, P. (2009). A meta-analytic review of depression prevention programs for children and adolescents: Factors that predict magnitude of intervention effects. *Journal of Consulting and Clinical Psychology, 77*, 486-503. doi:10.1037/a0015168
- Stockings, E. A., Degenhardt, L., Dobbins, T., Lee, Y. Y., Erskine, H. E., Whiteford, H. A., & Patton, G. (2016). Preventing depression and anxiety in young people: a review of the joint efficacy of universal, selective and indicated prevention. *Psychological Medicine, 46*, 11-26. doi:10.1017/S0033291715001725
- Streiner, D. L., & Norman, G. R. (2003). *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford; New York: Oxford University Press.
- Tabak, R. G., Khoong, E. C., Chambers, D., & Brownson, R. C. (2012). Bridging research and practice: Models for dissemination and implementation research. *American Journal of Preventive Medicine, 43*, 337-350. doi:10.1016/j.amepre.2012.05.024
- Teubert, D., & Pinquart, M. (2011). A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents. *Journal of Anxiety Disorders, 25*, 1046-1059. doi:10.1016/j.janxdis.2011.07.001
- Tiet, Q. Q., Bird, H. R., Hoven, C. W., Moore, R., Wu, P., Wicks, J., . . . Cohen, P. (2001). Relationship between specific adverse life events and psychiatric disorders. *Journal of Abnormal Child Psychology, 29*, 153-164. doi:10.1023/a:1005288130494

- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1-10. doi:10.1007/bf02291170
- Ukoumunne, O. C., Gulliford, M. C., Chinn, S., Sterne, J. A., Burney, P. G., & Donner, A. (1999). Methods in health service research. Evaluation of health interventions at area and organisation level. *BMJ (Clinical research ed.)*, *319*, 376-379.
- 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 & Adolescent Psychology*, *41*, 75-85. doi:10.1080/15374416.2012.632350
- Vismara, L. A., Young, G. S., Stahmer, A. C., Griffith, E. M., & Rogers, S. J. (2009). Dissemination of evidence-based practice: Can we train therapists from a distance? *Journal of Autism Developmental Disorders*, *39*, 1636. doi:10.1007/s10803-009-0796-2
- Wanless, S. B., & Domitrovich, C. E. (2015). Readiness to implement school-based social-emotional learning interventions: Using research on factors related to implementation to maximize quality. *Prevention Science*, *16*, 1037-1043. doi:10.1007/s11121-015-0612-5
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child & Adolescent Psychology*, *33*, 105-124. doi:10.1207/S15374424JCCP3301_11
- Wei, C., Cummings, C. M., Villabø, M. A., & Kendall, P. C. (2014). Parenting behaviors and anxious self-talk in youth and parents. *Journal of Family Psychology*, *28*, 299-307. doi:10.1037/a0036703
- Wei, C., Hoff, A., Villabø, M. A., Peterman, J., Kendall, P. C., Piacentini, J., . . . March, J. (2014). Assessing anxiety in youth with the Multidimensional Anxiety Scale for Children. *Journal of Clinical Child & Adolescent Psychology*, *43*, 566-578. doi:10.1080/15374416.2013.814541
- Weissman, M. M., Warner, V., Wickramaratne, P., Moreau, D., & Olfson, M. (1997). Offspring of depressed parents. 10 Years later. *Archives of General Psychiatry*, *54*, 932-940.
- Weist, M. D., Rubin, M., Moore, E., Adelsheim, S., & Wrobel, G. (2007). Mental health screening in schools. *Journal of School Health*, *77*, 53-58. doi:10.1111/j.1746-1561.2007.00167.x
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychological Bulletin*, *132*, 132-149. doi:10.1037/0033-2909.132.1.132
- Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist*, *60*, 628-648. doi:10.1037/0003-066X.60.6.628
- Wergeland, G. J. H., Fjermestad, K. W., Marin, C. E., Haugland, B. S.-M., Bjaastad, J. F., Oeding, K., . . . Heiervang, E. R. (2014). An effectiveness study of individual vs. group cognitive behavioral therapy for anxiety disorders in youth. *Behaviour Research and Therapy*, *57*, 1-12. doi:10.1016/j.brat.2014.03.007
- Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., & Christensen, H. (2017). School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review*, *51*, 30-47. doi:10.1016/j.cpr.2016.10.005
- Wood, J. J., McLeod, B. D., Sigman, M., Hwang, W.-C., & Chu, B. C. (2003). Parenting and childhood anxiety: Theory, empirical findings, and future directions. *Journal of Child Psychology and Psychiatry*, *44*, 134-151. doi:10.1111/1469-7610.00106
- Woods, S. A., & West, M. A. (2010). *The Psychology of Work and Organizations*. London (UK): CENGAGE.
- Woodward, L. J., & Fergusson, D. M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*, 1086-1093. doi:10.1097/00004583-200109000-00018
- World Medical Association. (2013). World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects World Medical Association declaration of Helsinki special communication. *JAMA*, *310*, 2191-2194. doi:10.1001/jama.2013.281053

Paper 1

Martinsen, K. D., **Rasmussen, L-M. P.**, Wentzel-Larsen, T., Holen, S., Sund, A. M., Loevaas, M. E., Patras, J., Kendall, P., Waaktaar, T., & Neumer, S.-P. (2018). Prevention of anxiety and depression in school-aged children: Effectiveness of the transdiagnostic EMOTION program. *Journal of Consulting and Clinical Psychology, 87*. doi:10.1037/ccp0000360

BRIEF REPORT

Prevention of Anxiety and Depression in School Children: Effectiveness of the Transdiagnostic EMOTION Program

Kristin D. Martinsen

University of Oslo and Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Oslo, Norway

Lene Mari P. Rasmussen

The Regional Centre for Child and Youth Mental Health and Child Welfare–North, Tromsø, Norway

Tore Wentzel-Larsen and Solveig Holen

Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Oslo, Norway

Anne Mari Sund and Mona Elisabeth S. Løvaas

Norwegian University of Science and Technology

Joshua Patras

The Regional Centre for Child and Youth Mental Health and Child Welfare–North, Tromsø, Norway

Philip C. Kendall

Temple University

Trine Waaktaar

University of Oslo

Simon-Peter Neumer

Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Oslo, Norway

Objective: The objective of the study was to examine the effectiveness of a transdiagnostic program (EMOTION, Coping Kids Managing Anxiety and Depression) targeting symptoms of anxiety and depression in schoolchildren by comparing the intervention condition (EC) to a control condition (CC).

Method: A clustered randomized design was used with schools as the unit of randomization. Children ($N = 1686$) aged 8–12 years in 36 schools completed screening using the Multidimensional Anxiety Scale (MASC-Child) and The Mood and Feelings Questionnaire Short version (SMFQ-Child). Scoring 1 SD above a population-based mean on anxiety and/or depression, 873 children were invited to participate. Intent-to-treat analyses were performed, and mixed-effects models were used. **Results:** Analyses revealed significant reductions of anxious and depressive symptoms as reported by the children, in which children in the intervention condition EC had almost twice the reduction in symptoms compared with the control condition CC. For parent report of the child's depressive symptoms, there was a significant decrease of symptoms in the intervention condition EC compared with CC. However, parents did not report a significant decrease in anxious symptoms in the intervention condition EC as compared with CC. **Conclusion:** A transdiagnostic prevention program, provided in schools, was successful in reducing youth-reported symptoms of anxiety and depression and parent-reported depression. The EMOTION program has the potential to reduce the incidence of anxious and depressive disorders in youth.

This article was published Online First December 13, 2018.

Kristin D. Martinsen, Department of Psychology, University of Oslo, and Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Oslo, Norway; Lene Mari P. Rasmussen, The Regional Centre for Child and Youth Mental Health and Child Welfare–North, Tromsø, Norway; Tore Wentzel-Larsen and Solveig Holen, Centre for Child and Adolescent Mental Health, Eastern and Southern Norway; Anne Mari Sund, Regional Centre for Child and Youth Mental Health and Child Welfare, Norwegian University of Science and Technology; Mona Elisabeth S. Løvaas, Institute of Psychology, Norwegian University of Science and Technology; Joshua Patras, The Regional

Centre for Child and Youth Mental Health and Child Welfare–North; Philip C. Kendall, Department of Psychology, Temple University; Trine Waaktaar, Department of Psychology, University of Oslo; Simon-Peter Neumer, Centre for Child and Adolescent Mental Health, Eastern and Southern Norway.

Two of the authors (Kristin D. Martinsen and Philip C. Kendall) receive royalties from sale of the EMOTION manuals in Norway.

Correspondence concerning this article should be addressed to Kristin D. Martinsen, Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Gullhaugveien 1-3, 0484 Oslo, Norway. E-mail: kristin.martinsen@r-bup.no

What is the public health significance of this article?

Anxiety and depression are common in youth and have unwanted effects on their functioning. Targeting both anxiety and depression in one protocol has important public health significance: Symptom levels can be reduced, thus preventing children from developing full-blown disorders.

Keywords: anxiety, depression, transdiagnostic intervention, prevention, early intervention

Anxiety and depression are prevalent and impairing disorders in childhood (e.g., Merikangas, Nakamura, & Kessler, 2009). The disorders often co-occur and may result in greater impairment and worse prognosis (Cummings, Caporino, & Kendall, 2014). Youth with subclinical levels of anxious and depressive symptoms experience significant impairment, and the symptoms predict later disorders (Kovacs & Lopez-Duran, 2010; Pine, 2007). Despite the high prevalence and negative sequela, there is a gap between the children in need and those few receiving care (Chavira, Stein, Bailey, & Stein, 2004; Heiervang et al., 2007). Prevention in a

school setting with early identification and initiation of early symptom-reducing interventions may bridge this gap. Previous research suggests modest but positive effects regarding prevention of anxiety and depression in school settings (e.g., Werner-Seidler, Perry, Calear, Newby, & Christensen, 2017). Transdiagnostic interventions targeting more than one disorder/problem are promising approaches to tackle both symptom presentations in anxious and sad children (Ehrenreich-May & Chu, 2014).

The present study evaluated the effectiveness of a 10-week transdiagnostic indicated prevention program (EMOTION, Coping

Table 1
Eligibility Criteria and Demographics

Eligibility criteria			
Inclusion		Exclusion	
Children between 8 and 12 years		Mental retardation	
Child scoring 1 <i>SD</i> > mean on measure of anxiety or/and	MASC-Child Girls: IC = 4.6 + 1 (<i>SD</i>) Boys: IC = 3.9 + 1.5 (<i>SD</i>)	Pervasive developmental disorders	
Child scoring 1 <i>SD</i> > mean on measures of depression	SMFQ-Child Boys/girls: IC = 3.8 + 3.6 (<i>SD</i>)	Not able to benefit from a group intervention	
	<i>M (SD)</i>	<i>% (N)</i>	Sign. diff. EC vs. CC
Baseline demographic characteristics of sample (<i>N</i> = 795)			
Child age, years (EC > CC)	10.1 (.90)		<i>p</i> < .013
Gender			<i>p</i> = .053
Male		42.0 (334)	
Female		58.0 (461)	
Participants per grade			<i>p</i> = .064
Third grade = younger		4.4 (35)	
Fourth grade = younger		36.6 (291)	
Fifth grade = older		45.7 (363)	
Sixth grade = older		13.3 (106)	
Race/ethnicity (mothers)			<i>p</i> = .176
Caucasian/Nordic/Western European		93.8	
Biological parents' marital status			<i>p</i> = .141
Married		78.5	
Not married, divorced, or widowed		21.5	
Living with			
Both parents		71.2	<i>p</i> = .139
Mother		7.6	<i>p</i> = .916
Father		.9	<i>p</i> = .754
Other		20.3	
Education			<i>p</i> = .921
More than 4 years of higher education		66.8	
Employment			<i>p</i> = .696
Full time		74.9	
Annual household income > NOK*			
500,000		82.4	<i>p</i> = .897
Stressing life events (EC > CC)			<i>p</i> > .001

Note. Sign. diff. = significant difference; MASC-Child = Multidimensional Anxiety Scale-child version (March, 1997); SMFQ-Child = Short Mood and Feelings Questionnaire-child version (Angold et al., 1995); NOK = Norwegian kroner.

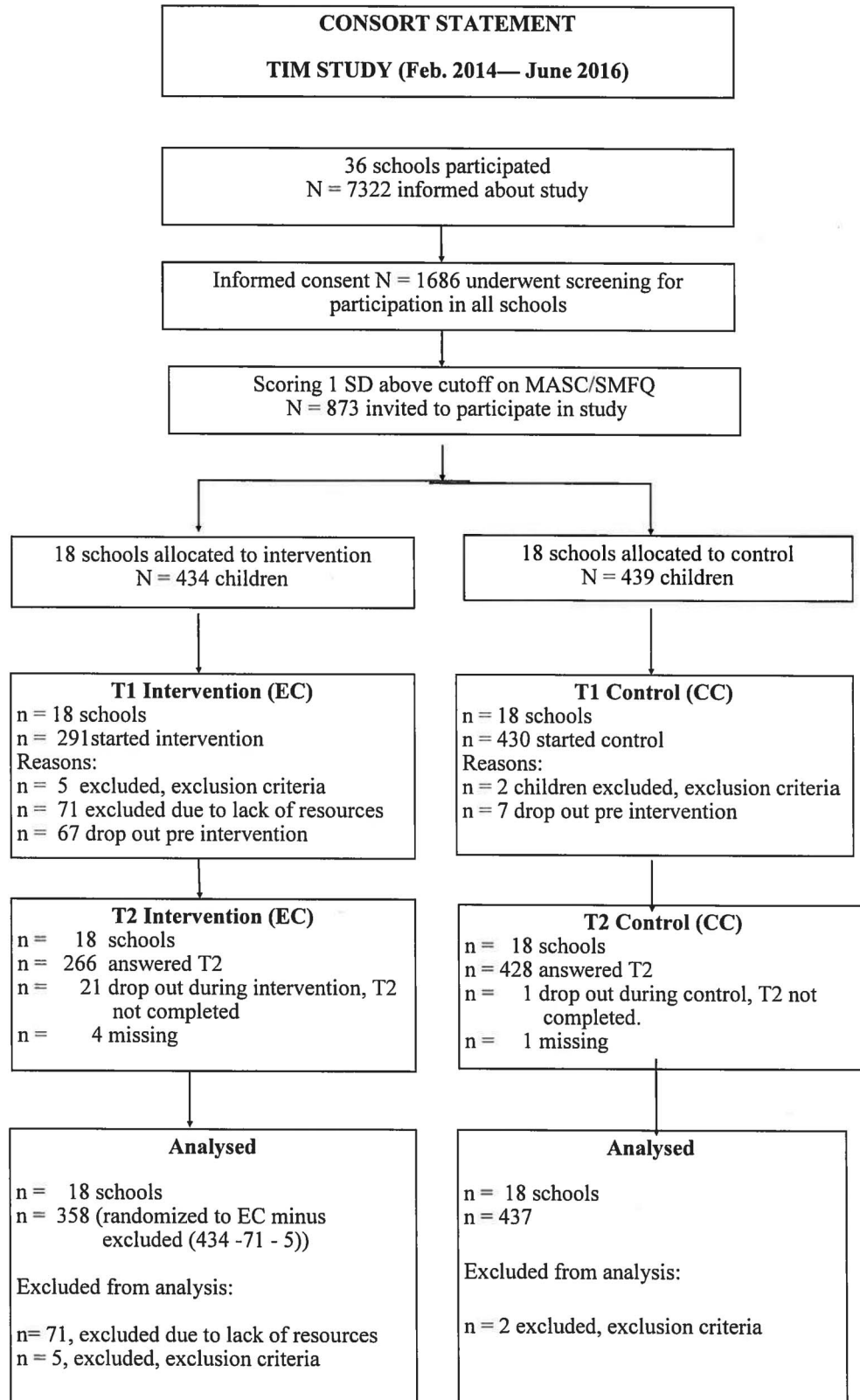


Figure 1. A consort style flow chart of participants through the study.

Kids Managing Anxiety and Depression; Kendall, Stark, Martinsen, O’Neil, & Arora, 2013) targeting anxious and depressive symptoms in children aged 8–12 years compared with a control condition (CC). We hypothesized that the intervention would be more effective than CC as measured by a decrease in symptoms of anxiety and symptoms of depression reported by children and parents. A prior study (Martinsen, Kendall, Stark, & Neumer, 2016) found high acceptability. The current study is the largest to date investigating the effectiveness of a transdiagnostic prevention program in schools.

Method

Study Design and Participants

This study used a clustered, randomized design; for a description of protocol, see Patras et al. (2016). Schools (36 from seven sites in Norway) were randomized. Allocation of the schools to (a) EMOTION intervention (EC) or (b) control condition (CC) involved pairing schools based on geography, school size, and demography and then randomly assigning schools. The Regional Committees for Medical and Health Research Ethics (2013/1909/REK South-East) approved the study.

Recruitment used multiple gating because symptomatic children were the target group for the intervention. Children and parents were informed about the study, and then children experiencing symptoms of anxiety and/or depression and with parental consent were screened. Inclusion/exclusion criteria are in Table 1. The parents of children scoring above the cutoff completed questionnaires. For demographics and flow of children in study, see Table 1 and Figure 1.

Measures

Multidimensional Anxiety Scale for Children (MASC-C/P; March, 1997). This 39-item, child self-report, assesses anxiety in youth ages 8–19 years during the last 2 weeks. Internal consistency of the MASC-Child in the present study was $\alpha = .91$ and $\alpha = .90$ for MASC-Parent.

Mood and Feelings Questionnaire Short version (SMFQ-C/P; Angold, Costello, Messer, & Pickles, 1995). The SMFQ has 13 questions assessing cognitive, affective, and behavioral-related depressive symptoms in youth ages 8–18 years during the last 2 weeks. Internal consistency of the SMFQ-Child in the present study was $\alpha = .94$ and for the parent version SMFQ-Parent, $\alpha = .88$.

The Intervention and Procedures

The indicated preventive intervention was the Norwegian version of the transdiagnostic EMOTION, Coping Kids Managing Anxiety and Depression program (Martinsen, Kendall, Stark, Rodriguez, & Arora, 2014) for youth aged 8–12 years considered at risk for emotional difficulties. EMOTION is cognitive behavioral and based on the notion that anxiety and depression arise from a combination of a diathesis that in the presence of stress leads to their expression. The intervention targets disturbances in cognition, affect regulation, problem solving, and coping skills that are indicated as transdiagnostic mechanisms of change (Kendall et al., 2014). The EMOTION intervention includes group meetings with children and with their parents (see Table 2).

Primarily psychologists and school health nurses provided the EMOTION intervention after a 3-day training. Cognitive behavioral therapy (CBT) supervisors gave weekly supervision to EMOTION group leaders. The CC involved normal contact with the school health nurse/physician.

Statistical Analysis

Power calculations accounted for multilevel data with an effect size of 0.35, power of 0.80, and an alpha of 0.05 (see also Patras et al., 2016). Accordingly, the number of children needed was 630 recruited from 36 schools.

Mixed-effects models were used, giving valid inference for missing at random values in dependent variables. Fixed effects included a time by randomization group interaction, and analyses were adjusted for gender and age group (third and fourth grade = younger; fifth and sixth grade = older). Subgroup analyses for gender and age group were performed; results can be obtained from the first author. The missing-at-random assumption was supported by statistical analysis.

Intent-to-treat analysis was used. The statistical program IBM SPSS (version 23; Oslo, Norway) was used for descriptive analyses. Estimation of mixed-effects models used the R (The R Foundation for Statistical Computing, Vienna, Austria) package nlme.

Results

Means on primary outcomes of anxiety and depression as reported by children and parents are presented in Table 3.

Table 2
EMOTION Program: Session Content for the Child and Parent meetings

Child meetings		Parent meetings	
Session	Content	Session	Content
1–4	Psychoeducation (model, feelings, coping strategies)	1–2	Psychoeducation and positive parenting
5–10	Problem solving and introduction to cognitive restructuring	3–4	Positive reinforcement, introducing exposure, and behavioral activation
10–17	Exposure/behavioral activation, cognitive restructuring, and self-schema	5–6	Problem solving, exposure, behavioral activation, and cognitive restructuring
17–20	Integration of skills, preparing closure	7	Integration and closure

Note. EMOTION = Coping Kids Managing Anxiety and Depression program (Martinsen et al., 2014).

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
Means and Standard Deviations of Primary Outcome Measures Before and After Intervention

Measure	Before intervention						After intervention					
	Intervention (N = 358)			CC (N = 437)			Intervention (N = 266)			CC (N = 428)		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Children												
MASC-Child												
All	358	64.7	13.4	437	62.4	13.6	266	53.5	18.2	428	57.8	16.0
Boys	137	60.5	13.5	197	58.5	13.6	96	48.3	18.2	192	52.4	16.5
Girls	221	67.3	12.7	240	65.6	12.8	170	56.4	17.6	236	60.5	16.1
Third and fourth grades	142	64.4	12.9	184	62.3	13.0	115	55.2	19.7	182	58.9	15.7
Fifth and sixth grades	216	64.9	13.8	253	62.4	14.0	150	52.1	16.9	246	57.0	16.1
SMFQ-Child												
All	358	10.4	5.2	437	9.5	4.6	265	8.0	5.7	428	8.0	5.3
Boys	137	10.1	5.2	297	9.4	4.8	96	7.5	5.5	192	7.3	5.3
Girls	221	10.4	5.2	241	10.1	4.6	169	8.3	5.8	236	8.7	5.4
Third and fourth grades	142	9.2	4.5	184	9.2	4.6	115	8.3	5.8	182	7.6	5.3
Fifth and sixth grades	216	11.1	5.5	253	9.7	4.6	150	7.9	5.6	246	8.3	5.4
	Intervention (N = 268)			CC (N = 301)			Intervention (N = 193)			CC (N = 228)		
Parents												
MASC-Parent	268	46.1	15.2	301	40.6	14.7	193	43.9	16.3	228	39.5	16.1
SMFQ-Parent	268	6.6	5.1	301	4.6	4.4	193	5.0	4.4	228	4.3	4.2

Note. MASC = Multidimensional Anxiety Scale—child and parent version (March, 1997); SMFQ = Short Mood and Feelings Questionnaire—child and parent version (Angold et al., 1995).

Intervention Effects: Children

We first ran the analyses with schools included. This multi-level model was unstable for anxiety and within some subgroups for depression, so models were run without the school level for child and parent data. The results are shown in Table 4. The interaction of Time and Condition was significant, indicating a larger reduction in anxious symptoms in the EC compared with CC. In the EC, there was a reduction in anxious

symptoms of 11.83 points, corresponding to a reduction between 17.4% and 19.7%, depending on gender and age group. In CC, the reduction was 4.63 points, corresponding to a reduction between 7.0% and 8.0%, depending on gender and age group. There was a significant difference between the EC and CC at posttreatment at which the CC youth were 5.35 points higher than the EC youth (see Figure 2a). We found a significant difference in the two conditions for gender, in which girls had

Table 4
Model-Based Estimates for Development in Anxious and Depressive Symptoms, Child and Parent Report

Variable	Child report			Parent report		
	Coefficient	95% CI	p value	Coefficient	95% CI	p value
MASC-Child/Parent						
Group by time interaction	7.20*	[4.92, 9.48]	<.001	.78*	[-1.24, 2.80]	.449
EC vs. CC at T1	1.85	[-.21, 3.92]	.079	5.00	[2.50, 7.50]	<.001
EC vs. CC at T2	-5.35	[-7.57, -3.12]	<.001	4.22	[1.54, 6.90]	.002
T2 vs. T1, EC	-11.83	[-13.60, -10.06]	<.001	-2.30	[-3.76, -.83]	.002
T2 vs. T1, CC	-4.62	[-6.07, -3.19]	<.001	-1.52	[-2.91, -.13]	.033
Girls vs. boys	6.99	[5.17, 8.81]	<.001	2.48	[.08, 4.88]	.043
Older vs. younger	-.97	[-2.79, .85]	.296	1.26	[-1.15, 3.66]	.305
SMFQ-Child/Parent						
Group by time interaction	.81*	[.04, 1.58]	.040	1.31*	[.64, 1.98]	<.001
EC vs. CC at T1	.73	[.01, 1.45]	.047	2.06	[1.31, 2.80]	<.001
EC vs. CC at T2	-.08	[-.85, .69]	.838	.75	[-.06, 1.55]	.069
T2 vs. T1, EC	-2.31	[-2.91, -1.71]	<.001	-1.66	[-2.15, -1.17]	<.001
T2 vs. T1, CC	-1.50	[-1.99, -1.01]	<.001	-.36	[-.82, .11]	.133
Girls vs. boys	.98	[.35, 1.62]	.003	-.17	[-.87, .53]	.636
Older vs. younger	.76	[.12, 1.40]	.020	.37	[-.33, 1.08]	.295

Note. MASC-Child/Parent = Multidimensional Anxiety Scale—child and parent version (March, 1997); SMFQ-Child/Parent = Short Mood and Feelings Questionnaire—child and parent version (Angold et al., 1995). Significant findings are in bold.

* Interaction coefficients: slope differences between intervention groups.

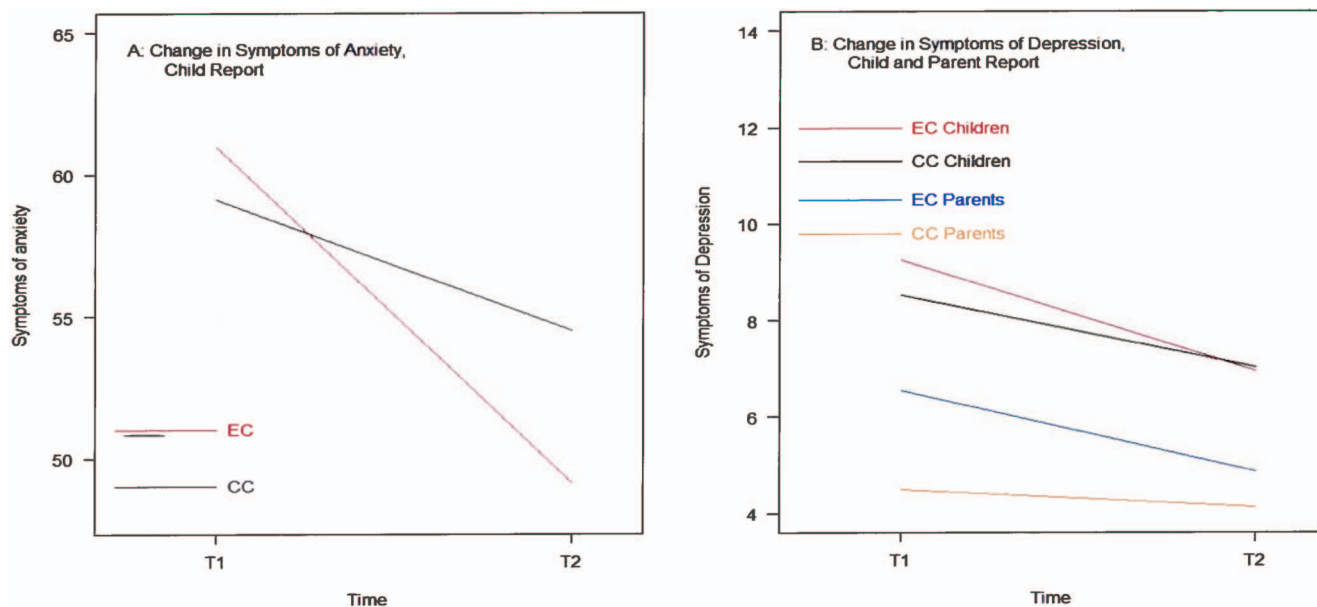


Figure 2. Significant interaction: symptoms of anxiety (a) and depression (b). See the online article for the color version of this figure.

6.99 higher scores than boys. The difference by age group was not significant in the two conditions.

For depressive symptoms, the Time \times Condition interaction was significant, $p = .04$. The intervention resulted in a decrease in depressive symptoms of 2.31 points, corresponding to a reduction between 21.0% and 25.0%, depending on gender and age group. The CC reduction was 1.50 points, corresponding to 14.6% and 17.6%. Before intervention, the difference between the conditions was significant, where CC was 0.73 points lower than EC. After intervention, the difference was not significant (see Figure 2b).

Intervention Effects by Parents' Report

Parent report was collected from 615 parents, in which 568 answered both primary outcome questions before ($n = 268$ EC, $n = 300$ CC), and 421 parents provided answers after intervention ($n = 193$ EC, $n = 228$ CC). Nonresponders at both Time (T1) and T2 were excluded from analysis.

The Time \times Condition interaction was not significant for parent-reported anxiety (see Table 4). There were significant differences between conditions both before and after intervention. Before and after intervention, the parent-reported EC scores were higher than CC.

There was a significant parent-reported Time \times Condition interaction on child depressive symptoms (see Table 4). The pre-intervention parent-reported symptoms were higher in the EC with 2.06 points, $p < .001$. After intervention the difference was not significant (see Figure 2b). The adjustment variables age and gender were not significant.

Discussion

The present results indicate that a transdiagnostic program produced significant reductions in anxious symptoms as reported

by the children. In fact, children who received the EC reported more than twice the reduction in anxious symptoms as compared with CC. The results also indicated a significantly higher reduction in child-reported depressive symptoms for the EC compared with CC. Hence, the EC condition was more effective than CC as measured by a decrease in child-reported depressive and anxious symptoms. Parents also reported significantly higher reductions in depressive symptoms in the EC compared with CC. Parent report of change in anxious symptoms was not significant.

The positive effect of the EMOTION intervention on child-reported anxious symptoms is in accordance with previous research in which children with anxious symptoms benefited from CBT (e.g., Teubert & Pinquart, 2011). Indeed, the findings are consistent with the summary of school-based CBT interventions by Mychailyszyn, Brodman, Read, and Kendall (2012): Youth with elevated levels of anxious symptoms who received an intervention had significantly greater reductions in symptomatology than did controls. Research has also shown that (a) childhood anxiety symptoms are a risk factor for the development of anxiety disorders (Pine, 2007) and (b) high levels of anxiety predict high levels of depressive symptoms later (Goodwin, Fergusson, & Horwood, 2004; Kovacs et al., 2010). It has been suggested that anxiety has depressogenic effects, in which anxiety-driven behaviors can result in feelings of sadness (Cummings et al., 2014; Garber & Weersing, 2010). Accordingly, reductions in anxiety could change the developmental trajectory—preventing later anxiety and depressive disorders.

CBT has been found to be effective for preventing depression in youth (e.g., Clarke et al., 2001). Some studies indicate lower response rates to CBT (March et al., 2004), whereas others have indicated better response rates (Stark, Streusand, Prerna, & Patel, 2012). Mychailyszyn et al. (2012) reported that youth with elevated symptoms of depression receiving an intervention did not

get greater symptom reductions than did controls. Stice, Shaw, Bohon, Marti, and Rohde (2009), however, reported that in 13 of 32 prevention programs, the interventions showed greater decreases in symptoms compared with controls. In our study, the EC condition had a significantly greater decrease of depressive symptoms than CC. Subclinical depressive symptoms are meaningful predictors for later development of disorders (e.g., Kovacs et al., 2010), and for each depressive symptom, the risk for a later disorder increases about 2-fold (Keenan, Feng, Hipwell, & Klostermann, 2009). Hence, even modest reductions in depressive symptoms may be important for long-term prevention. Preventing or delaying the onset of disorders can have public health benefits: Stockings et al. (2016) reported that preventive programs were associated with a decrease in risk for internalizing disorder onset.

Although the EC had larger symptom reductions than CC, both conditions showed a decrease in symptom levels. Some reductions among controls is not uncommon (e.g., Kendall, Hudson, Gosch, Flannery-Schroeder, & Suveg, 2008). It is also possible that controls learned coping skills because teachers in control schools attended workshops on how to help anxious/sad children.

Parents reported that children in the EC group had a significantly greater reduction in symptoms of depression than CC, although this was not the case for anxious symptoms. Note that parents reported lower symptom levels than the children. Although having multiple informants is recommended, parent-child disagreement is common (e.g., De Los Reyes et al., 2015). This is especially so for internalizing problems that are difficult for parents to identify (Comer & Kendall, 2004) and possibly to observe changes in these symptoms.

Before participating, EC children reported significantly higher depressive scores than CC children (see Table 3). This difference is surprising, given randomization. Examining parent-reported demographics (see Table 1) revealed higher preintervention child stress levels in the EC, which could contribute to the difference. Furthermore, there was a higher dropout before intervention in the EC condition than in CC. The intensity of the intervention may account for the higher dropout, and initiatives to make the intervention more flexible could be important for dissemination.

The study had several strengths: It was conducted in the real world with group leaders conducting EC groups in addition to usual work load. Children were recruited from urban and rural schools. Established measures were used to identify and recruit children, treatment integrity was secured, and sound statistical methods were used. However, limitations merit mentioning: A low rate of the overall school population participated in the study because at-risk children were targeted, knowledge about the school being in CC or EC condition could have influenced the recruitment and/or the reporting of symptoms, and recruitment was based on child report. Although screening all children could have increased the participation rate, this was not possible because of Norwegian ethical guidelines. Because the aim was to recruit children with elevated symptoms (i.e., an indicated approach), the sample exhibited more problems than many schoolchildren.

Conclusion

Children at risk for developing internalizing disorders benefitted from receiving a transdiagnostic intervention with significantly

higher reduction in both anxious and depressive self-reported symptoms and depressive symptoms as reported by parents.

Future research could focus on identifying which specific mechanisms account for the reduction in anxious and depressive symptoms, possibly done through dismantling studies. Such studies could include functional outcomes and innovative research designs. When implemented in community settings, the EMOTION program holds the promise of being an effective preventive intervention with the potential of reducing the incidence of anxious and depressive disorders in youth.

References

- Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, 5, 237–249.
- Chavira, D. A., Stein, M. B., Bailey, K., & Stein, M. T. (2004). Child anxiety in primary care: Prevalent but untreated. *Depression and Anxiety*, 20, 155–164. <http://dx.doi.org/10.1002/da.20039>
- Clarke, G. N., Hornbrook, M., Lynch, F., Polen, M., Gale, J., Beardslee, W., . . . Seeley, J. (2001). A randomized trial of a group cognitive intervention for preventing depression in adolescent offspring of depressed parents. *Archives of General Psychiatry*, 58, 1127–1134. <http://dx.doi.org/10.1001/archpsyc.58.12.1127>
- Comer, J. S., & Kendall, P. C. (2004). A symptom-level examination of parent-child agreement in the diagnosis of anxious youths. *Journal of the American Academy of Child & Adolescent Psychiatry*, 43, 878–886. <http://dx.doi.org/10.1097/01.chi.0000125092.35109.c5>
- Cummings, C. M., Caporino, N. E., & Kendall, P. C. (2014). Comorbidity of anxiety and depression in children and adolescents: 20 years after. *Psychological Bulletin*, 140, 816–845. <http://dx.doi.org/10.1037/a0034733>
- De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141, 858–900. <http://dx.doi.org/10.1037/a0038498>
- Ehrenreich-May, J., & Chu, B. C. (Eds.). (2014). *Transdiagnostic treatments for children and adolescents: Principles and practice*. New York, NY: Guilford Press.
- Garber, J., & Weersing, V. R. (2010). Comorbidity of anxiety and depression in youth: Implications for treatment and prevention. *Clinical Psychology: Science and Practice*, 17, 293–306. <http://dx.doi.org/10.1111/j.1468-2850.2010.01221.x>
- Goodwin, R. D., Fergusson, D. M., & Horwood, L. J. (2004). Early anxious/withdrawn behaviours predict later internalising disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 45, 874–883. <http://dx.doi.org/10.1111/j.1469-7610.2004.00279.x>
- Heiervang, E., Stormark, K. M., Lundervold, A. J., Heimann, M., Goodman, R., Posserud, M.-B., . . . Gillberg, C. (2007). Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors, and service use. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46, 438–447. <http://dx.doi.org/10.1097/chi.0b013e31803062bf>
- Keenan, K., Feng, X., Hipwell, A., & Klostermann, S. (2009). Depression begets depression: Comparing the predictive utility of depression and anxiety symptoms to later depression. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 50, 1167–1175. <http://dx.doi.org/10.1111/j.1469-7610.2009.02080.x>
- Kendall, P. C., Hudson, J. L., Gosch, E., Flannery-Schroeder, E., & Suveg, C. (2008). Cognitive-behavioral therapy for anxiety disordered youth: A randomized clinical trial evaluating child and family modalities. *Journal*

- of Consulting and Clinical Psychology, 76, 282–297. <http://dx.doi.org/10.1037/0022-006X.76.2.282>
- Kendall, P. C., Rodriguez, K. A. O. N., Villabo, M. A., Martinsen, K. D., Stark, K. D., & Banneyer, K. (2014). Cognitive-behavioral therapy with children and adolescents. In J. T. Ehrenreich & B. C. Chu (Eds.), *Transdiagnostic treatments for children and adolescents: Principles and practice* (pp. 161–182). New York, NY: Guilford Press.
- Kendall, P. C., Stark, K. D., Martinsen, K., O'Neil, K. A., & Arora, P. (2013). *EMOTION: "Coping kids" managing anxiety and depression; group leaders manual*. Ardmore, PA: Workbook Publishing.
- Kovacs, M., & Lopez-Duran, N. (2010). Prodromal symptoms and atypical affectivity as predictors of major depression in juveniles: Implications for prevention. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 51, 472–496. <http://dx.doi.org/10.1111/j.1469-7610.2010.02230.x>
- March, J. S. (1997). *MASC—Multidimensional Anxiety Scale for Children—Technical Manual*. Toronto, Canada: Multi-Health Systems.
- March, J., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J., . . . the Treatment for Adolescents With Depression Study (TADS) Team. (2004). Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *Journal of the American Medical Association*, 292, 807–820. <http://dx.doi.org/10.1001/jama.292.7.807>
- Martinsen, K. D., Kendall, P. C., Stark, K., & Neumer, S.-P. (2016). Prevention of anxiety and depression in children: Acceptability and feasibility of the transdiagnostic EMOTION program. *Cognitive and Behavioral Practice*, 23, 1–13. <http://dx.doi.org/10.1016/j.cbpra.2014.06.005>
- Martinsen, K. D., Kendall, P. C., Stark, K. D., Rodriguez, K. A. O. N., & Arora, P. (2014). *Mestrende barn, gruppeledermanual barn*. Oslo, Norway: Gyldendal Akademisk.
- Merikangas, K. R., Nakamura, E. F., & Kessler, R. C. (2009). Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neuroscience*, 11, 7–20.
- Mychailyszyn, M. P., Brodman, D. M., Read, K. L., & Kendall, P. C. (2012). Cognitive-behavioral school-based interventions for anxious and depressed youth: A meta-analysis of outcomes. *Clinical Psychology: Science and Practice*, 19, 129–153. <http://dx.doi.org/10.1111/j.1468-2850.2012.01279.x>
- Patras, J., Martinsen, K. D., Holen, S., Sund, A. M., Adolfsen, F., Rasmussen, L. P., & Neumer, S.-P. (2016). Study protocol of an RCT of EMOTION: An indicated intervention for children with symptoms of anxiety and depression. *BMC Psychology*, 4, 48. <http://dx.doi.org/10.1186/s40359-016-0155-y>
- Pine, D. S. (2007). Research review: A neuroscience framework for pediatric anxiety disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 48, 631–648. <http://dx.doi.org/10.1111/j.1469-7610.2007.01751.x>
- Stark, K. D., Streusand, W., Prerna, A., & Patel, P. (2012). Childhood depression: The ACTION Treatment Program. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (4th ed., pp. 190–233). New York, NY: Guilford Press.
- Stice, E., Shaw, H., Bohon, C., Marti, C. N., & Rohde, P. (2009). A meta-analytic review of depression prevention programs for children and adolescents: Factors that predict magnitude of intervention effects. *Journal of Consulting and Clinical Psychology*, 77, 486–503. <http://dx.doi.org/10.1037/a0015168>
- Stockings, E. A., Degenhardt, L., Dobbins, T., Lee, Y. Y., Erskine, H. E., Whiteford, H. A., & Patton, G. (2016). Preventing depression and anxiety in young people: A review of the joint efficacy of universal, selective and indicated prevention. *Psychological Medicine*, 46, 11–26. <http://dx.doi.org/10.1017/S0033291715001725>
- Teubert, D., & Pinquart, M. (2011). A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents. *Journal of Anxiety Disorders*, 25, 1046–1059. <http://dx.doi.org/10.1016/j.janxdis.2011.07.001>
- Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., & Christensen, H. (2017). School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review*, 51, 30–47. <http://dx.doi.org/10.1016/j.cpr.2016.10.005>

Received August 21, 2017

Revision received August 28, 2018

Accepted September 10, 2018 ■

Paper 2

Rasmussen, L-M. P., Patras, J., Neumer, S-P., Adolfsen, F., Martinsen, K. D., Holen, S., Sund, A. M., & Martinussen, M. (2019). Facilitators and barriers to the implementation of EMOTION: An indicated intervention for young schoolchildren. *Scandinavian Journal of Educational Research*, doi:10.1080/00313831.2019.1596976.



Facilitators and Barriers to the Implementation of EMOTION: An Indicated Intervention for Young Schoolchildren

Lene-Mari Potulski Rasmussen, Joshua Patras, Simon-Peter Neumer, Frode Adolfsen, Kristin Dagmar Martinsen, Solveig Holen, Anne Mari Sund & Monica Martinussen

To cite this article: Lene-Mari Potulski Rasmussen, Joshua Patras, Simon-Peter Neumer, Frode Adolfsen, Kristin Dagmar Martinsen, Solveig Holen, Anne Mari Sund & Monica Martinussen (2019): Facilitators and Barriers to the Implementation of EMOTION: An Indicated Intervention for Young Schoolchildren, Scandinavian Journal of Educational Research, DOI: [10.1080/00313831.2019.1596976](https://doi.org/10.1080/00313831.2019.1596976)

To link to this article: <https://doi.org/10.1080/00313831.2019.1596976>



Published online: 02 Apr 2019.



Submit your article to this journal [↗](#)






Article views: 10



View Crossmark data [↗](#)



Facilitators and Barriers to the Implementation of EMOTION: An Indicated Intervention for Young Schoolchildren

Lene-Mari Potulski Rasmussen ^a, Joshua Patras ^a, Simon-Peter Neumer^b,
Frode Adolfsen^a, Kristin Dagmar Martinsen^b, Solveig Holen^b, Anne Mari Sund^a and
Monica Martinussen ^a

^aThe Regional Centre for Child and Youth Mental Health and Child Welfare, UIT – The Arctic University of Tromsø, Tromsø, Norway; ^bCentre for Child and Adolescent Mental Health – Oslo, Norway

ABSTRACT

Research on implementation of school-based transdiagnostic interventions, conducted by personnel from the municipal services is limited. We investigated facilitators and barriers regarding implementation of EMOTION, an intervention targeting symptoms of anxiety and depression in children 8–12 years. Trained health- and childcare professionals completed one questionnaire before ($N=63$) and a separate questionnaire after running an EMOTION group ($N=66$). Twelve of the group leaders were interviewed to provide additional information regarding implementation. Results indicated that factors such as a perceived need for the intervention and positive attitudes from the group leaders facilitated implementation. Hindering factors were related to time constraints, workload, unsupportive leaders, and lack of cooperation from the schools. Allocating resources to implementation specifically could promote future use of the program.

ARTICLE HISTORY

Received 20 March 2018
Accepted 7 March 2019

KEYWORDS

School-based intervention;
implementation; emotional
problems; municipal services

Preventive interventions for children and adolescents are becoming an important part of children's mental health and municipal services (Greenberg, Domitrovich, & Bumbarger, 2001; Skogen, Smith, Aarø, Siqveland, & Øverland, 2018; Weisz, Sandler, Durlak, & Anton, 2005). Yet, children with internalizing problems are often overlooked and fail to receive adequate help from the municipal services (Chavira, Stein, Bailey, & Stein, 2004; Heiervang et al., 2007). Identifying and reaching these children is important, because of the disturbing consequences if left untreated (Kendall, Safford, Flannery-Schroeder, & Webb, 2004; Woodward & Fergusson, 2001). Using schools as a setting for delivering interventions has many advantages, such as easier identification and greater access to children with psychosocial difficulties (Ginsburg, Becker, Newman, & Nichols, 2008; Levitt, Saka, Hunter Romanelli, & Hoagwood, 2007). It also reduces barriers regarding locations and time scheduling, which could be a barrier for some to seek help (Masia-Warner, Nangle, & Hansen, 2006).

It is a challenge though, to recruit qualified personnel from the municipal services to deliver the interventions in school settings. Collaboration between the services running the intervention and the schools is required, and increases the complexity regarding implementation in an already intricate intra- and inter-organizational context (Novins, Green, Legha, & Aarons, 2013). Resources and support to participate and implement the intervention from the services, as well as fitting the intervention into the school context (e.g., scheduling conflicts, staffing, etc.) may be challenging for both the schools and the services (Lyons, Charlesworth-Attie, Vander Stoep, & McCauley, 2011).

The many barriers of implementing interventions in complex environments is constantly in conflict with the need for these interventions, and it is incumbent upon researchers to identify implementation factors within these contexts. For example, employees who work in municipal health services, but who deliver interventions in schools, create a circumstance where it is important to focus on the individual providers. They are in the front line conducting the intervention, hence, the individual- and organizational factors of the providers contribute directly to implementation outcomes. Aarons and colleagues (2011) proposed a model of implementation, the Exploration, Preparation/Adoption, Implementation and Sustainment (EPIS), which targets several important factors to consider during implementation in services for children and families. The implementation process can be explained through the model's four separate phases; exploration, preparation, active implementation and sustainment, together with relevant contextual factors (i.e., intervention characteristics, inner context and outer context) which acts differently within each phase (Aarons, Hurlburt, & Horwitz, 2011). Although the model comprises different phases with a multitude of variables, the *active implementation* phase which refers to the specific factors pertinent during actual implementation is most applicable to this study as we are focusing on relevant issues during ongoing implementation of a new intervention. Especially the inner context issues within this phase, directed at the organizational characteristics (i.e., structure, priorities and goals), and how different organizational factors, such as culture (shared beliefs and expectations) and climate (shared perceptions) affect implementation is of interest. Organizational culture and climate have large impact on the organizations (Glisson & Green, 2006), and constitutes together with readiness for change (Armenakis, Harris, & Mossholder, 1993; Glisson & James, 2002), some of the main factors within this phase.

The EPIS model also highlights the innovation-values fit and individual adopter characteristics (Aarons, Hurlburt, et al., 2011), which are relevant factor within implementation research (Durlak & DuPre, 2008; Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004). Innovation fit is defined as the organizations' and the individual's understanding of how the intervention incorporates the values, purpose, and service providers' tasks and responsibilities (Aarons, Hurlburt, et al., 2011; Durlak & DuPre, 2008). If the intervention itself is not well received among the providers, the willingness to implement is reduced. Hence, individual characteristics of the providers and their personal suitability are also necessary factors to consider in the implementation process, particularly in the active phase of implementation. Demographic variables, adaptability, beliefs, and attitudes toward interventions are all characteristics that could affect future utilization (Damschroder et al., 2009; Greenhalgh et al., 2004).

Another essential organizational factor in implementation is leadership (Harvey et al., 2011; Weiner, 2009), which Aarons and colleagues (2011) also identify. Leaders greatly influence the organizational climate and culture needed for adoption of new interventions, as well as managing the actual process (Aarons, Hurlburt, et al., 2011). Implementation of a new intervention might lead to changes in the organization where leadership becomes particularly important, as it may hinder a negative organizational climate and staff turnover (Aarons, Sommerfeld, & Willging, 2011). More research on leadership and other organizational factors related to implementation is however needed (Ogden & Fixsen, 2014).

Previous studies have sought to identify facilitators and barriers for implementation of evidence-based treatments within community settings (Ringle et al., 2015; Stein, Celedonia, Kogan, Swartz, & Frank, 2013). Within the school context, Forman and colleagues (2009) identified many factors regarding implementation of interventions (e.g., support, financial resources, training and consultation, association between intervention and school philosophy, visible outcomes and ways to address turnover), but these were merely aligned to school staff. For mental health providers working closely with schools, Lyon et al. (2011) observed that from one of the organizations, none of the employees continued using psychotherapy after initial training, indicating that the organizational climate influenced further implementation and continuation. Beidas et al. (2012) also investigated provider- and organizational factors relevant for training and implementation with school mental health

providers conducting CBT for anxious children. They did not find a relationship between organizational variables and implementation outcomes, indicating that more research is needed within the school setting on these issues (Beidas et al., 2012).

According to the literature, there are few interventions targeting symptoms of anxiety and depression at the same time in high-risk children (Werner-Seidler, Perry, Calear, Newby, & Christensen, 2017). Even fewer studies are investigating the implementation of such an intervention simultaneously, focusing particularly on the impact of the interventional and organizational factors. Hence, it is important to identify factors that may hinder or promote implementation within this specific context, which requires comprehensive personal resources, as well as cooperation between services, the different service providers and the schools involved.

The Current Study

The present study was part of a multi-site randomized controlled trial (RCT) investigating the effectiveness of the prevention program EMOTION: *Kids Coping with Anxiety and Depression* (Martinsen, Stark, Rodriguez, & Kendall, 2014). The study took place within the active implementation phase of EMOTION, in a Norwegian school setting (Patras et al., 2016). The program is a newly developed, group-based intervention, which aims to reduce symptoms of anxiety and depression in children aged 8–12 years. Throughout the 10-week intervention period, the children attended 20, one-hour sessions, twice a week during or immediately after school hours. In the sessions, the children focused on learning different coping skills, and strategies to handle sadness and/or anxiousness. The first ten sessions focused on psychoeducation, coping strategies and problem solving, while the last ten sessions focused merely on cognitive restructuring, exposure/behavioral activation and self-image. In addition, the parents were offered seven sessions, four of which were attended by the children. During the intervention, children (and parents) actively participated through games, role-play, exposure training/behavioral activation and different tasks which were meticulously chosen to enhance the knowledge and coping of anxiousness and sadness.

The primary aim of this study was to identify factors described in the active phase of the EPIS model that promote or inhibit the implementation of the EMOTION program within the group leaders' organizational context. Secondary aims were to explore predictors of group leader satisfaction with the EMOTION program, and group leader intention to continue using the EMOTION program in their practice. Qualitative data were gathered to explore the group leaders' experiences with implementing the intervention within the municipal services more deeply and elaborate on questions not captured with the questionnaires.

Method

Participants

Of the 68 group leaders trained in the intervention, 63 completed the group leader questionnaire prior to running groups (93% participation rate), and 97% ($N = 66$) completed the post-intervention questionnaire after the groups were finished. The group leaders were qualified health care and child-care professionals from different municipal services (e.g., health care services, educational and psychological services [EPS], and one regional Child and Adolescent Psychiatric Clinic). In Norway, EPS counsellors are employed in municipal services, but work closely with schools to help children and families with various difficulties (e.g., learning disabilities, psychosocial problems, etc.). The group leaders were recruited from seven municipalities within the three participating regions (North, Mid, and South East) in Norway. The total sample consisted of 94% women, with a mean age of 39.6 years ($SD = 9.7$).

A subsample of the participants ($n = 12$) were selected for in-person, qualitative interviews. Selection for the interview was conducted with a purposeful sampling (Creswell & Plano Clark, 2011) of

the total participants. The selection was stratified upon geographic location providing at least one representative from each of the seven locations. We also chose participants with different professions, educations, age, and experiences with the program. Previous experience working with children and/or manual-based interventions was also included, with the intention of gathering as rich information as possible. The informants were all women; four health care nurses, four psychologists and four educators. Eleven worked in the local municipal health services (e.g., EPC, school health services) and one was working in a child mental health outpatient clinic. All informants had conducted both children and parent groups, however experience with the program varied as some were conducting their first EMOTION groups, while others had run several groups (ranging from 1 to 4 groups, with a mean of 2.6).

Measures

The group leaders completed two different self-report questionnaires, chosen on the basis of available measures at the time guided by the implementation theory relevant for this study. In the startup-phase of a new EMOTION group, the participants completed the *group leader questionnaire*, assessing organizational and personal characteristics. This questionnaire included demographic variables, work environment, perceived need of the intervention and organizational factors relevant to implementation related to the group leaders' organization (e.g., leadership, organizational culture, autonomy, etc.). Within three weeks after the groups were finished, the group leaders completed a second survey (*post-intervention questionnaire*) with a different set of questions regarding their experience of being a group leader.

Group Leader Questionnaire

Demographics

The demographic questions consisted of 14 variables regarding gender, age, work place, municipality, profession, percentage of full-time employment (e.g., 50%, 100%, etc.), clinical or other specialties, and work experience. Experience was reported both in terms of number of years, and former experience working with anxiety and depression, treatment methods (e.g., cognitive behavioral therapy), and use of manuals (e.g., Coping Cat; Kendall & Hedtke, 2006) as part of a treatment method.

Work Environment and Intervention Fit

Ten questions regarding the work environment of the group leaders' and intervention fit, were developed for this study to address issues within this specific context. The items were rated on a 7-point Likert scale ranging from 1 (Strongly disagree) to 7 (Strongly agree). Based on the results of an exploratory factor analysis (EFA), the items indicated three subscales, which were labelled "Innovation fit" ($\alpha = .95$), "Organizational Support" ($\alpha = .71$), and "Attitudes towards Evidence Based Programs" ($\alpha = .82$). None of the items were dropped, but one item ("I have such large workload that it will be difficult to find time to run the EMOTION program"), did not fit into any of the scales, and was reported separately.

Organizational Readiness for Change

Inspired by the Organizational Readiness for Change (ORC; Lehman, Greener, & Simpson, 2002), a subset of items from this scale was used (the Employee Problem Scale), but re-phrased to be more relevant to the present study. For instance, the subject of the questions was changed from second person (you) to first person (I) to match the wording of the other questions in the study. Further, some of the subscales were modified. Four items were added to the efficacy scale to focus more on work-related abilities (e.g., "I manage to do positive changes through my work"), and two items that did not fit in this context were removed (e.g., "You have the skills needed to conduct

individual counseling”). One question was added to the adaptability scale to assess overall flexibility of the workplace, and one question was added to the autonomy scale to address the freedom to organize work priorities. Lastly, two items from the program goals scale were removed because they did not fit the present context (e.g., “Management here has a clear plan for this program”). This yielded us with a measure of 32 questions with six subscales; Adaptability (five items, $\alpha = .62$), Program goals (three items, $\alpha = .59$), Cohesion (six items, $\alpha = .77$), Efficacy (seven items, $\alpha = .80$), Autonomy (six items, $\alpha = .66$) and Communication (five items, $\alpha = .80$). Reliability analyses showed that the items added for this study maintained or increased reliability in terms of Cronbach’s alpha. Inter-item correlations were sufficiently large (Piedmont, 2014), indicating a fair degree of correspondence between the items. All items were rated on a 7-point Likert scale from 1 (Strongly disagree) to 7 (Strongly agree).

Readiness for Organizational Learning and Evaluation Instrument

Also, based on the Readiness for Organizational Learning and Evaluation Instrument (ROLE; Pre-skill & Torres, 1999) subsets of the questionnaire regarding work culture (e.g., Employees respect each other’s perspectives and opinions), and leadership (e.g., Managers and supervisors set realistic obligations for employees (e.g., time, resources, workload)) were included in the overall survey. Seventeen questions were included in the scale called Work culture ($\alpha = .91$) all rated on a Likert-scale from 1 (Strongly disagree) to 5 (Strongly agree). The Leadership scale ($\alpha = .89$) was also rated from 1–5 and included nine questions.

Post-intervention Questionnaire

After the groups were completed, the group leaders reported how many groups and sessions (out of 27) they had led. The group leaders also completed questions regarding the satisfaction with being a group leader in EMOTION and the supervision they received. This was rated on a scale from 1 (Very dissatisfied) to 5 (Very satisfied). The last question regarding the intention to continue with the program was rated from 1 (Very unlikely) to 5 (Very likely).

Qualitative Interview

The semi-structured interview guide was primarily based on Aarons and colleagues’ (2011) conceptual model of implementation, which led to the major topics (i.e., the intervention itself, organizational setting in relation to implementation of EMOTION, and demographic variables such as experience) with accompanied questions relevant for this study. The interviews focused on the group leaders’ organizational settings, particularly aimed towards the strengths and weaknesses regarding the implementation of EMOTION. It also focused on strengths and weaknesses concerning the program, as well as specific questions around feasibility and potential threats to a further use of the program. The semi-structured form however, allowed for flexibility to elaborate on relevant issues which emerged during the interviews.

Procedure

The intervention was mostly run as part of the group leaders’ regular practice, and delivered at schools. Schools were recruited via key-personnel (i.e., principals at schools) through general information assemblies where those interested signed a contract to participate in the study. Beyond being the context of delivery, the schools had limited liability to the completion of groups. The active phase of the project period lasted from spring 2014 until spring 2016, with a new group running at the participating schools each semester (i.e., up to two groups per school per year). Data were collected electronically by using the Confrmfit software system. The Regional Committee for Health and Medical Research Ethics (2013/1909/REK South-East) approved the study.

Interviews were conducted in person (by the first author) and took place at the group leaders' workplace or other suitable settings (e.g., a nearby café). The interviews were audio taped and lasted approximately 1-1.5 h. Data collection lasted from August 2015 to February 2016. Due to practical reasons, two informants were present at the same time in one of the interviews. A total number of eleven interviews were carried out, and a verbatim transcription then followed.

Implementation of EMOTION

Implementation of the EMOTION program being developed and investigated by the research staff could be seen as an active top-down implementation strategy (Ogden & Fixsen, 2014), supported by the following activities.

Recruitment

Professionals from different municipal and regional health services were recruited as group leaders mainly through meetings with leaders of the respective services and/or leaders of the local municipalities.

Training and Supervision

The group leaders received a three-day training in the intervention; one day with general introduction in cognitive behavioral therapy (CBT), and a two-day workshop going through the program session by session. Training was conducted by two of the research staff members (program developer and project manager). To avoid bias, all sites received training from both trainers. The group leaders were also offered a one-day booster session after most semesters, to discuss some of the challenges met during the execution of the intervention.

Supervision of EMOTION groups was conducted by certified CBT supervisors. The supervisors met with the group leaders one session prior to startup, and then every week during the ten-week program period (two on-site meetings, the remainder via Skype/telephone or face-to-face meetings). Additionally, the supervisors had regular Skype-meetings with the trainers to discuss important issues during the intervention period, and to secure a similar execution of the intervention across sites.

Quality Assurance

During the study, video tapes of 17% of the total number of sessions were obtained to ensure fidelity to the program. By using the Competence and Adherence Scale for Cognitive Behavioral Therapy (Bjaastad et al., 2016) the total adherence to the program (rated from 0 = *None* to 6 = *Thorough*) was $M = 3.53$ ($SD = 1.25$), and the mean competence score (rated from 0 = *Poor skills* to 6 = *Excellent skills*) was $M = 3.59$ ($SD = 1.26$). This indicated that the group leaders followed the manual to a large extent, and carried out the program with good skills.

Data Analyses

Quantitative Surveys

Data were analyzed using IBM Statistical Package for the Social Sciences (SPSS 24.0). Descriptive data was mainly the focus of the quantitative analyses. We also investigated the association between background variables and organizational factors to examine whether they were related to the group leaders' intention to continue with the program and satisfaction with being a group leader using Pearson's r (two-tailed). To simplify interpretation and give a meaningful impression of the group leaders' ratings of organizational characteristics (since there are no other studies to compare the results with), the three highest categories for the seven-point scales (5 [*slightly agree*], 6 [*agree*], and 7 [*strongly agree*]), and the two highest categories in the five-point scales (4 [*agree*] and 5 [*strongly agree*]) were grouped together and reported as percentages of the mean scale scores.

Qualitative Interview

The analysis for this study is mainly a descriptive presentation of the informants viewpoints, and was conducted as follows; The transcriptions were entered into the NVivo data management software program (QSR International, Cambridge, MA, USA), and analyzed using the analytic framework as described by Lacey and Luff (2001). This thematic analysis is inductive and tends to be theoretical driven, which is often applied in health research to gain specific information on a topic. The analysis consists of five key stages; the first stage, *familiarization*, took place during transcriptions and initial reading of the interviews, which generated ideas and preliminary codes linked to the data. The second stage was *to identify a thematic framework*. The theoretical frame in this study was based mostly upon the active phase within Aarons and colleagues (2011) conceptual model of implementation. In this stage, Nvivo was used to systematically work through the entire data set and generate codes and first-impression themes. Coding helped to develop a systematic overview of both pre-existing questions (i.e., strengths and weaknesses with the intervention), and newly emerging issues from the previous stage (i.e., collaboration, group process). Then, in stage three, *indexing* the data to the theoretical framework was initiated, searching for themes. Examples of themes generated from the data were work-related issues, benefits with the intervention and school participation. The fourth stage involved *charting* the material to create an overview and organizing the coded data into the different themes. This provided the opportunity to recode some of the information and grouping the data systematically into themes such as organizational factors, interventional aspects and school investment. Lastly, *interpretation of the material* took place, searching for patterns and associations relevant to the theoretical framework and the main aims of this study. During this phase, the first author finalized the thematic structure. The results were validated by a coauthor with qualitative experience, who read and discussed the data until agreement was reached. The remaining co-authors participated in the analytic process through reading and commenting on the qualitative material, by evaluating the clarity and relevance of the coded categories.

Results

Quantitative Surveys

The group leaders' professional background were mainly psychologists/specialists (35%), educational-psychological counsellors (18%) and health nurses (14%) primarily working in the municipal mental health services closely connected to the schools. The group leaders' had an average of eight years' experience in the field ($SD = 6.5$), and almost 70% had experience working with anxious and/or depressed children. See [Table 1](#) for an overview of the different background variables.

The results for the mean subscale percentage agreement were 70% or more for all the organizational subscales, except communication, leadership and work culture. The highest endorsed subscale was the efficacy scale reflecting the group leaders' self-efficacy regarding work, where 95% of the participants slightly agreed, agreed or strongly agreed with the statements. Similar results were obtained for the innovation fit scale (90%), which explored the group leaders' opinions of whether such an intervention was needed in their organizational setting. The leadership scale reflects how the group leaders perceived the general leadership and leadership support within their organization. This scale, together with the work culture scale indicating how the participants perceive the overall culture in the organization, received the lowest endorsements. Less than 30% of the group leaders agreed or strongly agreed with the statements. An overview of the percentages endorsing the highest response categories on the different mean subscales, are presented in [Table 2](#).

According to the post-group survey, the group leaders ($N = 66$) received an average of 7.08 ($SD = 2.53$) hours of supervision, which 62% reported to be satisfied or very satisfied with ($M = 3.92$, $SD = 0.92$). On the question of whether they would continue with the program in the future, approximately 53% indicated that they would likely or most likely continue ($M = 3.44$, $SD = 0.96$). The

Table 1. Background variables.

	<i>M (SD)</i>	%
Gender		
Men		6
Women		94
Age	39.6 (9.7)	
Profession		
Health nurses		14
Psychologists/Specialists		35
Educational and psychological counsellor (EPC)		18
Psychology student		5
Educator/special educator		11
Child-care worker		6
Occupational therapist		3
Other (e.g., counsellors, family therapists, lecturer)		8
Region		
South East		36
North		19
Mid		44
Experience		
Years in the field	7.6 (6.5)	
Clinical specialty		14
With anxiety and/or depression		68
With other treatment methods		
CBT		38
Coping cat		18
Other methods		21
Manual-based treatments		41

Note: $N = 63$.

correlation between satisfaction with the program and intention to continue was moderate, $r = .42$, $p < .01$. There were small, but significant associations between intention to continue and the organizational factor innovation fit $r = .26$, $p < .05$, as well as efficacy $r = .26$, $p < .05$. Autonomy showed a weak, negative correlation with satisfaction with being a group leader $r = -.28$, $p < .05$. Heavy workload was negatively associated with both satisfaction ($r = -.36$, $p < .01$) and intention to continue ($r = -.29$, $p < .05$). Further analyses showed no significant associations between the demographic variables and satisfaction with being a group leader nor intention to continue. A list of the correlations is presented in Table 3.

Table 2. Descriptive statistics for the different subscales in the Group Leader Questionnaire.

Instrument	Subscale/item	No. of items	<i>M</i>	<i>SD</i>	% positive endorsements ^c
Work environment/ intervention fit	Innovation fit	2	5.80 ^a	0.80	90
	Org. support	3	5.49 ^a	1.11	71
	Attitudes towards EBP	4	5.68 ^a	0.98	81
	Workload	1	4.95 ^a	1.26	73
ORC	Adaptability	5	5.27 ^a	0.60	70
	Program goals	3	5.32 ^a	0.92	71
	Cohesion	6	5.35 ^a	0.77	70
	Efficacy	7	5.69 ^a	0.47	95
	Autonomy	6	5.22 ^a	0.72	70
	Communication	5	4.83 ^a	0.83	54
ROLE	Leadership	9	3.48 ^b	0.67	27
	Work culture ^d	17	3.73 ^b	0.51	30

Note: $N = 63$.

^aRated on a scale from 1 (Strongly disagree) to 7 (Strongly agree).

^bRated on a scale from 1 (Strongly disagree) to 5 (Strongly agree).

^cMean scale scores percentages for the highest categories in the subscales.

^d $N = 51$ for the Culture scale.

Org. support = organizational support. EBP = Evidence Based Programs. ORC = Organizational Readiness for Change (Lehman et al., 2002). ROLE = Readiness for Organizational Learning and Evaluation Instrument (Preskill & Torres, 1999).

Table 3. Correlations between satisfaction and intention to continue with demographic and organizational variables.

	N	Satisfaction	Intention to continue
Demographic variables			
Clinical specialty	62	-.15	-.08
Exp. within the field/area	62	.03	.09
Exp. with depression and anxiety	61	-.12	.17
Exp. with CBT	42	-.05	-.03
Exp. coping cat	42	.26	.11
Exp. other treatment methods	42	.17	.11
Exp. manuals	62	-.11	.03
Organizational variables			
Innovation fit	62	.22	.26*
Organizational support	62	.02	.12
Attitudes towards EBP	62	-.14	.06
Workload	62	-.36**	-.29*
Adaptability	60	-.00	.05
Program goals	60	.25	.20
Cohesion	60	-.06	.08
Efficacy	60	.26*	.19
Autonomy	60	-.28*	-.25*
Communication	60	-.07	-.01
Leadership	60	.07	.01
Work culture	50	.01	-.03

* $p < .05$, ** $p < .01$ (two-tailed).

Exp. = Experience, CBT = Cognitive Behavioral Therapy, EBP = Evidence Based Programs.

Qualitative Interview

Results from the qualitative analysis show that all three main findings, *organizational factors*, *interventional aspects* and *school investment* include both facilitators and barriers for implementing the EMOTION program. There was a high concordance between the group leaders' responses, and quotations from different individuals have been chosen to illustrate the results.

Organizational Factors

Our first main finding showed that organizational factors promoting and inhibiting implementation was closely related to where the informants worked. Generally, they considered their workplace as a suitable setting for implementing EMOTION, particularly group leaders in the educational and psychological counsellor services (EPC). Reasons expressed were the thematic relevance of their work with the schools, the significance of the work tasks they were assigned by the Directorate for Education, and the need for helpful tools to respond to these demands. They also highlighted the opportunity to reach the children at an early stage because of the closeness to the schools, and potentially preventing larger difficulties to evolve.

The advantage with working, or to implement EMOTION in EPC is that we work closely with the schools and we know the schools pretty well, including the people working there ... and we know their challenges.

The informants had different types of employment in the organizations. Some were counsellors and health nurses, and others worked as mental health professionals in municipal services where children were referred to care (e.g., family centers, municipal psychology teams, etc.). The mental health professionals reported being able to adjust the time schedule instead of adding the EMOTION groups on top of the everyday work tasks, and hence had a larger capacity to run groups.

It will probably be room for it, if I say I want to work with this, I would be allowed – definitely. And my leader is very supportive to the whole project, and yes ... it was kind of her who brought it up, and gave us the opportunity to work with it if we wanted to ...

However, the informants also expressed some barriers regarding implementation of EMOTION in the municipal context. The main issues were the amount of mandatory work in the normal course of

their jobs, lack of support from the leadership regarding the intervention, limited time and resources, and issues reflecting the structure of the services and practitioners' main tasks (e.g., focusing on educational vs. mental health job tasks).

Recurring issues such as lack of leadership support and time and resources were particularly relevant. The informants who experienced a negative leader emphasized the lack of support as a major threat for the implementation of the program.

I feel that it stops when it comes to resources! [...] Our leader closes his eyes and ears to what I'm doing [...] and then I feel that it was my choice to participate, and I can't complain [...] but if he had been more positive, then I might get some help ...

Further, most of the informants did express issues regarding how this affected their everyday work situation. The greatest concern revolved around how to fit EMOTION into their work schedule. Everyone stated it was time consuming and for many of the group leaders it was challenging running groups in addition to regular mandatory work tasks.

We are doing this on top of everything else; no one is taking away the other work tasks, rather the opposite that we are getting more. It becomes very work demanding and intense periods [...]

Interventional Aspects

The second main finding was related to the intervention and aspects involving the EMOTION program. Every informant agreed that there was a need for an intervention targeting children with these difficulties, and accentuated elements with the program that was important and helpful for the children.

Yes, these children are everywhere ... [...] So, I think it is a need, and I think it is very important that we set in motion these kinds of preventive interventions for these children ... [...] In the end [if left untreated], they are not in on anything, not school, not work, nothing.

Additionally, the group leaders highlighted the usefulness and the learning outcomes of the intervention, which they could use in other areas at work. Overall, they likewise emphasized the group process as a major benefit of the program and how much fun it was to conduct the groups. However, regarding the less positive aspects, group leaders' also stated that the extent of the program was a challenge and that the manuals need some adjustments and minor revisions. The majority expressed a necessity to reduce the number of sessions, moderate some of the text and introduce more colors and age appropriate tasks.

The manual in relation to all the material you were supposed to cover in one session – it was a lot of text in the manual. A little bit like ... are you sure ... is it expected that you should say all of this? [...] You read it, and you try to communicate in a way the most important topics, but it is ... it was impossible to us at least.

School Investment

Our last main finding was particularly derived from this study context. For many of the informants, the completion of the groups was highly dependent on the schools' involvement and participation. The groups were conducted during regular school hours, but many of the informants worked in other offices outside the school premises. If the principal and the teachers were positive, practical issues, such as scheduling groups, became much easier for the group leaders. Also, the group leaders stated that involvement from the schools made a difference in how the parents and children understood and interpreted the information forwarded to them.

Now I am at a school I know pretty well, so the teachers I work together with this year I know, and they are positive and then it works very well! The first round it was not okay, they were ... some of the teachers didn't think much about [EMOTION] [...] and then it becomes hard!

Discussion

The main goal of this study was to identify facilitators and barriers regarding the implementation of an indicated transdiagnostic intervention, the EMOTION program, being delivered in schools by group leaders' from mental health and municipal services. Results showed that multiple factors within the active phase of implementation could influence the providers' likelihood to continue with the program. Especially factors related to the intervention and the implementing organization, as well as factors linked to the context of delivery (i.e., schools) were important issues derived from this study. There was an overall positive attitude towards the program and an obvious need for a program targeting anxiety and depressive symptoms in schoolchildren. However, time pressure and heavy workload seem to greatly affect further implementation. As the results from this study indicate, without a supportive leadership implementation becomes very difficult. Also, collaboration between the services and the context of delivery – in this case the schools – are highly important.

In this study, the perceived fit of the innovation within the organization were associated with the desire to continue with the program. This is an important component to promote implementation. Similar results have been reported previously, indicating that satisfaction with an intervention were one of the most important predictors among school psychologists' intention to continue the implementation (Forman, Fagley, Chu, & Walkup, 2012). Efficacy was also rated high, indicating that the group leaders considered themselves proficient within their work and having the skills needed to conduct EMOTION groups. Self-efficacy is an important characteristic when implementing new interventions as it also reflects the group leaders' perceived ability to adapt to change (Jim-mieson, Terry, & Callan, 2004). Results from the interviews supported this notion, as the group leaders stated that their service was suitable for running EMOTION groups and that they found themselves in a good position to work with issues related to anxiety and depression. The perceived need for an intervention targeting emotional problems was also evident, and the group leaders found the skills learned to be useful in other areas of their work. The group leaders indicated, however, that to strengthen an overall implementation of the program, some adjustments of the manual would be beneficial.

Previous research highlights the importance of the organizational factors during program adoption and sustainability (e.g., Aarons, Hurlburt, et al., 2011; Durlak & DuPre, 2008; Glisson et al., 2012). Although small, we did find a negative correlation between work load and satisfaction, as well as intention to continue. This suggests that time pressure may threaten the continued use of the program, especially when we consider the intensity of the program compared to similar interventions (Reynolds, Wilson, Austin, & Hooper, 2012). Hence, one of the main characteristics of the program, its intensity, could be a barrier for the implementation in municipal and school health services. This is a typical dilemma in these services, where you want to enhance the intensity of the treatment, but still keep the intervention within manageable limits (i.e., not conducting therapeutic counselling). Future implementation studies should address this effort by investigating how to increase the efficiency without exceeding time constraints or limits at work. Use of combined internet interventions or as part of the regular school curriculum could be adaptations to be explored.

Reinforcing concerns related to time and workload, nearly half of the participants in this study were uncertain if they would continue as group leaders after the project period had ended. Further, a majority (73%) of the group leaders indicated that they had such a heavy workload that it would be difficult to find time to run EMOTION. This is a substantial threat to further implementation of the EMOTION program. An explanation of this result could be limited time and resources allocated to run the intervention, which were recurring themes among the group leaders during the interviews. Unfortunately, this is not an issue uniquely derived from this study, as this has been addressed in several studies recently (Beidas et al., 2016; Bond et al., 2014). To diminish this barrier, reducing other work tasks or hiring enough people to carry out the intervention is required. Furthermore, as previous research has shown, resources and funding to the schools are also important aspects

to promote successful implementation and long-term sustainability of interventions with in the school context (Eriksen, Hegna, Bakken, & Lyng, 2014).

Similar issues were reflected upon regarding autonomy at work. In fact, the quantitative results indicated that the participants with more autonomy were less satisfied with the program. Findings from the interviews, however, indicated that the group leaders who experienced a possibility to set aside time to do the intervention also spoke more positively of continuing with the program. The group leaders who had to conduct EMOTION groups on top of their normal workload were also more reluctant to continue. Previous findings support the results from the interviews, suggesting that individuals who have the autonomy to create their own schedule and regulate work tasks also feel less overwhelmed and fatigued by the work they do (Hornung & Rouseau, 2007; Ringle et al., 2015). This inconsistency could be the result of unsupportive and disengaged leaders, who apply too much pressure on the group leaders to accomplish a rather comprehensive program in addition to their other work tasks. Hence, a high degree of autonomy is positive for arranging the work tasks and working schedule, however, it may also lead to negative feelings overload related to their role at work, thereby decreasing the satisfaction with the program. Generally, this implies that for further implementation of EMOTION, autonomy should be given attention, as it seems to affect the practitioners' attitudes towards the program, and impact future use.

In addition, the level of support the practitioners received from the leaders was important for implementing the EMOTION program. In general, the group leaders rated leadership support moderately and much lower than other constructs in the survey. In the interviews, the importance of leadership support became clearer in the sense that the program would not continue without managers' intention nor acceptance to proceed. There should be a particular focus on training or informing leaders in relevant issues within the implementation process (e.g., how to lead through change, be aware of potential threats), so that EMOTION, and similar interventions, receives adequate support. Future research could also focus on the higher management levels and investigate whether making more structural changes within the organizational systems (e.g., dedicating employees' time to do this intervention as part of their regular job) would enhance implementation.

The group leaders emphasized the schools level of involvement as an important facilitator for implementation of the EMOTION program into school settings. Children spend many hours at school, and some of the issues anxious/depressed youth come across are present during school hours (e.g., reading or talking aloud in the classroom). Also, the practical matters would be more easily organized if the collaboration between municipal health care services and the schools were enhanced. Domitrovich et al. (2008) have highlighted the importance of understanding the school's role and influence when implementing interventions. The lack of a natural meeting point in the Norwegian context between the mental health providers and school officials, except for when individual cases are referred to further assessment, serves as a potential barrier for implementation of EMOTION and other school interventions. Thus, increased emphasis on how to bridge the gap between the municipal services and schools could possibly result in a better collaboration. Further, a better collaboration could again promote implementation of effective interventions.

Furthermore, in this study, group leaders' reported that it was easier to run the groups at schools where the teachers and other school staff were better informed about the program. The schools volunteered to participate, but did not have an active role besides providing locations and conducting the surveys with the children. One way to deal with this in the future could be to have the group leaders present the program to school staff thoroughly and discuss with the teachers how to conduct the intervention most adequately within in the school context. This could have the effect of supporting the program in the classroom and raise awareness within the school, as well as reinforcing the collaboration between schools and the municipal services. In the future, gathering information from the schools (e.g., school leadership, teachers) on how to implement EMOTION as best as possible, should be undertaken as this could strengthen future use of the program. Also, testing different implementation strategies (e.g., applying additional training, supervision, feedback systems; Proctor,

Powell, & McMillen, 2013), and investigating specific implementation outcomes to overcome some of the barriers extracted from this study should be executed.

Study Limitations

Despite the number of practical implications from this study which are applicable for the further development of the intervention in the next phase, there are some limitations. First, the number of respondents are relatively small. Although including implementation research within the context of effectiveness studies is valuable (Proctor, 2009), this presented us with a challenge regarding sample size of group leaders. Future efforts should be made, to design and conduct implementation research to comply with these issues.

There was little variation within the response categories for some of the questions as indicated by the high mean scores and small standard deviations. One possible explanation could be that the group leaders mainly volunteered to participate, which indicate that they were interested in sad and anxious children and at least open to working with manual-based interventions. Some of the measurement scales also had marginal reliability. However, the lack of variation may be normal for these scales, which introduces other issues: lack of normative data and psychometric analyses of the measures used in particular and implementation measures in general (Lewis et al., 2015). Having reliable and valid measures should potentially reduce the need to develop self-constructed questions and questionnaires, like the adaptations to the ORC, which is an additional limitation in this study. This could further advance the statistical analyses, addressing issues such as mass significance, which could imply that the correlations derived from this study must be treated with some caution. Also, due to a technical error, the Culture subscale was omitted from the survey during data collection for two semesters. This caused the number of respondents on that scale to decrease. Finally, during one of the interviews, two of the participants were present at the same time, which could have influenced their responses. Only having one person code the qualitative material is an additional shortcoming of the study.

Conclusions

The current study provides important aspects regarding facilitators and barriers of implementing new interventions targeting anxiety and depression in the municipal services and bringing it into school settings. This challenging maneuver requires organizational leaders, group leaders, supervisors, and school personnel working together to adapt and implement a program.

The group leaders' clearly indicated a need for a program such as EMOTION targeting children at risk for developing anxiety and/or depression. They also found the municipal services as a suitable setting for implementing the intervention, and schools an appropriate arena to reach the children. However, the main barriers were associated with time constraints, lack of resources and capacity to run groups. Supportive leaders are also important for further use of the program, as well as a close cooperation with the schools involved. For EMOTION, revising the manual to fit it more adequately within the municipal service context, setting aside time and allocating resources in the services or reduce other work tasks, as well as establishing a cooperation with the schools involved will be important. To promote implementation of interventions delivered in schools in the future, the mentioned issues needs to be addressed and incorporated into a strategic implementation plan guided by the results from this study.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study. Parents consented on behalf of their children.

Disclosure Statement

Kristin D. Martinsen receives royalties from sale of the Norwegian version of the EMOTION program manuals and workbook. Anne Mari Sund has received travel support from Medice, a company which provides medication for ADHD treatment.

Funding

This work was supported by the Norwegian Research Council (Norges Forskningsråd) under Grant 228846/H10.

ORCID

Lene-Mari Potulski Rasmussen  <http://orcid.org/0000-0002-7092-2495>

Joshua Patras  <http://orcid.org/0000-0001-5233-6188>

Monica Martinussen  <http://orcid.org/0000-0002-0937-1479>

References

- Aarons, G. A., Hurlburt, M., & Horwitz, S. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*, 38, 4–23. doi:10.1007/s10488-010-0327-7
- Aarons, G. A., Sommerfeld, D. H., & Willging, C. E. (2011). The soft underbelly of system change: The role of leadership and organizational climate in turnover during statewide behavioral health reform. *Psychological Services*, 8, 269–281. doi:10.1037/a0026196
- Armenakis, A. A., Harris, S. G., & Mossholder, K. W. (1993). Creating readiness for organizational change. *Human Relations*, 46, 681–703. doi:10.1177/001872679304600601
- Beidas, R. S., Mychailyszyn, M. P., Edmunds, J. M., Khanna, M. S., Downey, M. M., & Kendall, P. C. (2012). Training school mental health providers to deliver cognitive-behavioral therapy. *School Mental Health*, 4, 197–206. doi:10.1007/s12310-012-9074-0
- Beidas, R. S., Stewart, R. E., Adams, D. R., Fernandez, T., Lustbader, S., Powell, B. J., & Barg, F. K. (2016). A multi-level examination of stakeholder perspectives of implementation of evidence-based practices in a large urban publicly-funded mental health system. *Administration and Policy in Mental Health and Mental Health Services Research*, 43, 893–908. doi:10.1007/s10488-015-0705-2
- Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W., Torsheim, T., Havik, O. E., Heiervang, E. R., & Öst, L.-G. (2016). Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. *Psychological Assessment*, 28, 908–916. doi:10.1037/pas0000230
- Bond, G. R., Drake, R. E., McHugo, G. J., Peterson, A. E., Jones, A. M., & Williams, J. (2014). Long-term sustainability of evidence-based practices in community mental health agencies. *Administration and Policy in Mental Health and Mental Health Services Research*, 41, 228–236. doi:10.1007/s10488-012-0461-5
- Chavira, D. A., Stein, M. B., Bailey, K., & Stein, M. T. (2004). Child anxiety in primary care: Prevalent but untreated. *Depression and Anxiety*, 20, 155–164. doi:10.1002/da.20039
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.) Thousand Oaks, CA: Sage.
- Damschroder, L., Aron, D., Keith, R., Kirsh, S., Alexander, J., & Lowery, J. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, 4(50). doi:10.1186/1748-5908-4-50
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., ... Jalongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, 1, 6–28. doi:10.1080/1754730X.2008.9715730
- Durlak, J. A., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–350. doi:10.1007/s10464-008-9165-0
- Eriksen, I. M., Hegna, K., Bakken, A., & Lyng, S. T. (2014). *The psychosocial school environment in Norwegian primary and secondary education [Felles fokus. En studie av skolemiljøprogrammer i norsk skole]* (NOVA Report 15/14). Oslo and Akershus University College of Applied Sciences (HiOA) Centre for Welfare and Labour Research (NOVA).
- Forman, S. G., Fagley, N. S., Chu, B. C., & Walkup, J. T. (2012). Factors influencing school psychologists' "willingness to implement" evidence-based interventions. *School Mental Health*, 4, 207–218. doi:10.1007/s12310-012-9083-z

- Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health*, 4, 26–36. doi:10.1007/s12310-008-9002-5
- Ginsburg, G. S., Becker, K. D., Newman, J., & Nichols, T. (2008). Transporting CBT for childhood anxiety disorders into inner-city school-based mental health clinics. *Cognitive and Behavioral Practice*, 15, 148–158.
- Glisson, C., & Green, P. (2006). The effects of organizational culture and climate on the access to mental health care in child welfare and juvenile justice systems. *Administration and Policy in Mental Health and Mental Health Services Research*, 33, 433–448. doi:10.1007/s10488-005-0016-0
- Glisson, C., Hemmelgarn, A., Green, P., Dukes, D., Atkinson, S., & Williams, N. J. (2012). Randomized trial of the ARC organization intervention with community-based mental health programs and clinicians serving youth. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51, 780–787. doi:10.1016/j.jaac.2012.05.010
- Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior*, 23, 767–794. doi:10.1002/job.162
- Greenberg, M., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention and Treatment*, 4(1), 1–62.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly*, 82, 581–629.
- Harvey, G., Fitzgerald, L., Fielden, S., McBride, A., Waterman, H., Bamford, D., & Boaden, R. (2011). The NIHR collaboration for leadership in applied health research and care (CLAHRC) for greater Manchester: Combining empirical, theoretical and experiential evidence to design and evaluate a large-scale implementation strategy. *Implementation Science*, 6, doi:10.1186/1748-5908-6-96
- Heiervang, E., Stormark, K. M., Lundervold, A. J., Heimann, M., Goodman, R., Posserud, M.-B., & Gillberg, C. (2007). Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors, and service use. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46, 438–447. doi:10.1097/chi.0b013e31803062bf
- Hornung, S., & Rousseau, D. M. (2007). Active on the job – proactive in change: How autonomy at work contributes to employee support for organizational change. *The Journal of Applied Behavioral Science*, 43, 401–426. doi:10.1177/0021886307307555
- Jimmieson, N. L., Terry, D. J., & Callan, V. J. (2004). A longitudinal study of employee adaptation to organizational change: The role of change-related information and change-related self-efficacy. *Journal of Occupational Health Psychology*, 9, 11–27. doi:10.1037/1076-8998.9.1.11
- Kendall, P. C., & Hedtke, K. (2006). *Cognitive-behavioral therapy for anxious children: Therapist manual* (3rd ed.). Ardmore, PA: Workbook publishing.
- Kendall, P. C., Safford, S., Flannery-Schroeder, E., & Webb, A. (2004). Child anxiety treatment: Outcomes in adolescence and impact on substance use and depression at 7.4-year follow-up. *Journal of Consulting and Clinical Psychology*, 72, 276–287. doi:10.1037/0022-006X.72.2.276
- Lacey, A., & Luff, D. (2001). Qualitative data analysis. *Qualitative Research*, 1, 303–323.
- Lehman, W. E. K., Greener, J. M., & Simpson, D. D. (2002). Assessing organizational readiness for change. *Journal of Substance Abuse Treatment*, 22, 197–209. doi:10.1016/S0740-5472(02)00233-7
- Levitt, J. M., Saka, N., Hunter Romanelli, L., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, 45, 163–191. doi:10.1016/j.jsp.2006.11.005
- Lewis, C. C., Stanick, C. F., Martinez, R. G., Weiner, B. J., Kim, M., Barwick, M., & Comtois, K. A. (2015). The society for implementation research collaboration instrument review project: A methodology to promote rigorous evaluation. *Implementation Science*, 10(2), doi:10.1186/s13012-014-0193-x
- Lyon, A. R., Charlesworth-Attie, S., Vander Stoep, A., & McCauley, E. (2011). Modular psychotherapy for youth with internalizing problems: Implementation with therapists in school-based health centers. *School Psychology Review*, 40, 569–581.
- Martinsen, K. D., Stark, K., Rodriguez, K. O., & Kendall, P. C. (2014). *Mestrende barn manual*. Oslo: Gyldendal Norsk Forlag.
- Masia-Warner, C., Nangle, D. W., & Hansen, D. J. (2006). Bringing evidence-based child mental health services to the schools: General issues and specific populations. *Education & Treatment of Children*, 29, 165–172.
- Novins, D. K., Green, A. E., Legha, R. K., & Aarons, G. A. (2013). Dissemination and implementation of evidence-based practices for child and adolescent mental health: A systematic review. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52, 1009–1025.e1018. doi:10.1016/j.jaac.2013.07.012.
- Ogden, T., & Fixsen, D. L. (2014). Implementation science: A brief overview and a look ahead. *Zeitschrift für Psychologie*, 222, 4–11. doi:10.1027/2151-2604/a000160
- Patras, J., Martinsen, K. D., Hølen, S., Sund, A. M., Adolfsen, F., Rasmussen, L.-M. P., & Neumer, S.-P. (2016). Study protocol of an RCT of EMOTION: An indicated intervention for children with symptoms of anxiety and depression. *BMC Psychology*, 4(48), doi:10.1186/s40359-016-0155-y
- Piedmont, R. L. (2014). Inter-item correlations. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and well-being research* (pp. 3303–3304). Dordrecht: Springer Netherlands.

- Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.
- Proctor, E. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research*, 36, 24–34.
- Proctor, E., Powell, B. J., & McMillen, J. C. (2013). Implementation strategies: Recommendations for specifying and reporting. *Implementation Science*, 8(139), doi:10.1186/1748-5908-8-139
- Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, 32, 251–262. doi:10.1016/j.cpr.2012.01.005
- Ringle, V. A., Read, K. L., Edmunds, J. M., Brodman, D. M., Kendall, P. C., Barg, F., & Beidas, R. S. (2015). Barriers to and facilitators in the implementation of cognitive-behavioral therapy for youth anxiety in the community. *Psychiatric Services*, 66, 938–945. doi:10.1176/appi.ps.201400134
- Skogen, J. C., Smith, O. R. F., Aarø, L. E., Siqveland, J., & Øverland, S. (2018). *Mental health among children and adolescents. Health-promoting and preventive public health interventions. A summary of evidence about effects. [Barn og unges psykiske helse: Forebyggende og helsefremmende folkehelseiltak. En kunnskapsoversikt]*. Retrieved from Oslo.
- Stein, B. D., Celedonia, K. L., Kogan, J. N., Swartz, H. A., & Frank, E. (2013). Facilitators and barriers associated with implementation of evidence-based psychotherapy in community settings. *Psychiatric Services*, 64, 1263–1266. doi:10.1176/appi.ps.201200508
- Weiner, B. J. (2009). A theory of organizational readiness for change. *Implementation Science*, 4, 67–67. doi:10.1186/1748-5908-4-67
- Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist*, 60, 628–648. doi:10.1037/0003-066X.60.6.628
- Werner-Seidler, A., Perry, Y., Calear, A. L., Newby, J. M., & Christensen, H. (2017). School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review*, 51, 30–47. doi:10.1016/j.cpr.2016.10.005
- Woodward, L. J., & Fergusson, D. M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 1086–1093. doi:10.1097/00004583-200109000-00018

Paper 3

Rasmussen, L-M. P., Patras, J., Handegård, B. H., Neumer, S-P., Martinsen, K. D., Adolfsen, F., Sund, A. M., & Martinussen, M. (*In review*). A validation of the adapted version of the Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT).

A validation of an adapted group-based version of the Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS CBT)

Abstract

The Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS CBT; Bjaastad et al., 2016) was developed to evaluate the delivery of cognitive therapies for children with clinical anxiety. The present study is an evaluation of the adapted version of the CAS CBT using a sample of group leaders delivering a newly-developed, CBT-based group intervention: *EMOTION: Kids Coping with Anxiety and Depression* (Martinsen, Stark, Rodriguez, & Kendall, 2014). We used a confirmatory factor analysis (CFA) approach in Mplus to test the factor structure of the 11-item instrument. Six raters evaluated a total of $N = 239$ video-recorded sessions of the EMOTION program. Results showed that we were not able to obtain adequate model fit for the unidimensional 11-item scale $\chi^2 = 497.076, p < .05, df = 44$; RMSEA = 0.208, $p < .05$; CFI = .953; and TLI = .941 or the alternate two-subscale solution (i.e., adherence and competence). The final tested model, which removed the items related to session goals, yielded improved but not excellent model fit, $\chi^2 = 23.26, p < .05, df = 11$; RMSEA = 0.068, $p = .19$; CFI = .998; and TLI = .997. Further revision of the CAS CBT instrument in order to address group-based interventions may be warranted.

Key words: youths – anxiety – depression – transdiagnostic – observation – validation – confirmatory factor analysis (CFA)

Manual-based interventions consist of prescribed procedures with specified goals and activities designed to produce changes in the target group. Treatment fidelity, or treatment integrity, refers to the therapists' ability to follow the program's core components, which are necessary to produce the desired outcomes (Bond, Evans, Salyers, Williams, & Kim, 2000; Dane & Schneider, 1998; Dusenbury, Brannigan, Falco, & Hansen, 2003). Perepletchikova, Treat, and Kazdin (2007) refer to treatment fidelity using three characteristics. These are 1) adherence, reflecting the therapists' utilization of prescribed intervention procedures, 2) competence, which represents how well the intervention is conducted, and 3) treatment differentiation, which indicates if the treatment differ from others. A high degree of fidelity to an effective program is associated with better treatment outcomes (Carroll et al., 2007; Durlak & DuPre, 2008; Perepletchikova & Kazdin, 2005), but fidelity has received less attention in treatment studies compared to the effectiveness of the intervention (Perepletchikova et al., 2007). It is therefore important to measure both treatment fidelity as well as treatment outcome when evaluating a manual-based intervention. Fidelity is also an important implementation outcome because it measures how well staff have been trained and supported to use the new intervention (Carroll et al., 2007).

Adherence and Competence

Adherence and competence have received a great amount of interest with regard to manualized therapies, mainly to assess and monitor treatment integrity (Perepletchikova & Kazdin, 2005). The present study focuses on a measure comprised of two factors labelled structure and process (Bjaastad et al., 2016), which were originally developed for measuring adherence and competence in CBT for children with anxiety disorders. The structural dimension is analogous to adherence; indicating whether the key components or active ingredients of the program were delivered, and to what extent the intervention was completed (O'Donnell, 2008; Odom, 2008). The process dimension of the measures reflects the quality

of the interaction and relationship between the therapists` and the child(ren) (Justice, Mashburn, Hamre, & Pianta, 2008; O'Donnell, 2008). However, as Bjaastad et al., (2016) found, there are considerable overlap between the adherence questions and competence questions in the two suggested factors, in which the factor structure contains both adherence- and competence-items and vice versa for the process dimension. Researchers recognize the need to address both these dimensions to understand how interventions impact outcome (Odom, Collet-Klingenberg, Rogers, & Hatton, 2010; Webb, DeRubeis, & Barber, 2010). Commonly used methods to assess fidelity are self-reports and observations of the sessions.

Self-report and Observations

In the field of cognitive behavioral therapy (CBT), self-reports such as the Cognitive Behavioral Therapy Checklist (CBTC; Kendall, Gosh, Albano, Ginsburg, & Compton, 2001) have the advantages of being easier to administer and demanding less resources than observations. Filling out self-reports and checklists following delivery can also serve as a reminder to the interventionist about program contents, which in turn can serve to reinforce intervention core components (Bellg et al., 2004). Self-reports, however, rely on individuals' ratings of their own performance, which allows for potential reporter bias (Bellg et al., 2004). Observations, by contrast, are conducted by third parties and are therefore considered a more rigorous and objective measure of treatment adherence (Hogue, Liddle, & Rowe, 1996). Observations, however, require the presence of recording equipment or trained observers in the intervention sessions, which can be time-consuming and expensive. Few of the measures that exist for CBT with children have been evaluated psychometrically, regardless of whether they are self-report or based on observation.

Bjaastad and colleagues (2016) developed the Competence and Adherence scale for CBT (CAS CBT), which is designed for assessing adherence and competence during therapy on youths with anxiety disorders. Anxiety and depression in children are among the most

prevalent psychological problems (Merikangas, Nakamura, & Kessler, 2009), and therefore it is important to develop instruments to measure fidelity when targeting these problems. Also, considering that CBT a commonly used therapy to address these mental health problems (Crowe & McKay, 2017), adequate measures to evaluate adherence and competence in the different interventions being used is therefore evident. Further, to help determine the successfulness of a specific intervention in relation to outcomes, it is important to focus on fidelity (Durlak & DuPre, 2008). It may also help clarify if failures reflect the intervention itself or how it was implemented.

Validation

Construct validity refers to whether a test measures what it is supposed to measure, and is often investigated using a confirmatory factor analysis (CFA) (EFPA, 2013; Floyd & Widaman, 1995). Instrument validation is important to ensure that the instrument used can be applied to similar contexts. For instance, CBT-based interventions for indicated prevention share a lot of common features with clinical therapy, however, conducting interventions in the prevention field involves a number of unverifiable factors (e.g., undefined symptoms in the children, scheduling issues, etc.). Also, resources aligned to support implementation are often limited (Forman, Olin, Hoagwood, Crowe, & Saka, 2009), and typically, assessing adherence and competence is often omitted from prevention studies (Cross & West, 2011; Dane & Schneider, 1998; Dumas, Lynch, Laughlin, Phillips Smith, & Prinz, 2001). Observations of fidelity are particularly rare given the extra resources needed (Hogue et al., 1996; Schoenwald et al., 2011). Furthermore, although highly educated and experienced within their field, many of those working in prevention services do not have formal CBT training. This may impact delivery of a CBT-based program, and is therefore also a reason to evaluate how the interventions are delivered.

The current study

The main goal of the current study was to test the reliability and examine the factor structure of the Competence and Adherence Scale for Cognitive Behavioral Therapy (Bjaastad et al., 2016) with a sample of group leaders delivering a newly developed preventive, CBT-based intervention; EMOTION: *Kids Coping with Anxiety and Depression* (Martinsen et al., 2014). This study was part of a Norwegian multi-site randomized controlled trial (RCT), investigating the effectiveness and the implementation of the EMOTION program (Patras et al., 2016). The CAS CBT has mainly been used with trained therapists working in outpatient clinics treating youth with clinical anxiety. Bjaastad and colleagues (2016), conducted an exploratory factor analysis (EFA) with oblique rotation (direct oblimin principal component analysis), in which two factors were obtained; 1) CBT structure and session goals and 2) Process and relational skills. CAS CBT also showed good internal consistency ($\alpha = .87$), good to excellent interrater reliability (ICC = .83 for Adherence and .64 for Competence) and high rater stability with an ICC = .89 for Adherence and .92 for Competence when the videos were rescored after an average of 17.4 months (Bjaastad et al., 2016). There is a need however, also highlighted by the CAS CBT developers (Bjaastad et al., 2016), to independently validate the instrument using manualized interventions targeting related problem areas, but with different delivery modalities and target groups.

Method

Participants

Participants were trained group leaders ($N = 68$) from different municipal mental health and child welfare services in Norway (e.g., school health services) delivering the EMOTION program. The study sample was 94% women with a mean age of 39.6 ($SD = 9.7$ years). The group leaders were psychologists/specialists (35%), health nurses (14%), educational and psychological counsellors (18%), educators (11%), child-care workers (6%), occupational

therapists (3%) as well as psychology students (5%) and 8% “others” (e.g., counsellor, project leader etc.). Almost 70% of the participants had former experience working with anxiety and depression in youths, and 38% had previously worked with cognitive behavioral therapy (CBT). They received a three-day training, with one-day introduction in general CBT, followed by a two-day workshop in the specific program components of the EMOTION program. The children ($N = 266$) in the RCT study undergoing the EMOTION program had a mean age of 9.64 years ($SD = 0.93$), where 56.9 % were girls. The children were recruited based on scores above a predetermined cutoff on anxiety and/or depression instruments. A total of $N = 239$ sessions (17% of all sessions) were recorded and scored for $N = 52$ groups.

The EMOTION intervention

The EMOTION program (Martinsen et al., 2014) is a group-based preventive intervention aimed at reducing symptoms of anxiety and depression in children 8-12 years. The intervention builds on regular CBT principles, and during the 20 sessions (one hour sessions, twice per week in a school setting), the main goals were to teach children different sets of skills and strategies to be able to handle their anxiousness or sadness. Additionally, parents received a seven-session course where four of these sessions were together with the children. The parent sessions consisted of themes like positive reinforcements and punishment, positive time with the child, in addition to learning some the same skills as the children learned in their groups. Both the child and parent group sessions were run by two group leaders who were trained in the EMOTION intervention and received regular supervision from an expert in CBT.

Procedure

The research staff distributed video cameras to the intervention group leaders before starting new groups. At the same time a list of the sessions that each leader was to record was distributed. A block of four concurrent children sessions and two concurrent parent sessions

were chosen for each group. The first of each session block was chosen randomly in order to get coverage of a variety of sessions. Sessions were chosen in blocks to simplify the data collection for the group leaders. For example, a group leader may have been randomly assigned to start with session 10, and then follow with sessions 11, 12, and 13. The first and the last session were excluded from the fidelity checks due to the content (introduction and finalization of the groups, respectively). When the groups were finished, the project staff collected and stored the video files at a secure server at one of the participating sites. Ethical approval was obtained from The Regional Committee for Health and Medical Research Ethics (2013/1909/REK Sør-Øst), and the study was registered in clinical trials (NCT02340637).

Measure

The CAS CBT consists of 11-items, divided into three main sections, covering key domains in CBT for children with anxiety (Bjaastad et al., 2016). This includes cognitive therapy structure (e.g., homework, session structure, and progress), process- and relational skills (e.g., reinforcement, collaboration, and flexibility) and treatment goals (specific goals for the session based on the treatment protocol). Adherence is assessed by different items within each of the main sections (e.g., homework, session structure, and progress), while competence is scored globally for each of the main sections. This means that the competence item “cognitive therapy structure” includes an overall competence assessment of both homework and session structure/progress. This potentially causes one of the items being more emphasized than the other(s) within the same section. For instance, the question regarding homework reflects whether the therapists reviewed the participants’ homework, and handed out new for next session. The structure/progress question reflects whether the therapists sets an agenda and follows this (including reviewing/presenting new homework), time administration, and general flow of the session. Hence, the structure/progress item generally

receives more emphasis. Further, the item “Flexibility” is rated as a competence score. In addition, there are two questions assessing the overall adherence and competence of the session. These are scored globally, and was added as supplementary items to the scale. The adherence score was rated from 0 = *None* to 6 = *Thorough*. The competence score ranges from 0 (*Poor skills*) to 6 (*Excellent skills*), with an explanation attached to the ratings, indicating different qualities which needed to be fulfilled. Furthermore, there are three questions about the video quality and challenges with the session (i.e. “What is your evaluation of how challenging this session was?”).

In this study, we made a few adaptations of the instrument to fit the EMOTION program in collaboration with the CAS CBT developer. In the original CAS CBT, the parents were included with one item called “parental involvement” (Bjaastad et al., 2016). In EMOTION, the parents received seven individual sessions and therefore this item was removed. The seven parent sessions were rated separately with the same structure as the CAS CBT for children. Also, in the original version, there were two program goals to be rated, but in our version we had up to three goals, so one item was added. The instrument developer(s) approved the modifications.

Rating of items, particularly regarding goals for the sessions were different for each time. During scoring, the items were assessed independently starting at the highest score (6 = *Thorough*) and subtracted accordingly for each element within each goal that was omitted (by the group leaders). If one of the goals for the session was not conducted, the score was 0 (*None*), regardless the reason for this (i.e., prioritizing underway, external factors, time etc.). The group leaders had to present the specific goals in the session, and follow them as described in the program manual to be evaluated by the raters. During the sessions, the two group leaders were scored as one unit, and not assessed individually as they were not assigned a primary and secondary role.

Raters

A scoring team consisting of six people, including both an experienced researcher with previous experience using the instrument, and students with a master's degree or higher in psychology or child care, rated a total number of 239 (17%) videos (170 child sessions and 69 parent sessions). The experienced researcher (scoring 40 individual videos and 66 videos for ICC), with previous clinical practice and video rating experience became the expert rater, which the other raters were tested against. The scoring team received one day of training by the instrument developer in the core elements of the scoring instrument (CAS CBT). In addition, they received a two-day training in the EMOTION program; similar to the group leader-training, focusing on key aspects of the program, session by session. Prior to start up, the raters had to score three of the same videos for training purposes and checking for inter-rater reliability (ICC), and if consensus was met with the expert rater, they could continue. During the project period, ongoing reliability tests were conducted which resulted in a total of 66 randomly selected videos used for interrater reliability (See Table 1 for an overview). Additionally, the team had regular meetings to calibrate, reach consensus and limit drifting. The raters received randomly assigned video recordings for scoring. All raters signed a declaration of confidentiality.

[Insert Table 1 near here]

Statistical analyses

Interrater reliability. The reliability analyses and descriptive analyses were conducted using SPSS statistical packages (24.0). Interrater reliability between raters was calculated using intraclass correlations (ICC, [3, 1]; Shrout & Fleiss, 1979). The ICC's were calculated by using the model (3, 1) with absolute agreement, which is a Two-Way Mixed Effects Model where people effects are random and measures effects are fixed. The videos were scored by the expert rater and compared against the other observers using the single

measure option. Results were guided by Chicetti`s (1994) principles were ICCs < .40 is considered poor agreement, ICCs between .40 to .59 indicate fair agreement, ICCs between .60 to .74 reflect good agreement and ICCs > .75 show excellent agreement.

[Insert Table 2 near here]

Confirmatory Factor Analyses (CFA). We employed a confirmatory factor analysis (CFA) using Mplus 7.0 statistical software with the weighted least squares estimator (WLSMV; Muthèn & Muthèn, 1998-2010), where the indicators were set as ordered categorical (ordinal). Based on the origin of the instrument and the structure of the items (i.e., competence items depending on the adherence items) the test strategy were as follows (for an overview, see Table 2);

Model 1: Based on the theory behind the instrument, and because the items assessing competence is closely connected to the adherence-items (indicating high correlations), we first tested the fit of a unidimensional model with all items loading on one factor – fidelity.

Model 2: Then, given the structure of the instrument and scoring instructions, we examined a model with the two other factors primarily being evaluated and scored by the instrument (adherence and competence).

Model 3a: Further, we investigated if the originally proposed two-factor structure of CAS CBT (Bjaastad et al., 2016) would be replicated in the current sample (CBT structure and goals and relational skills).

Model 3b: As an extension of the previous model (model 3) and based on methodological issues with reference to how the session goal items were rated during scoring, we tested a two-factor model including 1) structure and goals, and 2) relational skills with correlated item-residuals. The competence items depends highly on the items assessing adherence within the same topic (i.e. competence score on structure and process depends very much on the adherence score structure and process and homework). In addition, the adherence

items are emphasized unevenly (i.e. item 2 “Structure and progress” is emphasized more than item 1 “Homework review and planning new homework”), which indicate a higher correlation between item 2 and item 3, and therefore substantiate the reason to correlate the residuals of these two items.

Model 3c: Furthermore, the item evaluating competence on process and relational skills (item 7) depends highly on the items within the same topic, especially with item 6 (“Flexibility”). Because these items share some common features (i.e., both assesses flexibility and competence, while the two others within this topic is adherence-items), it is expected that the residuals for these items will correlate, and we therefore incorporated this in our model. The correlation between these items were $r = .91$, supporting our indication that these items measure similar constructs.

Model 4: Built on a modified version of the previous model, we investigated a two-factor model; 1) structure and goals, and 2) relational skills, with an alternative structure. Based on the how the items regarding goals for the session were assessed during scoring, not being able to capture why some goals were excluded for instance, could indicate that the latent factor “Structure” is not capable of modelling these particular items adequately. Therefore, we tested whether removing the items evaluating the session goals (items 8-11) would improve model fit. The low correlation between the adherence items within this topic ($r = .04$ to $.42$), reinforced our reasons to examine this model.

To assess how well the model fits the data, multiple fit indices were examined. Chi square is a commonly reported measure of model fit, where a significant result ($p < .05$) indicates misfit (Kline, 2011). Also, Root Mean square Error of Approximation (RMSEA; Steiger and Lind 1980) $< .08$, and a Bentler’s Comparative fit index (CFI; Bentler, 1990) and Tucker-Lewis Index (TLI; Tucker & Lewis, 1973) $> .90$ indicate adequate model fit. Preferably, for a good model fit, RMSEA should be $< .05$, and the CFI and TLI should be

> .95. Also, for RMSEA, a p-value is given, and is interpreted as the probability that the RMSEA < .05).

Correlations. Inter-item correlations between the items were computed using polychoric correlations, which takes into account the ordinal measurement level of the Likert-scale. Correlations between the total scores of structure and CBT session goals, as well the adherence and competence total scores were computed using Person's r .

Internal consistency. Internal consistency for the subscales were estimated with Cronbach's alpha, where values > 0.70 reflect adequate consistency (EFPA, 2013).

Results

Approximately 20% ($N = 267$) of the total number of sessions were video recorded and scored using the modified version of CAS CBT (Bjaastad et al., 2016). However, some of the videos were not scored (e.g., only parts of the session were recorded due to technical issues, poor video quality or camera placement made scoring impossible, or the group leaders failed to record the whole session). This resulted in 239 (17 %) individually recorded child and parent sessions for 52 groups ($M = 3.0$, $SD = 1.61$ sessions per group).

Interrater reliability

Results showed fair to good interrater reliability (from $\alpha = .40$ to $.74$) on all items, and on the mean adherence and mean competence score across all raters compared with the expert rater. See Table 3 for a complete overview of the Mean (SD), and ICC scores between the expert rater and the student raters. The items reflecting process and relational skills generally received the lowest scores (.42 to .52), indicating that these were more difficult for the raters to evaluate.

[Insert Table 3 near here]

Confirmatory factor analyses (CFA)

Model 1: Based on the theory behind treatment fidelity and development of the instrument, as well as the structure of the instrument and correlations between the different items, we tested a unidimensional model to check whether the different items loaded on one latent factor (i.e. fidelity). Results showed poor model fit ($\chi^2 = 497.076, p < .05, df = 44, RMSEA = 0.208, p < .05, CFI = .953, and TLI = .941$), which indicates that there are underlying issues within the model which is not taken into account.

Model 2: Further, because the structure of the CAS CBT advocates an adherence- and a competence evaluation of the sessions, we investigated whether the unexplained variance contributed to model by these two constructs which is primarily being assessed by the raters during scoring (i.e. items scored as “adherence” explained by a latent Adherence variable and items scored as “competence” explained by a latent Competence variable). This was also supported by the high inter-item correlations (Table 4). However, results showed that the model covariance matrix was not positive definite, and the technical output in Mplus showed a correlation between the latent factors “adherence” and “competence” > 1 ($r = 1.074$), which probably implies a model misspecification and that the results are not trustworthy.

Model 3a: Assessing the two component model identified by Bjaastad and colleagues (2016), indicated good fit according to the CFI and TLI, but poor model fit according to the RMSEA and significant misfit according to the chi-square test, $\chi^2 = 183.69, p < .05, df = 43; RMSEA = 0.117, p < .05; CFI = .985; and TLI = .981$.

Model 3: Based on theory, we modified model 3 by allowing the residuals of item 2 (Adherence: Structure and progress) and item 3 (Competence: Cognitive therapy structure) to

correlate, which improved the model results slightly ($\chi^2 = 162.10, p < .05, df = 42; RMSEA = 0.109, p < .05; CFI = .987; and TLI = .984$), but we were not able to get an adequate fit.

Model 3c: Further, as a particularly strong association was expected between item 6 (“Flexibility”), which is also a competence item, and item 7 (“Competence score for Process and relational skills”), we tested a model where we correlated the residuals between item 6 and item 7, but it did not improve model fit ($\chi^2 = 163.37, p < .05, df = 41; RMSEA = 0.112, p < .05; CFI = .987; and TLI = .983$).

Model 4: Lastly, we tested a model where all the session goal items were omitted. This model yielded an improved fit over the previous models, $\chi^2 = 23.26, p < .05, df = 11; RMSEA = 0.068, p = .19; CFI = .998; and TLI = .997$. See Table 5 for an overview of the different models tested with pertinent model fit indices.

[Insert Table 4 near here]

[Insert Table 5 near here]

Internal consistency

Based on the models tested in this study, the alpha for the subscale “CBT structure” (excluding session goal items) was .85 which indicated a good reliability, and an excellent alpha for the “Process and relational skills” subscale ($\alpha = .93$).

Discussion

This study was conducted to examine the factor structure and reliability of the Cognitive and Adherence Scale for Cognitive Behavioral Therapy (CAS CBT; Bjaastad et al., 2016) in a population of children receiving a preventive group intervention for symptoms of anxiety and depression. Results from our study showed that we were not able to estimate a good model fit when conducting a CFA in our sample, particularly when we included the items evaluating

the session goals. We were unable to replicate the results from Bjaastad and colleagues (2016).

Different models were tested to investigate the structure of CAS CBT beyond the hypothesized replication. A unidimensional model, examining if a higher order latent factor could explain the observed variance-covariance matrix was discarded, indicating that it was too simplistic. This implies that an overall dimension, such as fidelity (or treatment integrity), is difficult to model with all 11 items included. Further, the model evaluating the two dimensions being assessed by the raters during scoring (i.e., adherence and competence) was also discarded due to misspecification of the model, as we were not able to estimate all the parameters reliably. Possibly, this was caused by the high correlations and the high dependency between the items in the two suggested factors, particularly the strong associations between the competence-item and the adherence-items within the same topic, which the model is not able to account for.

Issues regarding high dependency between items was associated to different aspects of the instrument. For instance, while we were able obtain an acceptable internal consistency, we did not receive an adequate model fit when we conducted a CFA. The items cohered to such a large extent that it was easy to compare the scores to each other (scoring high on one item ultimately indicated a high score on the next item). Further, the competence questions were consistently being evaluated based on a global assessment of two or three adherence questions, where the previous questions seemed to explain much of the variance within this factor. For instance, within the topic reflecting “Cognitive therapy structure, the competence question (item 3) is strongly associated with the adherence-questions (item 1-2). The competence question (item 7) in the topic “Process and relational skills” is highly based on the adherence-questions (item 4-5), in addition to the other competence question in the same topic (item 6). Within the “Session goal” topic, the competence score for the session goals

(item 11) is also strongly linked to the accompanying adherence questions (item 8-9). During scoring, the raters will therefore base the competence score mostly on the adherence-ratings, but emphasize them differently. Furthermore, in global video observations such as this, the observers are interpreting and evaluating a concept of interest (i.e., adherence and competence). Guided by a scoring manual and a Likert-type scale, the raters make a judgement based on what they observe, and it will therefore have an impact on the scores. Although the raters in the present study met regularly, there should be more focus on training and how the items are rated, along with more frequent meetings to discuss scoring. Another aspect to consider is that the individuals being observed are aware of the situation and might be affected by having the camera nearby (i.e. nervous or more adherent), which may not always give an accurate picture of what is going on (Breitenstein et al., 2010). This could be partly addressed by having the group leaders record all the sessions and then choose sessions randomly for scoring.

In general, the correlations between the different items were high across the two factors (structure and relation) which originally were meant to be tested. High inter-correlations between the items in the instrument supposedly loading on different factors, as well as low correlations between items loading on the same factor, have an impact on the results. Especially, the adherence items rating the goals for the sessions showed low inter-item correlations, and further showed higher correlations with the items reflecting relational skills. This indicated that some of these items did not adequately fit the model. Looking closer at the individual items, the lack of correlation is not a total surprise. The goals for the sessions are independent, indicating that you do not have to complete one before moving to the next one. This may imply some issues regarding the scoring of these items, and possibly elements of the items which were not captured by the scoring, such as the difference between missing (not completed at all) and a total lack of adherence.

The different goals also varies from session to session, which makes them difficult to fit adequately in the instrument structure. This could be reflected by checking the distribution of the response categories, which was highly positively skewed (on a scale from 0 = *None* to 6 = *Thorough*), but very high on category 1 (almost not present). This indicate an uneven distribution of the adherence to the session goals. One reason for this could be the transdiagnostic origin of the EMOTION manual, which was quite comprehensive, including many elements for each session. For the program developers, choosing two or three main goals per session was challenging, and this could have impacted the completion, and therefore also the scoring of these particular items. This was partly reinforced by removing the items evaluating the session goals (model 4), which improved the model slightly, although good model fit was not achieved. We recommend a focus on the session goals, both in relevance to training and scoring, but also how they should be interpreted and analysed. For example, including information on outer factors that might impact the adherence to the session goal items (i.e., external factors such as time constraints or ongoing prioritizing affecting lack of adherence).

In Bjaastad and colleagues' (2016) original article, they assessed therapists' individual treatment of anxiety, but were not able to conduct a factor analysis on the group condition due to sample size. This could have been a bias in our study, as group condition, with up to 10 children, could potentially contribute with some issues that is not present during individual treatment and which we were not able to assess with the instrument in its current state (e.g. group dynamics, conflicts between the children, noise etc.). This might have affected the completion of the session goals, and subsequently the scoring of the session. Future studies could adapt for this by including additional questions to assess group dynamics. Also, as this was a preventive intervention targeting children with symptoms of anxiety and depression,

many of the children had unspecific symptoms and unestablished issues, which is more difficult to treat. Hence, the session outcome could be more difficult to evaluate.

One possible interpretation of the lack of model fit could also simply be rooted in methodological issues, and that the original factor structure was tested with a different approach (principal component analysis, with oblique rotation) than in our study. Lastly, another element that should be mentioned is that the name of the instrument and the topics being evaluated during scoring (adherence and competence) are not the same as those the structure of instrument reflects according to the factor analyses (i.e., CBT structure and goals *and* Process and relational skills). This introduces some confusion regarding interpretation and use of the instrument, which merit some consideration in future applicability.

Limitations

Group leaders were rated as one unit and were not given specific tasks or roles in advance. Preferably, a unique score for the two individuals would be optimal to be able to detect any variation between the group leaders. Alternately, assigning the group leaders` different roles as primary and secondary, would produce individual scores which is not influenced by the other group leader.

Further, as the instrument was slightly modified to fit our study, we had to remove one of the items assessing parental involvement, and also added one question assessing adherence for a third goal for the session. This implies that the validation of the instrument is not conducted on the exact same items as the original version of the instrument.

Optimally, we would have performed a multilevel confirmatory factor analysis (MCFA) to assess the between (groups) and within-level (sessions) of the data. However, in our study, the group leaders delivered the intervention in pairs, and therefore the groups were treated as the unit of analysis at the between level because the groups comprised different combinations of group leaders. Further, as the individual-level CFA model did not fit the data,

the MCFA was not warranted (Hox, 2002). We did however consider if the group level had an impact, by testing between-group variance on the two factors primarily being tested. The alpha was low ($\alpha = .16$ for Structure and session goals and $\alpha = .20$ for Process and relation skills), and therefore not relevant in this study.

Although within the acceptable range, some of the inter-rater reliability scores were in the lower range $< .50$, particularly for the items assessing process and relational skills (e.g. Positive reinforcement, Collaboration, Flexibility). This implies that either it was difficult to come to an agreement regarding these items, or there was something with the instrument that makes it difficult to calibrate and reach consensus when scorings these items. Also, the more raters, the more difficult it will be that everyone completely agrees on all items. Focusing on training and conducting accuracy testing frequently should be obliged, as well as keeping the number of raters to a minimum.

Conclusion

In this study, we found similar factors as originally proposed, however, we were not able to replicate the factor structure with adequate model fit in a CFA. There were some issues with the model, particularly when session goal-items were included, mostly reflected by the high correlations and dependence between the items. This indicates that further development of the measure is warranted and that a revision may be in order to adequately assess the use of different manualized CBT group interventions applied on children with symptoms of anxiety and depression.

References:

- Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology, 23*, 443-451. doi:10.1037/0278-6133.23.5.443
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246. doi:10.1037/0033-2909.107.2.238
- Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W., Torsheim, T., Havik, O. E., Heiervang, E. R., & Öst, L.-G. (2016). Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. *Psychological Assessment, 28*, 908-916. doi:10.1037/pas0000230
- Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H.-W. (2000). Measurement of Fidelity in Psychiatric Rehabilitation. *Mental Health Services Research, 2*, 75-87. doi:10.1023/a:1010153020697
- Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*, 40.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. In (Vol. 6, pp. 284-290). US: American Psychological Association.
- Cross, W. F., & West, J. C. (2011). Examining implementer fidelity: Conceptualizing and measuring adherence and competence. *Journal of Children's Services, 6*, 18-33. doi:10.5042/jcs.2011.0123
- Crowe, K., & McKay, D. (2017). Efficacy of cognitive-behavioral therapy for childhood anxiety and depression. *Journal of Anxiety Disorders, 49*, 76-87. doi:[10.1016/j.janxdis.2017.04.001](https://doi.org/10.1016/j.janxdis.2017.04.001)
- Dane, A. V., & Schneider, B. H. (1998). Program Integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review, 18*, 23-45. doi:[10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3)
- Dumas, J. E., Lynch, A. M., Laughlin, J. E., Phillips Smith, E., & Prinz, R. J. (2001). Promoting intervention fidelity: Conceptual issues, methods, and preliminary results from the EARLY ALLIANCE prevention trial. *American Journal of Preventive Medicine, 20*, 38-47. doi:[10.1016/S0749-3797\(00\)00272-5](https://doi.org/10.1016/S0749-3797(00)00272-5)
- Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350. doi:10.1007/s10464-008-9165-0
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237-256. doi:10.1093/her/18.2.237
- EFPA. (2013). *EFPA Review model for the description and evaluation of psychological tests: Test review form and notes for reviewers version 4.2.6*. In.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286-299. doi:10.1037/1040-3590.7.3.286
- Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy: Theory, Research, Practice, Training, 33*, 332-345. doi:10.1037/0033-3204.33.2.332
- Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Justice, L. M., Mashburn, A., Hamre, B., & Pianta, R. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly, 23*, 51-68. doi:10.1016/j.ecresq.2007.09.004
- Kendall, P. C., Gosh, E. A., Albano, A. M., Ginsburg, G. S., & Compton, S. (2001). *CBT for child anxiety: Therapist treatment integrity checklist*. Philadelphia: Temple University.
- Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3 ed.). New York: Guilford Publications.
- Martinsen, K. D., Stark, K., Rodriguez, K. O., & Kendall, P. C. (2014). *Mestrende barn Manual*. Oslo: Gyldendal Norsk Forlag.
- Merikangas, K. R., Nakamura, E. F., & Kessler, R. C. (2009). Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neuroscience, 11*, 7-20.
- Muthèn, L., & Muthèn, B. (1998-2010). *Mplus : Statistical analysis with latent variables : User's guide*. Los Angeles, CA: Muthen & Muthen.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*, 33-84. doi:doi:10.3102/0034654307313793
- Odom, S. L. (2008). The Tie That Binds: Evidence-Based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*, 53-61. doi:10.1177/0271121408329171
- Odom, S. L., Collet-Klingenberg, L., Rogers, S. J., & Hatton, D. D. (2010). Evidence-Based Practices in Interventions for Children and Youth with Autism Spectrum Disorders. *Preventing School Failure: Alternative Education for Children and Youth, 54*, 275-282. doi:10.1080/10459881003785506
- Patras, J., Martinsen, K. D., Holen, S., Sund, A. M., Adolfsen, F., Rasmussen, L.-M. P., & Neumer, S.-P. (2016). Study protocol of an RCT of EMOTION: An indicated intervention for children with symptoms of anxiety and depression. *BMC Psychology, 4*, 48. doi:10.1186/s40359-016-0155-y
- Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12*, 365-383. doi:10.1093/clipsy.bpi045
- Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of consulting and clinical psychology, 75*, 829-841. doi:[10.1037/0022-006X.75.6.829](https://doi.org/10.1037/0022-006X.75.6.829)
- Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research, 38*, 32-43. doi:10.1007/s10488-010-0321-0
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*. doi:10.1037/0033-2909.86.2.420
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10. doi:10.1007/bf02291170
- Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical psychology, 78*, 200-211. doi:10.1037/a0018912

Tables

Table 1

Distribution of Videos per Observer (Single Scored and ICC)

	Observer						Total
	Expert	R1	R2	R3	R4	R5	
Single videos scored	40	37	22	82	27	31	239
Videos used for ICC		19	10	15	12	10	66
ICC Adherence		.67	.54	.83	.69	.86	
ICC Competence		.45	.45	.63	.58	.53	

Note: ICC = Intraclass correlation [3, 1] by Shrout and Fleiss, 1979; Two-Way Mixed Effect model, single measurement (absolute agreement).

Table 2

Overview of the Test Strategy

Model	Reasoning	Factor(s)
1. Unidimensional	Based on the theory behind the instrument (treatment fidelity/integrity).	1) Fidelity
2. Two factor alternative structure	Based on the structure and scoring instructions of the instrument	1) Adherence 2) Competence
3a. Two factor structure	Based on the original EFA from Bjaastad et al. 2016	1) Structure and goals 2) Relational skills
3b. Two factor; correlating residuals (item 2 and 3)	An extension of the previous model, based on the dependence between adherence and competence-items and unequal emphasis on the items (2-3)	1) Structure and goals 2) Relational skills
3c. Two factor; correlating residuals (item 6 and 7)	An extension of the previous model, based on the dependence and correlation between items and unequal emphasis on the items (6 and 7)	1) Structure and goals 2) Relational skills
4. Two factor; removing session goals	Modified version of Model 1; Based on methodological grounds (scoring of the sessions goals)	1) Structure and goals 2) Relational skills

Table 3

Inter-Rater Reliability Between Expert and Student Raters for the 11-item CAS CBT Scale and Mean Adherence/Competence.

Item/variable	M (SD)			ICC
	Total	Expert rater	Student raters (n = 5)	
N videos	<i>n</i> = 239 ^a		<i>n</i> = 66 ^b	
1. Homework review/planning homework	3.46 (1.97)	4.00 (2.05)	3.21 (2.07)	.60
2. Structure and progress	3.59 (1.52)	3.20 (1.69)	3.12 (1.66)	.60
3. Cognitive therapy structure (items 1-2)	3.36 (1.48)	3.29 (1.74)	3.03 (1.49)	.52
4. Positive reinforcement	3.91 (1.32)	3.83 (1.47)	3.55 (1.54)	.48
5. Collaboration	4.06 (1.38)	4.24 (1.18)	3.83 (1.38)	.40
6. Flexibility	4.00 (1.36)	4.15 (1.26)	3.64 (1.44)	.42
7. Process and relational skills (items 4-6)	3.90 (1.32)	4.23 (1.25)	3.44 (1.42)	.52
8. Session goal 1	3.53 (1.61)	3.15 (2.12)	3.15 (1.85)	.63
9. Session goal 2	2.93 (2.10)	2.58 (2.32)	2.82 (2.15)	.74
10. Session goal 3 ^c	2.61 (1.95)	1.65 (1.60)	1.68 (1.67)	.55
11. Session goals (items 8-10)	3.19 (1.47)	3.08 (1.76)	2.75 (1.49)	.56
12. Global adherence	3.60 (1.47)	3.18 (1.87)	3.23 (1.37)	.49
13. Global competence	3.60 (1.40)	3.55 (1.38)	3.21 (1.37)	.51
Mean score adherence (7 items)	3.55 (1.24)	3.43 (1.33)	3.19 (1.23)	.60
Mean score competence (4 items)	3.61 (1.26)	3.69 (1.34)	3.22 (1.30)	.60

Note: Total scale (11 items); Item 1, 2, 4, 5, 8, 9, 10 (and 12) represents adherence scores. Item 3, 6, 7, 11 (and 13) represents competence scores. The adherence score was rated from 0 = *None* to 6 = *Thorough*. The competence score ranges from 0 (*Poor skills*) to 6 (*Excellent skills*).

^a*N* = 239 individual videos scored only once,

^b*N* = 66 videos used for interrater reliability calculations;

^c*N* = 140 videos scored with session goal 3 (not applicable to all sessions).

Table 4

Polychoric Correlations Between Items

Item	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
2.	.58**									
3.	.71**	.89**								
4.	.56**	.57**	.66**							
5.	.55**	.54**	.69**	.72**						
6.	.51**	.56**	.69**	.74**	.82**					
7.	.56**	.56**	.75**	.85**	.88**	.91**				
8.	.46**	.58**	.59**	.43**	.55**	.51**	.51**			
9.	.39**	.60**	.58**	.40**	.33**	.34**	.39**	.24**		
10.	.24**	.50**	.44**	.29**	.24**	.17	.22*	.04**	.42**	
11.	.57**	.78**	.83**	.67**	.67**	.67**	.73**	.67**	.71**	.53**

Note. $N = 239$. *Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

Table 5

Overview of the Models Tested with Model Fit Indices

Model	df	Chi square	RMSEA	CFI	TLI	WRMR
1.	44	497.076*	0.208*	0.953	0.941	1.779
2.	43	477.588*	0.206*	0.955	0.942	1.728
3a.	43	183.694*	0.117*	0.985	0.981	0.974
3b.	42	162.104*	0.109*	0.987	0.984	0.898
3c.	41	163.370*	0.112*	0.987	0.983	0.890
4.	11	23.263*	0.068 ^a	0.998	0.997	0.362

Note: $N = 239$; * $p < .05$; ^a $p = .19$