Social disparities in interventions for behavior problems

Truls Tømmerås

Dissertation presented for the degree of Philosophiae Doctor (PhD) 2017

Department of Sociology and Human Geography

Faculty of Social Sciences

University of Oslo

Series of dissertations submitted to the
Faculty of Social Sciences, University of Oslo
No. 677

ISSN 1564-3991

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Acknowledgements

I would like to thank Terje Ogden at the Norwegian Center for Child Behavioral Development for allowing me the chance to take my Ph.D. I am very lucky to have had the chance to be part of such a creative and friendly environment at the Norwegian Center. First of all, I have to thank John Kjøbli for setting up the TIBIR research project and subtly leading me into it. First, Dr. Kjøbli persuaded me to take my master thesis using TIBIR data, and second, he encouraged me to pursue a Ph.D.-degree. I am also grateful for all the inspiring conversations we have had about business and pleasure, and for you being one of my supervisors. I also have to thank Elisabeth Askeland and Terje Christiansen for all your encouraging words and support. After a while, I started to believe in it myself.

I also want to thank my supervisor from University of Oslo Torkild Hovde Lyngstad for his direct and clarifying feedback on the papers and the Ph.D.-thesis. Moreover, I would also like to thank my fellow co-authors in the four papers. Terje Ogden and John Kjøbli are already mentioned, and Marion Forgatch, thanks for your quick responses and also for your supportive and encouraging feedback. You are an inspiration.

I am very grateful for the day to day support, research feedback, and numerous shared jokes between me and my fellow colleges in the research department; Agathe Backer Grøndahl, Ane Nærde, Gunnar Bjørnebekk, Imac Maria Zambrana, Ivar Frønes, Kristin Berg Nordahl, Mari-Anne Sørlie, Henrik Daae Zachrisson, Silje Hukkelberg, Thormod Idsøe, Luisa Antunes Ribeiro, Serap Keles, and Kristine Amlund Hagen, you have all been of priceless value to my work. Without your research knowledge and experience my Ph.D.-thesis and I would have suffered. Not to forget my equally important colleges in the research logistics team; Gørel Eriksson Bringedal, Asgeir Olseth, Bjørn Arild (BA) Kristiansen, Sivarajan Rajah, and Kristian Rognstad. Moreover, I have to thank you in the administrative department of the center and for your excellent day to day efforts to prepare the ground for all our work;

Inger-Hege Utgarden, Rukiye Tantar Akturk, Thomas Åker-Furre, Astri Vannebo, and least but not last Øivind Magnus (support) Hoff.

This Ph.D.-thesis is also largely influenced by you skilled clinicians, practitioners, and implementers in; the Department of Child Program Development, the National Implementation Team, and you TIBIR practitioners. You all have inspired the work in this thesis - no names mentioned, no one forgotten. And not to forget the children and families receiving help in TIBIR; it is for you that we are here. Thus, if this thesis can contribute to inform and improve the services for vulnerable children, my most important goal in this thesis will be accomplished.

Finally, my greatest debts are to my closest family Marie, Marte, and Mikkel for always being there to provide inspiration, frustration, support, and relief. And thank you for putting up with my absent-mindedness, particularly in the final stages of this thesis. I also want thank to my family, Solveig and Gjert for providing the genes and good enough parenting for my sister Lisa and me. Thanks to my friends; (Lille-) Kristian Heggebø for the long, late, and dozier sociological conversations; Espen, Rune, Øyvin and (Store-) Kristian for distracting me; and to the rest of you (you know who you are), thanks for putting up with my increasingly nerdier microsocial interactions.

Table of contents

List of papers

Paper I:

Tømmeraas, T. (2016). Social gradients and participant characteristics in child behavior problem interventions, *Children and Youth Services Review, 70*, pp. 57-64.

Paper II:

Tømmerås, T. and Kjøbli, J. (2017). Family resources and effects on child behavior problem interventions: A cumulative risk approach. *Journal of Child and Family Studies, 70* (26), pp. 2936-2947.

Paper III:

Tømmerås, T., Kjøbli, J., and Forgatch, M.S. (2017). Collateral benefits of interventions for child behavior problems on parent wellbeing, under revision for *Family Relations.*

Paper IV:

Tømmeraas, T. and Ogden, T. (2015). Is there a scale-up penalty? Testing behavioral change in the scaling up of Parent Management Training in Norway, *Administration and Policy in Mental health and Mental Health Services Research, 44* (203), pp. 203-216.

Summary

Children displaying early behavior problems are at risk of experiencing problems in education system and often face poor life chances later in life. As with other mental health problems, there is a social gradient in child behavior problems: Lacks of social and economic resources in the family are social risks in children's family environment. Providing effective help for this group of children and families is important. Accordingly, this thesis evaluates evidence-based parent training interventions targeting child behavior problems. The overall topic concerns health care disparities in service utilization and treatment outcomes, whether evidence-based parent training interventions may exacerbate or ameliorate social disparities in health care by being less used and having less beneficial outcomes for the low-resource populations.

The first paper addresses key dimensions in health care disparity research; service utilization. First, whether behavior problem interventions reaches and serves children from low-resource backgrounds. Second, whether there are social gradients relating to treatment intensity, if the low-resource families are less likely to use high-intensive parent training. Overall, I find that children and families who utilize evidence-based parent training offered in the Norwegian municipal service level have low levels of social and economic resources compared to the Norwegian normal population of families with children. Second, I find that low-resource background predict utilization of the more intensive parent training intervention. However, the results showed that the most disadvantaged families, having more than three cumulative family risks, were less likely to receive the high-intensive treatment.

The second paper addresses another key dimension in health care disparities, namely whether children from low-resource backgrounds have less beneficial outcomes in treatment. Firs, we examined whether evidence-based parent training interventions generally contributed to health care disparities by producing less behavioral change for the children

from low-resource backgrounds. Second, we investigated whether receiving low or high-intensive parent training had differential impact on children`s outcome. Results showed that evidence-based parent training interventions overall ameliorated health care disparities by being more effective for children from low-resource families. Moreover, results that the high-intensive intervention was particularly effective for the low-resource children.

In the third paper, we examined whether parent training interventions could reduce social risk by having collateral benefits on parent wellbeing, i.e. somatic health status, mental health status, and feeling of vitality. We found collateral benefits in the low-intensive parent training intervention that thus largely included lower risk participants. However, we did not find collateral benefits in the high-intensive parent training intervention when compared to regular care. Accordingly, the low-intensive parent training showed promising results by reducing children`s exposure to social risks six months after treatment termination.

Paper four marks a shift from the child and family oriented perspectives, and has a focus on parent training outcomes in different phases of implementation. The question was whether treatment outcomes attenuated in the large-scale implementation phase when the parent training intervention was disseminated across the whole service system intended to deliver intervention. Despite an increasing heterogeneity among service providers and target group, we did not find any indications of reductions in parent training treatment effects in the large-scale implementation phase.

1. Introduction

Mental health problems, and more specifically behavior problems such as inattentive behavior, conduct problems, oppositional problems, and antisocial behavior, affect both children's emotional states and social relationships at present and their life chances; (Cuellar, 2015; Patterson, Forgatch, & DeGarmo, 2010). In fact, childhood behavior problems has been found to harm long-term development by having consequences for school readiness and academic underachievement, work problems, criminal behavior, and poor health later in life (McLeod & Kaiser, 2004; Moffitt, Caspi, Harrington, & Milne, 2002; O'Connor, Dearing, & Collins, 2011). Early childhood conditions, and particularly the family environment, have implications for children´s level of behavior problems. Lacks of social and economic resources in the family have been found both to increase the likelihood for and intensify the development of behavior problems: There is a social gradient in behavior problems (Aneshensel, Phelan, & Bierman, 2013; Mazza et al., 2016; Piotrowska, Stride, Croft, & Rowe, 2015). Experiencing early behavior problems may hamper children from realizing their human potential (Aneshensel et al., 2013; National Academies of Sciences & Medicine, 2017). As a result, childhood behavior problems are costly to individuals, their family, and society at large (Patterson, 1996; Scott, Knapp, Henderson, & Maughan, 2001). To prevent and treat such problems, evidence-based parent training interventions have been developed and implemented since the 1960s (Forgatch & Kjøbli, 2016; Kaehler, Jacobs, & Jones, 2016).

Behavior problems are considered to be part of the mental health field (Cuellar, 2015). However, the dominant source of intervention is not within the traditional medical care. Instead, primary services in schools, child welfare, child protection services, and community health clinics are the dominant sources that deliver care for behavior problems (Askeland, Solholm, & Apeland, 2014; Burns et al., 1995). Somewhere between 3% and 5% of Norwegian children have serious behavior problems (Skogen & Torvik, 2013). If we include

children in sub-diagnostic levels who also struggle with the behavior problem consequences, the rates would probably be a lot higher. Accordingly, young children's behavior problems place a large burden upon social and health sectors; behavior problems have been identified as one of the most frequent reasons for children's referral to mental health services in Norway (Krogh & Kvello Bukten, 2013). Moreover, disrupted parenting style, for example lack of positive involvement and use of harsh and inconsistent discipline, was reported to be the most frequent reason for receiving intervention in the Norwegian child protection services (Kristoffersen, 2017).

At the end of the nineties, the Norwegian government enacted a policy to prevent and treat the negative consequences of childhood behavior problems by implementing evidence-based parent training interventions (Ogden, Forgatch, Askeland, Patterson, & Bullock, 2005). This thesis concerns the evaluation of these parent training interventions. The point of departure in this thesis is to apply a sociological health care disparity perspective on the evaluation of evidence-based interventions (EBI). Accordingly, this thesis adds to the sociological literature on mental health by focusing on health care disparities in care for children with behavior problems.

1.1 Health care disparities

Research on social disparities in mental health has long traditions (Aneshensel et al., 2013), and the sociological interest in mental health can be traced back to Durkheim's "Suicide" (1951). Social disparities refer to the unequal distribution of social and economic resources (McLeod, 2013). In line with the growing awareness on mental health problems, there has been a mounting sociological interest in social disparities in mental health care (Pescosolido, Boyer, & Medina, 2013). Although there are many different strands of mental health care research (Aneshensel et al., 2013), the sociological interest is often rooted in social disparities

and how unequal distribution of resources may be social determinants involved in processes that have dysfunctional consequences for individuals and the health care system. Regarding behavior problems, sociological scholars have often focused attention towards the pathways between family resources, parenting style and family stress processes in behavior problems development (Conger et al., 1992; Hanson, McLanahan, & Thomson, 1995; McLeod & Shanahan, 1993), and the long-term social consequences of having early behavior problems (Cuellar, 2015; Elder, Downey, & Cross, 1986; Evensen, Lyngstad, Melkevik, & Mykletun, 2016; McLeod & Almazan, 2003). In health care disparity research, central aspects are the social disparities and processes related to service utilization and outcomes in treatment (more extensively reviewed in section 2.2; Spencer & Grace, 2016).

There is consensus that health care disparities exist (Ceci & Papierno, 2005). However, little is known about the health care disparities in evidence-based parent training in Norway. When we know that there is a social gradient found in behavior problems, it would be concerning if the services and EBI were less beneficial for this vulnerable group of children and families. Accordingly, this thesis applies a focus on health care disparities in which the emphasis is directed towards the social disparities in the encounter between the families and the Norwegian service system for children displaying behavior problems. Thus, important topics in this thesis concerns; who utilizes the services; how do they benefit; how do type of treatment have implications for users; and what social processes may generate health care disparities.

1.2 A social policy to reduce health care disparities

In 2007, the Norwegian government issued a white paper (Government, 2007) on strategies to promote health equity. Prevention of health care disparities was listed as one of the strategies to promote health equity. It states that social policy should promote health care

equity: health care intervention should reduce health disparities without any groups having poorer health. If there are health care disparities in service utilization and outcomes of behavior problem interventions, we do not only fail to help a high-risk group of children and their families, but the social policy to help children would probably increase health disparities between advantaged and disadvantaged families. This is probably why providing effective help to this vulnerable and difficult-to-reach group of children is considered particularly important (Leijten, Raaijmakers, Orobio de Castro, van den Ban, & Matthys, 2015).

Recently, there have been calls for interventions that target problems as early in childhood as possible to more effectively help those in need (Heckman, 2006; Melhuish, 2011). According to Heckman, when you target developmental outcomes in early childhood and focus on those at greatest risk of long-term negative development, this would give greater returns to individuals and society. The Heckman tenet skill begets skills (2006), and his theory of human skill formation and returns from investment underpins this argument: When interventions target problems in early stages of development, less effort may be needed to produce favorable outcomes. Somewhat simplified, there are two kinds of interventions that are important to promote equity in health; universal intervention and selected intervention (Giæver, 2013). Universal interventions are considered as the primary objective to accomplish health equity (Government, 2007). In that regard, universal interventions, for example high quality childcare, a good public school system, and policies aimed at income and wealth redistribution, may limit the number of children ending up in high-risk positions. Universal interventions provide the opportunity to target the underlying risk factors for health problems. However, and for multiple reasons, universal interventions will not prohibit all children from obtaining high-risk positions. Selected interventions concerns the interventions that target individuals identified to be in risk positions. Evidence-based parent training aimed

at children`s behavior problems concerns the latter form of intervention, selected intervention targeting children at risk of negative development.

1.3 Evidence-based intervention and evaluation science

Since the 1970s, the concept of evidence-based intervention (EBI) has gained popularity in the biomedical and behavioral sciences (Kristiansen & Mooney, 2004). However, testing of intervention has longstanding historical roots. In 1747 Dr. Lind conducted the first known clinical experiment (Dunn, 1997). Due to the long time spent at sea, sailors tended to get scurvy. Dr. Lind randomly allocated 12 men into two groups and provided them with different dietary supplements. As a result, he could argue that the citrus fruits had a strong effect on reducing scurvy; he had created an effective intervention. Inventing what might have been the first EBI partly by chance; modern accounts of EBI depart from Dr. Lind's in that there are basic prerequisites related to testing of EBI. Interventions should have a theoretical base, a theory of change, EBIs needs to be adequately described, and the EBIs have to be rigorously implemented. The two latter prerequisites are particularly important in behavioral interventions that often involve complex contents and delivery settings (described in more detail in section 3.3). If the EBI content is not adequately described and implemented, it is impossible to know what is being evaluated. The concept of EBI has spread from medicine to other social sciences and practice fields such as psychology, education, nursing, and social work, and at present, the EBI concept has been gaining popularity in welfare policy.

The word evidence connotes proof in everyday language, but in science this concept is related to standards of scientific arguments. To a large degree, international criteria for EBI is based upon two leading American milieus, Blueprints for Healthy Youth Development (Mihalic, Irwin, Elliott, Fagan, & Hansen, 2004), and Society for Prevention Research (Flay

et al., 2005). In Norway, the Ungsinn database provides an overview of Norwegian criteria and EBIs (Eng, Lauritzen, Reedtz, Mørch, & Martinussen, 2014). Many of the criteria for EBI are overlapping and can be summed up as: Interventions should have a theoretical foundation, the effects of interventions should lead to positive change, interventions must have sustaining effects, scientific results are time and context-dependent, and thus, have to be replicated in different populations, cultural settings, and over time. Based on the above-mentioned criteria, an EBI is classified cumulatively and hierarchically according to the magnitude of, and the standard in, the available research associated with a particular intervention. Conducting the scurvy experiment, Dr. Lind did probably not have standards of criteria in mind, but he was the first known to develop research evidence using what today is called a randomized controlled trial (RCT). Today, the RCT design is considered as the "gold standard" in EBI evaluation science (Weisz & Kazdin, 2010). However, both the RCT design and EBIs are debated and certainly not without flaws and weaknesses (more extenesively reviewed in section 4; Berk, 2005; Cartwright, 2007; Weisz et al., 2013).

1.4 The aims of this thesis

In this thesis, we evaluate three versions of the Parent Management Training – Oregon model (PMTO) that are part of a comprehensive intervention program for children with behavior problems called TIBIR. TIBIR is a Norwegian acronym meaning Early Initiatives for Children at Risk (Tidlig Innsats for Barn i Risiko). Accordingly, this thesis has two interwoven contributions: The thesis adds to the sociological literature on mental health care disparities, and provides evaluations of parent training interventions for children with behavior problems. Thus, the sociological theoretical focus is merged with a practical focus on specific EBIs, where the goal is to derive new knowledge that contributes both to theory and practice. In other words, by using a theoretical focus on health care disparities, the goal is

to derive new knowledge whether evidence-based parent training may serve as an efficient way of addressing social policy goals to provide effective help to underserved low-resource populations. To a certain extent, this thesis bridges social policy with individual data.

The tenet, what works for whom under what conditions, is central in EBI and in this thesis. Research has shown that the PMTO based parent training interventions in TIBIR work better than the alternative of receiving regular care in the Norwegian service system (Bjørknes & Manger, 2013; Hagen, Ogden, & Bjørnebekk, 2011; Kjøbli & Bjørnebekk, 2013; Kjøbli, Hukkelberg, & Ogden, 2013; Kjøbli & Ogden, 2012; Ogden & Hagen, 2008). However, we know less about for whom and under what conditions the parent training interventions work. Accordingly, this thesis will - throughout three empirical papers – investigate different aspects of health care disparities related family resources. In paper I, the focus is on health care disparities among those who utilize TIBIR parent training interventions. Paper II examines health care disparities in the outcomes (i.e. child behavior change) of TIBIR parent training interventions, whether cumulative aspects of family resources moderate treatment effects. Paper III investigates whether targeting child behavior problems may have positive effects on the family environment by improving parents' wellbeing. Paper IV marks a shift from the family oriented approaches to a focus on disparities within the service system itself, whether parent training works equally well when implemented large-scale in multiple service institutions compared to initial and more stringent effectiveness testing in the specialist services for children with behavior problems.

The remainder of this introductory chapter is organized as follows. In Chapter 2, I elaborate on the concept of behavior problems including the family processes involved, before presenting more theory and research on health care disparities. In Chapter 3, I describe in detail the case study: the TIBIR parent training interventions and their implementation. Chapter 4 contains a description of the statistical methods used in the empirical papers and

discussions and critique directed towards the RCT design and the use behavioral EBIs in mental health care. Chapter 5 gives a brief description of the content in the four empirical papers. And finally, in Chapter 6, the findings will be discussed in terms of their implications for theory and practice.

2. Theory and previous research

2.1 Behavior problems – Consequences, risk factors, and family processes

Behavior problems are situated within a societal context and have long-term consequences for individuals, families, communities, and society at large. Accordingly, this section starts with an emphasis on some of the associated consequences of displaying early behavior problems before the focus is directed towards behavior problems risk factors and social processes in the family.

2.1.1 Consequences of behavioral problem

In advanced economies, there is a high need for human skills and knowledge in professional life (Frønes, 2016). Accordingly, modern childhood from infancy to adulthood is a long and demanding period of development where children spend many years in the education system (McLeod & Almazan, 2003). Forecasting future labor market demands in Norway, Bjørnstad et al. (2010) estimated that there would be a steady decline in unskilled and manual labor opportunities towards 2030. Children displaying early behavior problems are particularly vulnerable in this context as they more often fall behind in education and have limited opportunities as they emerge into adulthood (Frønes, 2016). As a result, there is reason to believe that the negative consequences of having early behavior problems would not diminish in the near future.

The high demand for human skills in our society is probably the reason why school dropout is a major problem in Norway today. Although definitions have been debated (Vogt, 2017), seven per cent of all young Norwegians between the ages of 16 and 24 have been found not to be in education, work, or training, and this number has been increasing in the last decades (Digre & Haugberg, 2016). Moreover, Markussen and Røed (2017) found that young adults from low resource backgrounds are increasingly falling behind in terms of

education and earnings. Children with behavior problems would typically be found in these numbers. For instance, early behavior problems have been found to affect academic problems in elementary school (O'Connor et al., 2011). Extending to later stages of education, behavior problems have been associated with poor educational outcomes in adolescence (McLeod & Kaiser, 2004; Sayal, Washbrook, & Propper, 2015) and in adulthood (Evensen et al., 2016). In fact, early behavior problems have been found to be an important predictor for later school dropout both internationally (Breslau, 2010) and in Norway (Sagatun, Heyerdahl, Wentzel-Larsen, & Lien, 2014). Accordingly, early behavior problems have been found to predict low earnings, work problems, poor health, and criminal behavior (Moffitt et al., 2011; Patterson, 1996; Robins, 1966). The reason why children with behavior problems have problems in the education system are probably rooted in children's ability for self-control and their capability to sit still and concentrate over long periods of time. Moreover, and opposed to internalizing problems, children with behavior problems are often a burden for teachers and other students as their externalizing behaviors harm the teaching environment.

The aphorism "Child is father of the man" (Caspi et al., 2016, p. 1) pinpoints the continuity from childhood problems to adult outcomes. Some of the children who display early behavior problems in the family environment are at risk of following disadvantaged developmental trajectories, starting with problems in school that prepare the ground for later marginalization. Understanding the complete picture of developmental trajectories is challenging because the relations between behavior problems and poor outcomes later in life are complicated, often involving comorbid problems such as internalizing problems, lower social competence, deficit hyperactivity disorder, and cognitive impairment (Galambos, Barker, & Almeida, 2003; McConaughy & Skiba, 1993; Patterson et al., 2010).

2.1.2 Behavior problems and associated risk factors

It is well established that there is social disparities in health (Marmot & Wilkinson, 2005). Likewise, it has repeatedly been demonstrated that mental disorders are unevenly distributed throughout society; they tend to cluster within lower social strata (Aneshensel et al., 2013). This finding is also relevant for behavior problems: children from families with low amounts of social and economic resources are more likely to develop behavior problems compared to their high-resource counterparts (McLeod & Shanahan, 1996; Piotrowska et al., 2015). Although individual risk (or biology) accounts for some of these relations, much of it can be traced back to children's social environment, social risk (Haskins, Garfinkel, & McLanahan, 2014). Here, a risk factor refers to correlates that precede development of behavior problems. There is a large body of research that connects development of behavior problems to risk factors in the children themselves, their immediate environment, and society at large (Caspi et al., 2016; Costello & Angold, 2001; Moffitt & Scott, 2009; Sroufe, Egeland, Carlson, & Collins, 2009).

Starting with the most proximal, individual risk refers to properties often innate or developed at a prenatal or early stage in life. Several individual characteristics, for example neurobiological disruptions, temperament, and aggression, have been associated with behavior problems (Deater–Deckard, Dodge, Bates, & Pettit, 1998; Moffitt & Scott, 2009). Social risk refers to a broad range of factors in children's close and distal surroundings together with process related factors. Examples of social risks are; lack of social and economic resources in the family, peer rejection and antisocial peers, poor neighborhood context, cultural context, societal context, negative parenting style, and family stress processes (Dishion & Patterson, 2006; Moffitt & Scott, 2009; Murali & Oyebode, 2004; Piotrowska et al., 2015; Sroufe et al., 2009). There is also Norwegian research showing that a lack of resources is related to behavior problems. For example, Bøe et al. (2012) and

Wichstrøm et al. (2012) have documented how social risk predicts a higher likelihood of experiencing behavior problems. Taken together, the relation between social risk and behavior problems seem robust. In addition, the social risk factors, such as low amounts of family access to resources, are associated with problems in other health domains which place children from disadvantaged backgrounds more broadly at risk of negative development (Shonkoff et al., 2012).

A contribution from the sociology of mental health has been to recognize that vulnerable families often are exposed to a broad array of risks (Turner, Wheaton, & Lloyd, 1995; Wheaton, 1994). Evidence from The Fragile Families and Child Wellbeing Longitudinal Study confirms the dynamic relation between social risks (Waldfogel, Craigie, & Brooks-Gunn, 2010). For instance, single parent families have been found to have fewer social and economic resources available; caregivers are more often poor, have lower education levels, as well as more somatic and mental health problems (Kalil & Ryan, 2010; Meadows, McLanahan, & Knab, 2009). Thus, social risk factors are correlated and the consequences of several risk factors operating together are probably stronger than the sum of effects of single risk factors considered independently of each other (Frønes & Strømme, 2014). This implies that some families may accumulate several social risks with the results of an intensified likelihood of behavior problems development. In that regard, social risk and behavior problems may both be parts in a process where families accumulate disadvantages, often termed as cumulative disadvantage (DiPrete & Eirich, 2006; Seabrook & Avison, 2012).

In psychology, a related term is often used to conceptualize exposure to multiple disadvantages, cumulative risk (Evans, Li, & Whipple, 2013). Cumulative risk differs from the sociological concept of cumulative disadvantage because it does not focus on the accumulation process of disadvantages. The concept of cumulative risk is more often used to measure quantitative exposure to several risks simultaneously and the implications

cumulative risk has for psychological outcomes, including behavior problems. Research has shown that children who are exposed to family environments with cumulative risk factors – in effect children from the most disadvantaged backgrounds - are more likely to experience behavior problems (Evans et al., 2013; Sameroff, Bartko, Baldwin, Baldwin, & Seifer, 1998; Trentacosta et al., 2008). As with cumulative disadvantage, it seems likely that cumulative risk exposure may have implications for both mental health and health care disparities.

2.1.3 Behavior problems - Social risk, family stress, and disrupted parenting style

It is well established among scholars from sociology and psychology that the family environment is a crucial factor in early behavior problems (Ge, Conger, Lorenz, Shanahan, & Elder Jr, 1995; Goodman & Gotlib, 2002). Thus, an important question is; what are the mechanisms that connect social risks and the family environment to behavior problems in children.

Many of the processes that link social risk and behavior problems in children are related to parents' wellbeing and family stress (Conger, Conger, & Martin, 2010). In his classic study on the consequences of the great depression, Elder et al., (1986) found that social risk in the form of economic problems created parental stress which in turn were followed by disrupted parenting style, and finally, behavior problems in children. In a meta-review, Conger et al. (2010) summed up the relations between social risk, stress and behavior problems in children; low access to resources affected family stress which in turn harmed parenting style, and thus increased the likelihood for a broad range of developmental problems in children, including behavior problems.

The evidence supporting the social risk and the family stress hypothesis seems robust across culture and ethnic groups (Benner & Kim, 2010; Conger et al., 2002; Parke et al., 2004; Solantaus, Leinonen, & Punamäki, 2004). Moreover, in addition to economic hardships,

associations between social risk, stress and child behavior problems involve other parental characteristics such as parental depression (Bank, Forgatch, Patterson, & Fetrow, 1993; Glied & Oellerich, 2014), and family structure and instability (Waldfogel et al., 2010). Moreover, in support of the stress hypothesis, social risk, in the form of low access to resources in the family, has been found to increase the likelihood of child exposure to harsh and inconsistent parenting style (Bank et al., 1993; Elstad & Stefansen, 2014; Glied & Oellerich, 2014; Rishel, 2012).

In keeping with the social risk and stress hypothesis, Gerald Patterson's Social Interaction Learning (SIL) model expands on social risk, family stress and behavior problem development by emphasizing the microsocial interactions that lie beneath the stress processes and disrupted parenting style (termed as parenting practices in SIL). Thus, the SIL model complements the family stress hypothesis by highlighting how microsocial coercive interactions in the family lead to development of child aggression (Patterson, 2002). Coercive interaction is characterized by conflict and emotional escalation. Typically, children get tantrums that are followed by emotional escalation in parents and children. In turn, parents tend to concede to meet children's demands. This creates negative reinforcement of aggressive and antisocial behavior in children (Patterson, 2002). Emotional escalation, withdrawal, and negative harsh and inconsistent parenting style are parents' contribution the coercive interaction process. In line with the family stress hypothesis, the SIL model explains how family and parent characteristics such as poverty, low parental education, single parenting, and parental mental and somatic problems are contextual factors that harm parenting style and create coercive interactions between family members.

The SIL model and the coercive mechanism highlight the reciprocity in child and parent interaction: children and parents reinforce each other's negative behavior, which thus has implications for the whole family system (Patterson, 2002). Similarly, Avison and

Comeau (2013) noted that children's mental health problems have consequences for other family members. Therefore, experiencing a child who has behavior problem is likely a factor that adds to the family stress load and to the accumulation of family disadvantage. For example, parents who have a child displaying behavior problems tend to experience more social isolation from family and friends (Patterson, 2002). As noted by Patterson and colleges (Patterson, 2002; Patterson et al., 2010), reducing behavior problems may have positive consequences for the involved child but also for their parents and other family members (Patterson, 2002; Patterson et al., 2010). And importantly, SIL model offers a social solution to the problems: reduction of negative parenting style and teaching of positive and effective parenting style.

2.2 Health care disparities

Prevention of health care disparities is one of the means to counteract health disparities (National Academies of Sciences & Medicine, 2017). Thus, the prevention of health care disparities rests both on moral arguments and issues connected to the efficiency of the health care system. This implies that if health care disparities exist, interventions fail to help an underserved population of disadvantaged families. Moreover, for interventions to be efficient, they must be beneficial for all risk populations, which, in the case of behavior problems, include children from families marked by social risk. It has been argued that mental health care interventions may intensify health disparities in two ways. Two different disparity-increasing mechanisms have been proposed: (i) by disproportionately serving high-resource populations with health care, and (ii), the interventions may work better for clients from high-resource backgrounds (Ceci & Papierno, 2005). As such, important elements in health care disparities are related to the social disparities in service utilization and outcomes in treatment.

2.2.1 Social disparities in service utilization and outcomes in treatment

The health care system itself has been characterized as a determinant of health (Marmot et al., 2008). Thus, extensive research has focused on how health care systems exacerbate health disparities (Spencer & Grace, 2016). Implicit in much of this research is that people will have improvements in their health if they utilize and benefit from health care. The impact of social and economic resources on health care disparities is well recognized (Marmot et al., 2008) and have been closely linked to: (i) access to care and service utilization, (ii) experiences of care, and (iii) treatment outcomes or benefits from care (Alegría, Pescosolido, Williams, & Canino, 2011). The effect of these health care disparities generating domains on individuals takes place in a complex ecological system spanning from the (welfare) state and health policy, via the functioning of the health care services, to more micro-level factors such as family dynamics, social support and patient-provider communication (Alegría et al., 2011; Spencer & Grace, 2016).

In this thesis, the focus on health care disparities relates essentially to the service utilization and the treatment outcomes of interventions. First, results from several meta-analyses has backed up the general consensus that there are social gradients in the outcomes of parent training interventions, meaning that such interventions seem to work better for children from high-resource backgrounds (Leijten, Raaijmakers, de Castro, & Matthys, 2013; Lundahl, Risser, & Lovejoy, 2006; Reyno & McGrath, 2006). However, others have pointed out that the social gradient in parent training outcomes are somewhat mixed (Deković et al., 2011). Second, several scholars have documented social gradients in utilization of evidence-based treatments for behavior problems, meaning that children from low-resource backgrounds are less likely to utilize help services (Bussing, Zima, Gary, & Garvan, 2003; Haggerty et al., 2002; Kazdin, Holland, & Crowley, 1997; Pettersson, Lindén-Boström, & Eriksson, 2009; Reedtz, Martinussen, Jørgensen, Handegård, & Mørch, 2011). Taken

together, if there are social disparities in both service utilization and the treatment outcomes of evidence-based parent training interventions, the effectiveness of the interventions and adequacy of service provision would be compromised.

The generalizability of the effects of family resources on both service utilization and treatment outcomes implies that there may be common mechanisms of change relating to the two health care disparity domains. In that regard, it has been suggested that socioeconomic resources is a fundamental determinant of health that structures individuals access to flexible assets (or resources) to avoid risks and to cope with health problems (Link & Phelan, 2005). Applied to behavior problems care, this suggests that low-resource families may lack several flexible assets that in turn may produce health care disparities. Hence, lacks of family resources are likely proxies for several types of change mechanisms that may impact on both service utilization and treatment outcomes.

Different classes of mechanisms might be at play in health care utilization and beneficial outcomes in EBI. For instance, socially graded network mechanisms, which offers beneficial support including, information, social influence, and advice, might produce health care disparities (Smith & Christakis, 2008; Thoits, 2011). Sociocultural mechanisms may be another class of mechanisms that may create health care disparities. For instance, factors such as attitudes towards help services may create health care disparities if low-resource families norms, values and practices may adhere less to those communicated in mental health services (Gillies, 2006), and parent training may less closely match the realities of low-resource families (Zilberstein, 2016). Moreover, low-resource parents have been found to behave in a less confident way when interacting with professionals. This may lead to less beneficial outcomes in treatment and probably also avoidance from utilizing services faced with certain need in their child (Gengler, 2014; Lareau, 2011; Weininger & Lareau, 2003). Finally, practical mechanisms associated with low access to family resources may also increase health

care disparities (Cuellar, 2015). For example, low-resource families more often consist of poor single parent households with several children. Thus, low-resource families are more often likely to face practical barriers such as money to pay for transport and childcare, and time to practice skills learned in parent training between sessions, which both could limit their service utilization and their potential to benefit from intervention.

3. TIBIR - Parent training interventions and implementation

3.1 The TIBIR program

To identify children at risk and to provide them with effective interventions, Norwegian government has invested in the development and implementation of the evidence-based TIBIR program. TIBIR targets children between 3 and 12 years of age. The complete TIBIR program contains six interventions: a screening intervention, a brief parent training intervention (BPT), individual and group versions of PMTO, a teacher training intervention, and a children's social skills training intervention (Solholm, Kjøbli, & Christiansen, 2013). TIBIR is a community-wide model that is designed to target children's behavior problems systematically and tailored to different risk levels in the main social arenas for children (i.e. home, day care, or school; Solholm et al., 2013). Accordingly, TIBIR is implemented across service sectors essentially in the municipal service level to be close to the target populations (Solholm et al., 2013).

In this thesis, the focus is on the parent training interventions that target behavior problems in the home environment. In parent training interventions, parents are supposed to be the agents of change in their children (Kaehler et al., 2016). Thus, in parent training, parents are given the "responsibility" for behavior change in their children, which avoids problem focus and stigma on children themselves. Evidence-based parent training is considered to be one of the most effective ways to target child behavior problems (Edwards, Céilleachair, Bywater, Hughes, & Hutchings, 2007; Greenwood, 1998; Serketich & Dumas, 1996). Accordingly, strengthening the family environment through parent training interventions is the cornerstone in the TIBIR program.

3.2 Parent training interventions in TIBIR

The TIBIR parent training interventions are based on the American PMTO program. Gerald Patterson, Marion Forgatch, and their colleges at the Oregon Social Learning Center, have developed PMTO over five decades of research and testing (Forgatch & Kjøbli, 2016). PMTO is theoretically based on the SIL model. Accordingly, the TIBIR parent training interventions aim at reducing behavior problems and promoting prosocial behavior by reducing negative parenting style and replacing it with positive and effective parenting style; with the ultimate goal of reducing family coercive interactions (Forgatch & Patterson, 2010).

The high-intensive PMTO interventions in TIBIR are delivered with a dosage of approximately 25 hours in the individual mode and 30 hours in the group mode (Solholm et al., 2013). The PMTO interventions target children of moderate to high risk of developing behavior problems (also referred to as selected and indicated level of risk). PMTO includes five parenting skill core components; positive involvement; praise and encouragement; problem solving; effective and consistent discipline; and monitoring. PMTO also contains five additional components to support positive family interaction; good directions; emotional regulation; and screening of child and parent behavior, together with a school or child care component to promote continuity and support of PMTO content in other arenas (Askeland et al., 2014). Provision of the core components is fairly fixed, but the therapy sessions and progress are customized to fit each individual family. Accordingly, the therapy starts with a screening of the family in which the therapy is anchored in the strengths and challenges in each family. The course of PMTO therapy always starts with the positive parenting components praise and reward and good directions before advancing the other components (Askeland et al., 2014). PMTO therapists uses several tools to effectively teach the parenting skills such as role play, homework, telephone contact between sessions, repetition and rehearsals of the parenting skills, to mention a few.

Brief Parent Training (BPT) is a short-form of PMTO, approximately five 1- hour sessions, and has the same overall aim as in the PMTO interventions. However, BPT contains a reduced version of the curriculum and components used in PMTO. BPT is a preventive intervention that targets children between low and moderate risk of developing behavior problems. In BPT, parents are taught the most important parenting skills, and the counselor customizes the content in relation to family needs (Askeland et al., 2014). Components used include good directions, praise and reward, effective discipline, and problem solving. Compared to PMTO, there is less room for rehearsal of components in BPT. In cases where BPT counseling proves to be insufficient, families are referred to more intensive therapy, for instance PMTO.

3.3 Implementation of TIBIR

Since the 1970s, the field of implementation science has gradually evolved, and today implementation and EBI are inseparable concepts. One of the main goals in implementation science is to bridge science into practice, and it has been defined as the study of methods to promote the systematic uptake of research findings and EBI into routine practice (see Ogden & Fixsen, 2014, p. 4).

There has been a growing awareness that development of EBI does not automatically come to benefit wider target populations. To give an example; in the case of a medical drug treatment, the implementation of new innovations might be straightforward; produce the drug and disseminate it to health care professionals who can effectively provide the drug. However, implementation may be a lot more complicated in the field of mental health and behavioral interventions, where the treatment consist of practitioners who must adhere to complex intervention components combined with mastery of therapeutic common factors. Adding up to these challenges is the fact that target groups often consist of individuals with complex

problems. The heterogeneity in target groups is often accompanied by heterogeneity among EBI practitioners; they are often highly varied in terms of background training, experience, and work place in large-scale implementation. Accordingly, the implementation of an EBI needs to be rigorously addressed if an EBI is to be effective in real world settings. Mathematically expressed, an effective intervention (x) multiplied with effective implementation (y) equals (=) social significant outcomes. Removing either of the variables from the equation, setting the x or y to zero equals (=) non-significant outcomes (0). As a consequence, several scholars have focused on what are the facilitators and obstacles in effective implementation (Fixsen, Blase, Naoom, & Wallace, 2009; Ogden & Fixsen, 2014; Welsh, Sullivan, & Olds, 2010).

In a health care system, the facilitators and obstacles relates to multiple implementation drivers (or components) in multiple service levels such as the organizational level, leadership level, and practitioner level (Fixsen et al., 2009), and to the feasibility of the EBI in the service system. In their meta-review, Fixsen and Blasé (2009) highlighted several core implementation drivers that were important to provide effective implementation support; training of practitioners, evaluation of staff performance, on-site coaching, and facilitative administrative support, to mention a few.

Addressing implementation drivers and provision of EBI support and quality control is thus important in the dissemination and sustainability of intervention effects in real world settings. Norwegian government has financed the Norwegian Center for Child Behavioral Development to administer the implementation of TIBIR interventions in Norway (Ogden et al., 2005). To uphold program effects and sustainability within diverse service settings, a great deal of effort is invested the implementation support and quality control such as facilitative administrative and leadership structures, time to practice interventions, on-site coaching and local collaborating teams of practitioners, data based quality control system for

treatment effects and EBI adherence, regular recertification of PMTO therapists, and minimum case load requirements (For a more thorough review of the TIBIR implementation process see Askeland et al., 2014; Ogden et al., 2005).

4. Methods

After a short presentation of the ethical standards in TIBIR research, this section proceeds with a short description of the study designs before I present the different statistical methods used in the four papers. All of the papers include EBI evaluated in RCT designs. Thus, I will present and discuss some of the strengths and challenges in the RCT design and the use of EBI.

All the studies in the four empirical papers comply with Norwegian and international ethical research standards. Consequently, all procedures in the studies used were in accordance with - and approved by - The Norwegian National Committee for Research ethics, Region South, and The Norwegian Social Data Services. Prior to inclusion in the evaluation studies, participants filled out written informed consents.

There are several similarities across the RCTs used in this thesis. Individuals were randomized in a 50:50 allocation to either the intervention or comparison group. In separate samples, the intervention group received one of the three TIBIR PMTO interventions and the comparison group received the alternative of regular care. Regular care consisted of the help normally provided to children displaying behavior problems in the Norwegian services at that time. To varying degrees, the regular care consisted of active treatments. Overall, the regular care treatments varied a lot in scope and intensity, and often consisted of unstructured counseling or therapy supplied by professionals in the services system. Importantly, regular care did not receive other EBIs.

4.1 Statistical methods used in the four papers

Following the first PMTO RCT in Norway (Ogden & Hagen, 2008), the US National Institute for Drug Abuse funded a data collection for an implementation study (Forgatch & DeGarmo, 2011). Paper IV is based on data from the quasi-experimental implementation

study combined with data from the Ogden and Hagen (2008) RCT. In paper I, II, and III, we used pooled data from the BPT and the PMTO (group mode) RCTs. Reasons behind the pooling of the data were both driven by the research questions and pragmatic concerns; we were only able to pool data from studies containing identical outcome measures. Since the four papers address different types of research questions, the four papers rely on different analytical strategies and statistical approaches.

In paper 1, group differences regarding the participant characteristics in PMTO (group mode) and BPT were compared to the characteristics in the Norwegian population of families with children. I used *t*-tests and chi-square tests for categorical and binary outcomes. In addition, relations between family resources and the outcome of intensity in treatment (binary outcome) were analyzed using multiple regressions. Conventionally, and due to problems with homoscedasticity and out-of-bounds predictions, logistic regression has been preferred for analyses of binary outcomes in regression analysis. However, this notion has been challenged (Hellevik, 2009; Mood, 2010). Since the main interest in these analyses was the coefficient parameters and not the probabilities, multiple regressions were preferred.

In the moderator analyses in paper II and in the mediator analyses in paper III, we used structured equation modeling (SEM) analyses. The interaction and mediation analyses in SEM were run in Mplus version 7 (Muthén & Muthén, 2012) using latent child and parent outcomes. SEM allows incorporating a measurement model of latent constructs estimated simultaneously with a structural model of regression path analysis based on the estimation of covariance matrixes (Kline, 2015). In SEM, a covariance matrix, in terms of a statistical model specified by the researcher, is compared with a non-specified covariance matrix in the original dataset. As a result, and in addition to testing coefficient parameters, SEM allows for evaluation of the statistical model fit in the analytical models specified by the researcher. In

other words, SEM allows for the statistical evaluation of the appropriateness of theoretical models applied on a sample (Hu & Bentler, 1999).

A feature in SEM is that it includes the opportunity to incorporate a measurement model of latent outcomes that is estimated together with the structural regression paths. Compared to sum scores, analyzing latent constructs have the advantage of minimizing error by only allowing the common variance in the observed indicators to tap the latent construct (Kline, 2015). Thus, statistical noise and error variance, which is not correlated to the specified latent construct, is partialled out from the latent outcome. Due to issues concerning adequate item to sample size ratio, and to prevent the analytic models being just identified, we chose to use parcels in our latent outcomes. Item parceling in SEM involve combining several observed indicators into parcels that tap the latent outcome. The use of parcels has been debated (see for example Little, Cunningham, Shahar, & Widaman, 2002). However, the use of parcels has been considered to be better than the alternative of using observed constructs (or sum scores; Rhemtulla, 2016). Considering the pitfalls of parceling (Little et al., 2002), we based our parcels on theoretically established construct dimensions. In addition, we invariance tested the dimensionality of the items in confirmatory factor analyses across time, treatment condition, and (pooled) studies; with the result of obtaining partial strong invariance across treatment condition groups and studies over time. Moreover, in paper II and III, we allowed for the correlation of the parcel error terms over the time-points used in the autoregressive models (i.e. time point 2 and 3 regressed on time point 1). These unanalyzed associations are a standard way to represent shared sources of variability over and beyond the common variance estimated in the latent outcomes (e.g. error variance due to resampling respondents in several time points; Kline, 2015). Note that the coefficient paths between the independent variables and the latent outcomes were identical with and without correlation of parcel error terms, however, model fit improved in the correlated models. Moreover, in the

paper II and III analyses, we used intent-to-treat analyses in which all participants sampled at time point 1 were part of analyses regardless of whether they received treatment or not, or whether they agreed to be sampled in time point 2 and 3.

In paper IV, we tested group differences in child behavior change between the effectiveness phase and the large-scale phase of implementation using a pre-post design. In the statistical analyses, we used within-subject factorial multiple analysis of covariance (MANCOVA) analyzed in SPSS version 22s' general linear modeling procedure (*F*-test statistics). MANCOVA models allow for the testing of composite outcome models with the advantage of preventing potential type 1 errors (Tabachnick & Fidell, 2001).

4.2 Limitations in the RCT design

The use of RCT data in this thesis has several advantages. However, there are also several potential limitations in a RCT design that needs to be considered. Validity within empirical research is commonly concerned with whether a conclusion or inference represents a good estimate of the true conclusion (Trochim, 2006). Validity in research concerns the cumulativeness of several factors often operationalized as; conclusion validity; construct validity; internal validity; and external validity. The validity in a RCT concerns all four, however, the last two concepts are particularly salient. Internal validity can be defined as the ability to argue that the observed correlations are causal (Roe & Just, 2009). In a RCT, the "magic" of randomization combined with a sufficiently powered experiment will produce high internal validity, which probably is the most important rationale for using a RCT. As a result, researchers have the possibility to make causal inferences about the correlations observed in a specific sample. However, whether a sample specific causal inference generalizes over and beyond a particular RCT sample does not follow automatically from the randomization procedure; it concerns the external validity in a RCT.

External validity can be defined as the ability to generalize the correlations found in research to other persons, times, and settings (Roe & Just, 2009). As stressed by Cartwright and Hardie (2012), the cost of high internal validity in a RCT may come at a cost of limited external validity due to narrowness in scope and sampling. This implies that external validity can only be claimed if the participants in a RCT are a representative group of individuals from the population they were sampled from. As highlighted by Cartwright and Hardie (2012), to draw policy conclusions from one RCT is problematic. Accordingly, results RCTs needs to be replicated in different contexts and time to plead any policy relevant evidence. Therefore, and as mentioned in chapter 1, replication across contexts and time is a build in feature in standardized EBI criteria (Flay et al., 2005; Mihalic et al., 2004; Weisz & Kazdin, 2010).

How can we increase the external validity in RCTs? Representativeness regarding target group and replication of results have been mentioned. However, representativeness also applies to other EBI elements. The matter of external validity calls for a testing of interventions in real world practice settings in effectiveness and large-scale implementation phases of implementation; testing in regular service systems with the regular practitioners, and with regular implementation support intended to support the EBI in practice. (Ogden & Fixsen, 2014; Weisz et al., 2013). Moreover, in many RCTs the control groups may consist of waiting list controls or essentially non-active treatments. In such designs, the RCT is designed to produce favorable outcomes in favor of an EBI. Such RCT is not designed to answer the basic question about whether the EBI should replace regular care (Weisz et al., 2013). Claims about this question can only be made when the comparison groups consist of regular care. Nevertheless, the extent of external validity in a RCT is always a matter of discussion. Weisz et al. (2013) stresses the need to report factors that affects external validity in RCTs.

In the TIBIR PMTO interventions, external validity is addressed by testing the EBIs in regular care settings, using regular care comparison groups, and the RCTs includes data collected in diverse service settings from all Norwegian health regions (Kjøbli et al., 2013; Kjøbli & Ogden, 2012; Ogden & Hagen, 2008). Moreover, there were no additional participant exclusion criteria in the RCTs other than those regularly practiced in the TIBIR parent training interventions. Although there are different versions of the PMTO-based parent training interventions in TIBIR, similar PMTO principles and content has been replicated in three RCTs in Norway. However, sampling size in the RCTs limits to $N = 112$ (PMTO individual mode), $N = 137$ (PMTO group mode), and $N = 216$ (BPT), indicating that the extent of external validity could be debated. Nevertheless, Solholm et al. (2014) have tested the external validity in the PMTO individual mode RCT (Ogden & Hagen, 2008). They found that the RCT participants were a representative sample of families from the Norwegian services. Also paper IV in this thesis relates to the arguments set forth by Cartwright and Hardie (2012), and whether treatment effects from the first PMTO RCT in Norway may generalize to PMTO as part of regular care practices. Results in paper IV indicate that the RCT treatment effects may generalize to PMTO in Norwegian services (paper IV is described in section 5).

There are also other and more technical limitations in the RCT design such as the assumption of "no interference" and the role of random assignment in statistical inference and representation of mechanisms by which the treatment has an impact (Berk, 2005). First, you have the assumption of "no interference" called stable unit treatment value assumption in which treatments provided in the two conditions should not have an impact on each other. In the case of TIBIR parent training intervention, this implies that the implementation of the EBIs do not affect the treatments given in regular care. It is likely that this may have happened in the case of TIBIR, where the PMTO interventions are implemented within the

same municipalities, and sometimes in the same institution as the regular care treatments were given. Thus, there might be a spillover effect from implementation of EBI to the regular care treatments. In real life this may be considered as a benefit, as knowledge about effective EBI components might be adopted in regular care. However, in RCT testing this may lead to downward bias of the treatment estimate for the intervention evaluated. In some cases, there might be a conflict between the no interference assumption and strive to achieve external validity in RCTs.

Second, in post hoc analyses where treatment effects are tested across individuals that differ on covariates, you are running the risk of data snooping. Meaning that if you look closely enough you will eventually find subgroups that differ on the outcome of interest (Berk, 2005). Hence, you may capitalize on idiosyncratic patterns within a sample that do not easily replicate. One way of avoiding such problems may be to rely on theory and previous research to inform the measurement of moderators and mediators in the statistical models. Nevertheless, other scholars have stated that conducting secondary moderator and mediator analyses are considered as sound and important ways to inform the evidence from RCTs (Kraemer, Wilson, Fairburn, & Agras, 2002). However, it should be noted that many RCTs are powered to detect main effects and not subgroup effects. Relatedly, adding covariates to RCT data will also have implications for the statistical inferences in that p-values may be too optimistic and confidence intervals to narrow (Berk, 2005). Statistically, the best solution may be to mount a new experiment based on the inductively discovered results from the first experiment. For example you may conduct and experiment where participants are randomized according to their level of cumulative risk from which you can draw more statistically robust conclusions. Accordingly, the less variation you have in a moderator, the more the analyses may capitalize on chance within a given sample. A solution is to use pooled samples from several studies to enhance both generalizability of treatment main

effects and subgroup effects (Berk, 2005; Bloom, Hill, & Riccio, 2003). Nevertheless, the limitations in the RCT design, and particularly in the secondary subgroup analyses, should be kept in mind when interpreting when interpreting the results in this thesis.

4.3 EBI critique

Today, many scholars argue that the EBI practice should replace regular treatments in everyday clinical care (Bøyum, 2013; Chambless & Hollon, 1998; Satcher, 2000), however, critics disagree (Addis & Waltz, 2002; Ekeland, 2007; Garfield, 1996; Zilberstein, 2016). Critics often argue from a philosophical standpoint that EBI is neo-positivistic and instrumentalist while others argue that EBI is too rigidly manualized to permit personalized treatment. Others argue and that such interventions are the product of the dominant middleclass Western culture disregarding cultural and ethnical diversities (Zilberstein, 2016). Weisz et al. (2013) insightfully notes that for or against EBI may ultimately be an empirical question. In their meta-analysis (2013), they found that EBI outperformed regular care treatments with a Cohen's *d* effect size of 0.29. However, and due to RCT design issues, he further noted that the effect size difference should be interpreted with caution.

I argue that the polarization and dichotomy between EBI and unstructured and eclectic treatments often projected in debates are somewhat misplaced. Certainly, regular services may outperform EBI, and EBI do not fit all clients. However, many EBIs, here in the form of parent training interventions, have a build-in component to contextualize treatment due family strengths and challenges (Askeland et al., 2014). In line with Weisz et al.'s (2013) argument above, a central question is whether EBI in general may provide many children with more effective treatments. Philosophical point of views may provide important arguments. However, such arguments would probably not help children that have problems. Nevertheless, critique is important to move the field of EBI forward, and to prevent overly

instrumentalist understandings of children's development and intervention. Hence, there should be room for both structured and unstructured treatments across diverse service settings: we need them both, and efforts should be made to combine the strengths from both approaches to develop more effective interventions.

5. Empirical results – summary of four papers

This section includes summaries of the four empirical papers. As noted, the overall focus in this thesis is on health care disparities in parent training interventions. However, the papers address different research questions, paper I, II, and III gravitates towards to social disparities in health care, whereas paper IV marks a shift from the family oriented focus to an emphasis on PMTO treatment effects in different phases of implementation. I am the sole author of paper I, whereas paper II was co-authored with John Kjøbli, paper III was co-authored with John Kjøbli and Marion Forgatch, and paper IV was co-authored with Terje Ogden. Importantly for this thesis, I am the first author in all the papers. I did all the statistical analyses and wrote the first draught for all the papers. Co-authors provided ideas and input for the research questions and commented on the written manuscripts, and were involved in the design of the RCTs.

Paper I: Social gradients and participant characteristics in child behavior problem interventions

Social disparities have been found in both the prevalence of child behavior problems and among those who use health care services that target such problems. If low-resource families are simultaneously at higher risk of having a child experiencing behavior problems and less inclined to utilize help services, health care disparities may be exacerbated. Accordingly, service use is one of the important domains relating to health care disparities. This paper focuses exclusively on health care disparities related to service use in two versions of PMTO, the high-intensive PMTO group mode and the low-intensive BPT. The paper extends the literature on health care disparities by focusing on the social determinants including several social and economic resources and their relations with service use. Both singular and cumulative relations between family resources and service use were examined. Social

determinates were investigated (i); overall in a pooled sample of the two parent training interventions compared with Norwegian normal population of families with children, and (ii); whether there where social gradients relating to the intensity of treatment.

In contrast to the established health care disparity hypothesis, the results revealed inverse social gradients among those families receiving PMTO and BPT. First, compared to other Norwegian families with children, the TIBIR families had markedly lower social and economic resources on a broad range of indicators. Second, the inverse social gradient was partly replicated through analyses that focused on the intensity in treatment; families with fewer resources – high social risk families – were more likely to receive the high-intensive PMTO. However, cumulative risk families with > 3 social risks were less likely to receive high-intensity treatment. This is concerning since these families are at high risk. Nevertheless, overall the results in this paper indicate that the Norwegian PMTO interventions do not exacerbate health care disparities as a result of who utilizes the services. To the contrary, the findings reveal that providing parent training interventions to selected target groups in different municipal services may prevent social gradients in service use.

Paper II: Family resources and effects on child behavior problem interventions: A cumulative risk approach

Lacks of family resources have both been associated with health care disparities in general and with the outcomes of parent training interventions for children with behavior problems. There is consensus that lack of family resources, often assessed as socioeconomic status, is a social determinant that creates health care disparities in the form of reduced beneficial outcomes in care for behavior problems. Accumulation of social and economic disadvantages in the family – that is cumulative risk – has been established as a social risk in behavior problems development. However, there is scarce evidence on whether cumulative risk may

create health care disparities by being associated with less beneficial outcomes in parent training interventions. Evaluating BPT and PMTO group mode, we constructed cumulative risk index tapping family amounts of social and economic resources together with parent health statuses. Thus, we focused on the effects of quantitative exposure to social risks, and whether cumulative risk moderated treatment effects on child behavior changes in PMTO and BPT. Data was analyzed in a pooled sample and separate analyses in a two case comparison of the interventions differing in intensity and target group. The overall question was whether exposure to cumulative risk created health care disparities in the PMTO interventions compared to regular care.

We did not find any indications of health care disparities following PMTO interventions. To the contrary, we found inverse social gradients relating to treatment condition and cumulative risk exposure: The children exposed to cumulative risk in the pooled sample of PMTO interventions displayed in average more reductions in behavior problem levels compared to families without cumulative risk exposure. Conversely, children in the pooled regular care groups, who were exposed to equal levels of cumulative risk, experienced rising levels of behavior problems compared to those families without cumulative risk. Further, we examined relations between treatment conditions and cumulative risk where we separately analyzed PMTO and BPT, thus relating cumulative risk and health care disparities to intervention intensity. We found that the low-intensity BPT intervention on average worked equally well both for families with and without cumulative risk. In contrast, we found inverse social gradients in the high-intensity PMTO intervention; children from cumulative risk backgrounds experienced high reductions in their level of behavior problems following treatment, and at six-month follow-up. On the one hand, providing evidence-based parent training seems to be an effective way to prevent health care disparities in the Norwegian service system for children with behavior problems. On the other hand, providing

children with the alternative of regular care seems to intensify health care disparities. Moreover, the high-intensity EBI seem to be a particularly effective way to help children from the low-resource and cumulative risk families.

Paper III: Collateral benefits of child behavior problem interventions on parent wellbeing

The family environment, family resources and parent wellbeing, have been found to affect a broad range of developmental outcomes for children, including behavior problems development. Thus, enhancing parent wellbeing may benefit children with behavior problems. In PMTO, parent wellbeing is not explicitly targeted. However, if PMTO has benefits beyond the targeted child behavior problems, and also has effect on parent wellbeing, such collateral benefits implicate that the outcomes of PMTO interventions may be underestimated. Regarding the overall topic of health care disparities in this thesis, paper III contributes by focusing on the capacities of PMTO and BPT to increase parent wellbeing, and thereby to reduce social risk.

We constructed a latent parent wellbeing index tapping parental mental health, somatic health and lack of energy. We used data collected at three time-points and examined the direct effects of treatment conditions on parent wellbeing and indirect effects via change in children's behavior problems, change in parenting style, and via parents level of self-efficacy. In that regard, this paper focuses on the change processes that occur in the family following evidence-based parent training. As in paper I and II, we apply a case study where we examine whether differences in treatment intensity and target groups have different implications for parent wellbeing. Importantly, the BPT sample contains families with lower risk levels compared to the PMTO sample. Albeit research evidence on collateral benefits is relatively limited, scholars have found long-term collateral benefits after three years and

beyond. In paper III we examined whether parent training interventions had more immediate collateral benefits 6 months after treatment termination.

Results revealed that the BPT had direct effects on improving parent wellbeing compared to the families receiving regular care, whereas families receiving PMTO did not experience significant changes on parent wellbeing compared to the regular care families. In addition, both PMTO and BPT had significant effects on improving parenting style, reducing child behavior problems, and predicted higher level of parenting self-efficacy compared to regular care. However, we did not find any significant indirect relations via those mentioned variables on parent wellbeing. The findings suggest that there were two parallel paths of change processes in children and parents six months after receiving BPT. Moreover, the results indicate that reaching children and their parents at early stages in negative developmental cycles may increase the potential for immediate collateral benefits on parent wellbeing. In turn, this may broadly come to benefit child development.

Paper IV: Is there a scale-up penalty? Testing behavioral change in the scaling up of Parent Management Training in Norway

In paper IV, the focus is on disparities related to the health care system itself; whether different phases of implementation were associated with attenuation of program effects. There is a common perception that when EBI is disseminated throughout the service apparatus the effects of interventions attenuates. It has been estimated that program effect reductions varies between 25% and 50%; that there is a scale-up penalty in large-scale implementation. The scale-up penalty has been explained by implementation challenges in large-scale implementation such as the feasibility of the evidence-based intervention, heterogeneity among the practitioners delivering the intervention, heterogeneity in the target group, and challenges to maintain effective implementation to support intervention fidelity

(i.e. adherence EBI manuals and therapeutic common factors). In a case study of the PMTO implementation in Norway, paper IV sets out to test whether there was a scale-up penalty when PMTO (individual mode) was transported out across all parts of the Norwegian service system intended to deliver the intervention. In addition to the testing of a potential scale-up penalty, this paper also describes the consequence of the large-scale dissemination upon practitioner heterogeneity, i.e. level of background training and work place, and target group characteristics, i.e. children's levels of behavior problems.

We found that there was increasing heterogeneity both among service practitioners and among those children and their families that received PMTO in the large-scale phase of implementation compared to the initial phase of effectiveness testing (essentially in Norwegian specialist services). However, and opposed to the hypothesized scale-up penalty, we did not find a scale-up penalty in terms of reduced levels of child behavior change. In fact, differences in the outcomes between the phases of implementation were marginal despite the facts that generalists in the municipal service level more often delivered PMTO in large-scale dissemination and that target children had markedly lower levels of behavior problems in the latter phase of implementation. The PMTO implementation support system was a constant in our case study, meaning that the level of support was essentially the same in both phases of implementation. The absence of a scale-up penalty in Norway suggests that the centralized center strategy of implementation support likely resulted in program maturation effects; in which implementers and support systems became more experienced, skilled, and more culturally adapted to fit the Norwegian context.

6. Concluding remarks

This thesis applies a sociological health care disparity perspective on evidence-based parent training interventions. Thus, the overall focus concerns social disparities in service utilization and treatment outcomes of PMTO interventions in Norway. The social gradient found in

behavior problems implies that children from low-resource backgrounds – children exposed to social risk - are important to serve with effective intervention. Therefore, the theoretical perspective on health care disparities may serve as starting point from which to evaluate whether evidence-based parent training may help this group of children, and thus, whether EBI exacerbate or ameliorate health care disparities and ultimately health disparities. The theoretical focus and the proximity to the practice field provides the opportunity to view the interconnected implications the results in this thesis have for theory and for practice. Therefore, this concluding chapter will discuss the implications for theory, practice, and social policy.

Paper I showed promising results regarding the ability to reach low-risk populations with selected interventions implemented essentially in the municipal service level. Overall participants had markedly lower levels of social and economic resources compared to the normal population of Norwegian families with children. Opposed to a strategy of self-selection and universal provision, the paper I results indicate that implementing both EBI and other interventions close to target groups in municipal services settings, and basing the inclusion into treatment on referrals and practitioner screening, are strategies that may counteract health care disparities in service utilization.

Low amounts of family resources have been recognized as a fundamental determinant that may cause health care disparities by structuring flexible assets to cope with different problems in diverse contexts (Link & Phelan, 2005). For example, flexible assets may include network mechanisms (e.g., friends and colleagues) for advice and support, and cultural mechanisms, for example attitudes towards treatment and self-confidence when interacting with professionals, which has been found to enhance professional help seeking, and thus, service utilization (Gengler, 2014; Gillies, 2006; Thoits, 2011; Weininger & Lareau, 2003). Moreover, high-resource families have also been found to have a greater capacity to

get family members into care (Pescosolido, Boyer, & Medina, 2013). Thus, how families view behavior problems symptoms and react upon them may be otther socially graded asset that influence service utilization. The social gradient found in other studies of service utilization, particularly the Norwegian study by Reedtz et al., (2011) where they found social gradients in a universally provided evidence-based parent training program, may suggest that the mechanisms and social processes associated with high amounts of flexible assets come less into play in selected intervention where provision is based on professional judgments. Moreover, it might also be the case that the close-to-target provision in the municipal services may undermine the practical barriers for service utilization that more often is a constraint for the low-resource families. Thus, when the treatment is provided in a local environment it might be more accessible for the low-resources parents. In sum, the TIBIR interventions, which are at the selected level and implemented in municipal services seem to attract adequate target groups of children with behavior problem symptoms who come from backgrounds marked by low levels of social and economic resources.

Another central issue in health care disparity research concerns the outcomes of treatment. It would increase health care disparities if the help services produce less favorable outcomes for the low-resource participants. In paper II, we examined the impact of family resources, measured as cumulative risk, and whether quantitative lacks of social and economic resources moderated treatment effects in the PMTO interventions and regular care. Overall, the results revealed opposite social gradients in the two treatment conditions according to cumulative risk exposure. Cumulative risk children receiving evidence-based parent training showed more positive behavior change whereas children exposed to equal levels of cumulative risk in regular care showed less positive change compared to children without cumulative risk exposure.

The results indicate that there were opposite mechanisms of change at play in evidence-based parent training and regular care. The results in the regular care group were more in line with the former theories that highlight how lacks of resources are related to fewer flexible assets, including social network mechanisms, cultural mechanisms, and practical mechanisms, that may have created the unfavorable outcomes for low-resource participants in regular care. In that regard, when receiving unstructured regular care, more resources and flexible assets may have impacted on the treatment provided and produced better returns to children from high-resource backgrounds. In contrast, provision of evidence-based parent training seems to counteract this tendency. When receiving the high-intensive PMTO, family resources did not appear to be a fundamental determinant. To the contrary, the low-resource families experienced compensatory effects. The reasons for this might be numerous. For instance, there might be less room for high-resource participants to influence the quality of care in EBI. Moreover, in the theory review, I highlighted how the link between low access to family resources and behavior problems in children are related to family stress and disrupted parenting style (Conger et al., 2010; Patterson, 2002; Capaldi et al., 2002). Evidence-based parent training specifically targets parenting style. As a result, the parent training interventions may more closely match the realties and concrete coercive social processes in the low-resource families, and therefore, work better for them. Thus, parent training may provide low-resource families with the right tools to help their child.

In paper 3, the focus is directed towards the parental outcomes of treatment. As noted in the theory review, the family environment is particularly important for young children (Caspi et al., 2016; Shonkoff et al., 2012; Sroufe et al., 2009). In that regard, if evidence-based parent training may have positive impact on family resources in terms of enhancing parent wellbeing, this would probably have impact on the amount of social risk children are exposed to in the family environment. In keeping with the family resource and stress

hypothesis and the SIL model, enhancing parent wellbeing would reduce the likelihood for disrupted parenting style and thus behavior problems in children. Accordingly the question is whether evidence-based parent training may have collateral benefits and reduce social risk exposure in families that are experiencing a child with behavior problems. The results showed promising results for the BPT intervention, however, in the high-risk PMTO sample no collateral benefits appeared. This suggests that when children and families are reached and helped early in negative developmental cycles, this may increase the likelihood for collateral benefits. In other words; when families are less exposed to cumulative disadvantages, in terms of social risks and lower levels of child problems, less effort is needed to stimulate processes that lead to improved wellbeing both for children and their parents. For example, when the family problems were less cemented, receiving the BPT intervention might have produced feelings of relief and comfort for parents, which in turn, could have affected their subjective feelings of wellbeing. Although research is limited, results from other studies suggest that collateral benefits progress over time (Patterson et al., 2010; McEachern et al., 2013). Thus, improving parents' wellbeing might be a way to promote healthy child development in general, but might also be a way to prevent parent social isolation, and thus, weak labor market attachments and poor family living conditions.

Taken together, the results in this thesis have showed that evidence-based parent training interventions might be a way of preventing and counteracting potential health care disparities for children with behavior problems. Moreover, the results have indicated that health care context and type of treatment could be important to determine whether family resources is a fundamental determinant of favorable flexible assets to cope with behavior problems. Viewed together, the results in paper I and II, and to a certain extent in paper III, underline this argument. Moreover, in the theory review I highlighted the relations between cumulative family disadvantages, family stress and poor outcomes for children (DePrete &

Eirich, 2006; Conger et al., 2010; Seabrook & Avison, 2012). These relations implies that children from the cumulative risk backgrounds are running the risk of being another accumulated disadvantage that adds to the family stress load. The results in this thesis suggest that evidence-based parent training might be a promising way to prevent such negative family trajectories by helping both children and their parents. However, more longitudinal data would have allowed us better opportunities to examine the relations between parent training and family trajectories. Moreover, we do not know what are the actual mechanisms involved in the processes of counteracting health care disparities in parent training. Accordingly, future studies applying a health care disparity perspective on EBI could shed more light on the mechanisms of change in EBI and parent training, also by applying this perspective on EBI targeting other mental health problems.

The results displayed in this thesis may have implications for social policy. In addition to the main effects showing that PMTO interventions were more effective in reducing behavior problems compared to regular care (Bjørknes & Manger, 2013; Kjøbli & Ogden, 2012; Kjøbli et al., 2013; Ogden & Hagen, 2008), the health care disparity perspective in this thesis has provided evidence that parent training interventions is an effective way of targeting behavior problems - also for low-resource families. However, in line with the arguments of Cartwright and Hardie (2012) and Berk (2005) one should be aware of limitations of RCTs to inform policy, due to potential lack of external validity and narrowness in scope. However, the replication of effectiveness in the four RCTs, together with the results from paper I, II, II, and particularly paper IV in this thesis, strongly indicates evidence-based parent training is efficient across diverse service settings and target groups in Norway. In that regard, strengthening the municipal service system with evidence-based parent training has the potential to fulfill two important social policy goals (Government, 2007); (i) to effectively help children identified to be in high-risk positions; (ii) to provide

effective help without the expense of increasing differences between high-resource and low-resource groups. The latter goal specifically relates to the results in this thesis; implementing evidence-based parent training seems to be a promising strategy to help an underserved group of children and families and thereby reduce unwanted variations in health care outcomes. Combining the large body of evidence in favor of evidence-based parent training interventions with the capacity to prevent health care disparities, it seems reasonable to suggest that this research should inform social policy. However, of the approximately 5% of Norwegian children that are in need of intervention, only about 8% receives adequate help (Skogen & Torvik, 2013). When we know the problems these children face in the school system and later in life, this is concerning. A conclusion that can be drawn from this is that evidence-based parent training should be scaled up in the Norwegian service system.

Synthesizing the results from paper I, II and III, the shorter BPT intervention showed positive outcomes for children and parents. In line with arguments regarding individual and societal returns from investment set forth by Heckman and others (Heckman, 2006; Knudsen et al., 2006), these findings suggests that when you target behavior problems at early stages of development, this is a promising way to give broader returns from intervention. From a social policy perspective, the 5 hour BPT is more scalable than the 25-30 hour PMTO interventions: It is less costly to support and provide, and it is more flexible and feasible for implementation in terms of practitioner background and workplace, all of which may indicate that the low-intensive parent training may have a larger potential for improving public health. However, this potential is also dependent on reaching at-risk children and families in the early stages of negative development. To achieve this, health clinics, day care centers, and schools may be the key contexts for early identification of at-risk positions. As shown in paper I, the preventive BPT intervention contained lower risk participants compared to the participants receiving PMTO. Thus, we do not know what benefits the shorter BPT

intervention would give in a higher risk sample. More research is needed to examine such relations, maybe in the form of new RCTs using randomization to different treatment intensity conditioned on social risk.

Acknowledging the public health potential in the shorter parent training intervention, results in this thesis have displayed that particularly the high-risk children from low-resource backgrounds seem to benefit from the high-intensive parent training. Accordingly, there should also be room for provision of high-intensive versions of PMTO in Norway. Moreover, in addition to child symptoms, the results in paper II may suggest that low levels of family resources could warrant provision of high-intensive parent training. This should be kept in mind when screening for behavior problems. Moreover, the results have shown that the most troubled families, having more than 3 cumulative risks, were less likely to partake in the group mode of high-intensive PMTO. This result may suggest that more effort should be focused on including the high-risk families in high-intensive parent training, and further, that providing individual mode of the high-intensive PMTO may be a better alternative for the most disadvantaged families. Moreover, the absence of collateral benefits in the high-risk PMTO sample suggests that cumulative risk families may be in need of additional intervention components that specifically targets parents, for example components targeting health and job-training.

A limitation in the TIBIR program is that it could be criticized for having a too narrow target group by having a one-dimensional focus on behavior problems. This critique may be supported by recent developments in Norwegian welfare policy issued in the Coordination Reform (Government, 2009) and in the later new Child Welfare Act amendment (Government, 2017). These new policy reforms involve a transfer of many health-related services from the state level to the municipal level. As a result, the municipal service providers will need effective interventions to prevent and treat a broad range of

mental health related problems in the years to come. Adapting to these recent developments in social policy, another way forward for TIBIR, and for EBI in the municipal mental health-related services, may be to target other mental health problems in addition to behavior problems. For example internalizing problems, attention deficit hyperactivity disorder, child maltreatment, and conditions in the autism spectrum. As documented particularly in paper IV in this thesis, the TIBIR program seems to be a well-designed platform to implement and provide EBI for behavior problems. Therefore, I believe that TIBIR has the potential to include other mental health problems as well. By broadening the inclusion criteria in TIBIR, the municipal services may reach and help more vulnerable children and families, and thereby produce beneficial public health impacts.

References

Addis, M. E., & Waltz, J. (2002). Implicit and Untested Assumptions About the Role of Psychotherapy Treatment Manuals in Evidence‐Based Mental Health Practice. *Clinical Psychology: Science and Practice, 9*(4), 421-424.

Alegría, M., Pescosolido, B. A., Williams, S., & Canino, G. (2011). Culture, race/ethnicity and disparities: Fleshing out the socio-cultural framework for health services disparities. In B. A. Pescosolido, J. K. Martin, J. D. McLeod, & A. Rogers (Eds.), *Handbook of the sociology of health, illness, and healing* (pp. 363-382): New York, US: Springer Science and Business Media.

Aneshensel, C. S., Phelan, J. C., & Bierman, A. (2013). The sociology of mental health: Surveying the field. In C.S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health* (pp. 1-19): New York, US: Springer Science and Business Media.

Askeland, E., Solholm, R., & Apeland, A. (2014). *PMTO: Foreldretrening for familier med barn som har atferdsvansker*: Oslo, Norway: Gyldendal akademisk.

Avison, W. R., & Comeau, J. (2013). The impact of mental illness on the family. In C.S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health* (pp. 495-515): New York, US: Springer Science and Business Media.

Bank, L., Forgatch, M. S., Patterson, G. R., & Fetrow, R. A. (1993). Parenting practices of single mothers: Mediators of negative contextual factors. *Journal of Marriage and Family, 55*(2), 371-384.

Benner, A. D., & Kim, S. Y. (2010). Understanding Chinese American adolescents' developmental outcomes: Insights from the family stress model. *Journal of Research on Adolescence, 20*(1), 1-12.

Berk, R. A. (2005). Randomized experiments as the bronze standard. *Journal of Experimental Criminology, 1*(4), 417-433.

Bjørnstad, R., Gjelsvik, M. L., Godøy, A., Holm, I., & Stølen, N. M. (2010). Demand and supply of labor by education towards 2030. Linking demographic and macroeconomic models for Norway. Oslo, Norway: Statistics Norway.

Bjørknes, R., & Manger, T. Can parent training alter parent practice and reduce conduct problems in ethnic minority children? A randomized controlled trial. *Prevention Science.* 14(1), 52-63.

Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare‐to‐work experiments. *Journal of Policy Analysis and Management, 22*(4), 551-575.

Breslau, J. (2010). Health in childhood and adolescence and high school dropout. *University of California Drop Out Resarch Project.* California, US.

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E., & Erkanli, A. (1995). Children's mental health service use across service sectors. *Health affairs, 14*(3), 147-159.

Bussing, R., Zima, B. T., Gary, F. A., & Garvan, C. W. (2003). Barriers to detection, help-seeking, and service use for children with ADHD symptoms. *The journal of behavioral health services & research, 30*(2), 176-189.

Bøe, T., Øverland, S., Lundervold, A. J., & Hysing, M. (2012). Socioeconomic status and children's mental health: results from the Bergen Child Study. *Social psychiatry and psychiatric epidemiology, 47*(10), 1557-1566.

Bøyum, S. (2013). Evidensbasert Undervisning-Eit Filosofisk Forsvar. *Nytt Norsk Tidsskrift, 1*. Oslo, Norway: Universitetsforlaget.

Capaldi, D., DeGarmo, D., Patterson, G. R., & Forgatch, M. (2002). Contextual risk across the early life span and association with antisocial behavior. In J. B. Reid, G. R. Patterson, & J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 123-145). Washington, DC, US: American Psychological Association

Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties, 2*(1), 11-20.

Cartwright, N., & Hardie, J. (2012). *Evidence-based policy: a practical guide to doing it better*. Oxford, UK: Oxford University Press.

Caspi, A., Houts, R. M., Belsky, D. W., Harrington, H., Hogan, S., Ramrakha, S., . . . Moffitt, T. E. (2016). Childhood forecasting of a small segment of the population with large economic burden. *Nature Human Behaviour, 1*, 0005. doi:10.1038/s41562-016-0005

Ceci, S. J., & Papierno, P. B. (2005). The rhetoric and reality of gap closing: when the" have-nots" gain but the" haves" gain even more. *American Psychologist, 60*(2), 149-160.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of consulting and clinical psychology, 66*(1), 7-18.

Conger, R. D., Conger, K. J., Elder, G. H., Lorenz, F. O., Simons, R. L., & Whitbeck, L. B. (1992). A family process model of economic hardship and adjustment of early adolescent boys. *Child development, 63*(3), 526-541.

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family, 72*(3), 685-704.

Conger, R. D., Wallace, L. E., Sun, Y., Simons, R. L., McLoyd, V. C., & Brody, G. H. (2002). Economic pressure in African American families: a replication and extension of the family stress model. *Developmental psychology, 38*(2), 179-193.

Costello, E. J., & Angold, A. (2001). Bad behaviour: An historical perspective on disorders of conduct. In J. Hill & B. Maughan, *Conduct disorders in childhood and adolescence*, 1-31 Cambridge, UK: Camebridge Univeristy Press.

Cuellar, A. (2015). Preventing and treating child mental health problems. *The future of children, 25*(1), 111-134.

Deater–Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1998). Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology, 10*(03), 469-493.

Deković, M., Slagt, M. I., Asscher, J. J., Boendermaker, L., Eichelsheim, V. I., & Prinzie, P. (2011). Effects of early prevention programs on adult criminal offending: A meta-analysis. *Clinical psychology review, 31*(4), 532-544.

Digre, K., & Haugberg, T. (2016). *The Education Mirror*. http://utdanningsspeilet.udir.no/2016/wp-content/uploads/2016/10/Utdanningsspeilet_2016_en.pdf: The Norwegian Directorate for Education and Training.

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual review of sociology*, 271-297.

Dishion, T. J., & Patterson, G. R. (2006). The development and ecology of antisocial behavior in children and adolescents. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol 3: Risk, disorder, and adaptation (2nd ed)* pp. 421-471. Hoboken, NJ, US: John Wiley & Sons.

Dunn, P. M. (1997). James Lind (1716-94) of Edinburgh and the treatment of scurvy. *Archives of Disease in Childhood-Fetal and Neonatal Edition, 76*(1), F64-F65.

Durkheim, E. (1951). Suicide: A study in sociology (JA Spaulding & G. Simpson, trans.). Glencoe, IL: Free Press.(Original work published 1897).

Edwards, R. T., Céilleachair, A., Bywater, T., Hughes, D. A., & Hutchings, J. (2007). Parenting programme for parents of children at risk of developing conduct disorder: cost effectiveness analysis. *Bmj, 334*(7595), 682-688.

Ekeland, T. (2007). Evidens og nypositivisme I: *Forskningspolitikk*. 30(2), 10-11.

Elder, D., Downey, G., & Cross, C. E. (1986). Family ties and life chances: Hard times and hard choices in women's lives since the 1930s. *Life-span developmental psychology: Intergenerational relations*, 151-183.

Elstad, J., & Stefansen, K. (2014). Social Variations in Perceived Parenting Styles among Norwegian Adolescents. *Child Indicators Research, 7*(3), 649-670. doi:10.1007/s12187-014-9239-5

Eng, H., Lauritzen, C., Reedtz, C., Mørch, W.-T., & Martinussen, M. (2014). Establishing evidence-based practices within services for children: knowledge transfer challenges. *BMC Health Services Research, 14*(2), P39.

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin, 139*(6), 1342-1396.

Evensen, M., Lyngstad, T. H., Melkevik, O., & Mykletun, A. (2016). The Role of Internalizing and Externalizing Problems in Adolescence for Adult Educational Attainment: Evidence from Sibling Comparisons. *European Sociological Review*, jcw001 552-566.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*(5), 531-540.

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*(3), 151-175.

Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining fidelity following the nationwide PMTO™ implementation in Norway. *Prevention Science, 12*(3), 235-246.

Forgatch, M. S., & Kjøbli, J. (2016). Parent Management Training—Oregon Model: Adapting Intervention with Rigorous Research. *Family Process, 55 (3),* 500-513.

Forgatch, M. S., & Patterson, G. R. (2010). Parent Management Training—Oregon Model: An intervention for antisocial behavior in children and adolescents. *Evidence-based psychotherapies for children and adolescents, 2*, 159-178.

Frønes, I. (2016). Høyterskelsamfunnet In I. Frønes, H. Eng, S. E. Ertesvåg, & J. Kjøbli (Ed.), *Risiko, intervensjon og evidens* (pp. 110-126). Oslo, Norway: Gyldendal Akademisk.

Frønes, I., & Strømme, H. (2014). *Risiko og marginalisering: norske barns levekår i kunnskapssamfunnet*. Oslo, Norway: Gyldendal akademisk.

Galambos, N. L., Barker, E. T., & Almeida, D. M. (2003). Parents do matter: Trajectories of change in externalizing and internalizing problems in early adolescence. *Child development, 74*(2), 578-594.

Garfield, S. L. (1996). Some problems associated with "validated" forms of psychotherapy. *Clinical Psychology: Science and Practice, 3*(3), 218-229.

Ge, X., Conger, R. D., Lorenz, F. O., Shanahan, M., & Elder Jr, G. H. (1995). Mutual influences in parent and adolescent psychological distress. *Developmental psychology, 31*(3), 406-419.

Gengler, A. M. (2014). "I Want You to Save My Kid!" Illness Management Strategies, Access, and Inequality at an Elite University Research Hospital. *Journal of Health and Social Behavior, 55*(3), 342-359.

Gillies, V. (2006). *Marginalised mothers: Exploring working class experiences of parenting*. New York, US: Routledge.

Giæver, Ø. (2013). *Nasjonal strategi for å utjevne sosiale helseforskjeller–i teori og praksis. Underlagsrapport til Sosial ulikhet i helse: En norsk kunnskapsoversikt*. Retrieved from: www.hioa.no.

Glied, S., & Oellerich, D. (2014). Two-generation programs and health. *The future of children, 24*(1), 79-97.

Goodman, S. H., & Gotlib, I. H. (2002). Children of depressed parents: Mechanisms of risk and implications for treatment. Washington, DC, US: American Psychological Association.

Government. (2009). *White paper: Report No. 47 The Coordination Reform: Proper treatement - at the right place and right time*. https://www.regjeringen.no/en/dokumenter/report.no.-47-to-the-storting-2008-2009/id567201/: Government.no.

Government. (2017). *Amendment to Child Welfare Act, Prop. 73L*. Government.no Retrieved from https://www.regjeringen.no/no/dokumenter/prop.-73-l-20162017/id2546056/.

Government. (2007). *Report No. 20 to the storting (2006-2007) National strategy to reduce social inequalities in health*. https://www.regjeringen.no/no/dokumenter/meld-st-20-20122013/id717308/.

Greenwood, P. W. (1998). *Diverting children from a life of crime: Measuring costs and benefits*: California, US: RAND Corporation.

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

Haggerty, K., Fleming, C., Lonczak, H., Oxford, M., Harachi, T., & Catalano, R. (2002). Predictors of Participation in Parenting Workshops. *Journal of Primary Prevention, 22*(4), 375-387. doi:10.1023/A:1015227623145

Hanson, T. L., McLanahan, S. S., & Thomson, E. (1995). Economic resources parental practices and child well-being. *Office of Population Research Working Paper* No. 95-9. New Jersey, US: Princeton University.

Haskins, R., Garfinkel, I., & McLanahan, S. (2014). Introduction: Two-generation mechanisms of child development. *The future of children, 24*(1), 3-12.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900-1902.

Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity, 43*(1), 59-74.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.

Kaehler, L. A., Jacobs, M., & Jones, D. J. (2016). Distilling Common History and Practice Elements to Inform Dissemination: Hanf-Model BPT Programs as an Example. *Clinical child and family psychology review, 19*(3), 236-258.

Kalil, A., & Ryan, R. M. (2010). Mothers' economic conditions and sources of support in fragile families. *The future of children, 20*(2), 39-61.

Kazdin, A. E., Holland, L., & Crowley, M. (1997). Family experience of barriers to treatment and premature termination from child therapy. *Journal of consulting and clinical psychology, 65*(3), 453-463.

Kjøbli, J., & Bjørnebekk, G. (2013). A randomized effectiveness trial of Brief Parent Training: Six-month follow-up. *Research on Social Work Practice, 23*(6), 603-612.

Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: reducing child conduct problems in real-world settings. *Behaviour research and therapy, 51*(3), 113-121.

Kjøbli, J., & Ogden, T. (2012). A randomized effectiveness trial of brief parent training in primary care settings. *Prevention Science, 13*(6), 616-626.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, US: Guilford Press.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences, 103*(27), 10155-10162.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of general psychiatry, 59*(10), 877-883.

Kristiansen, I. S., & Mooney, G. (2004). *Evidence-based medicine: In its place*. New York, US: Routledge.

Kristoffersen, L. B. (2017). *Oppvekstrapporten Section 5: Social inequality in child protection services*. Retrived from: https://www.bufdir.no/bibliotek/Dokumentside/?docId=BUF00003918.

Krogh, F., & Kvello Bukten, M. (2013). Aktivitetsdata for psykisk helsevern for barn og unge 2014. Retrived from: www.helsedirektoratet.no/Lists/Publikasjoner/Attachments/935/BUP-rapport-2014. pdf.

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. California, US: University of California Press.

Leijten, P., Raaijmakers, M. A., de Castro, B. O., & Matthys, W. (2013). Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. *Journal of Clinical Child & Adolescent Psychology, 42*(3), 384-392.

Leijten, P., Raaijmakers, M. A., Orobio de Castro, B., van den Ban, E., & Matthys, W. (2015). Effectiveness of the Incredible Years Parenting Program for Families with Socioeconomically Disadvantaged and Ethnic Minority Backgrounds. *Journal of Clinical Child & Adolescent Psychology*(ahead-of-print), 1-15.

Link, B. G., & Phelan, J. C. (2005). Fundamental sources of health inequalities. *Policy challenges in modern health care*, 71-84.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural equation modeling, 9*(2), 151-173.

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical psychology review, 26*(1), 86-104.

Markussen, S., & Røed, K. (2017). *Egalitarianism under Pressure Toward Lower Economic Mobility in the Knowledge Economy?* Retrieved from https://www.frisch.uio.no/docs/PDF/social-mobility-4.pdf:

Marmot, M., Friel, S., Bell, R., Houweling, T. A., Taylor, S., & Health, C. o. S. D. o. (2008). Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet, 372*(9650), 1661-1669.

Marmot, M., & Wilkinson, R. (2005). *Social determinants of health*. Oxford. US: Oxford University Press.

Mazza, J. R. S., Boivin, M., Tremblay, R. E., Michel, G., Salla, J., Lambert, J., . . . Côté, S. M. (2016). Poverty and behavior problems trajectories from 1.5 to 8 years of age: Is the gap widening between poor and non-poor children? *Social psychiatry and psychiatric epidemiology*, 1-10.

McConaughy, S. H., & Skiba, R. J. (1993). Comorbidity of externalizing and internalizing problems. *School Psychology Review,* 421-436.

McEachern, A. D., Fosco, G. M., Dishion, T. J., Shaw, D. S., Wilson, M. N., & Gardner, F. (2013). Collateral benefits of the family check-up in early childhood: Primary caregivers' social support and relationship satisfaction. *Journal of family psychology, 27*(2), 271-281.

McLeod, J. D. (2013). Social stratification and inequality. In C.S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health* (pp. 229-253): New York, US: Springer Science and Business Media.

McLeod, J. D., & Almazan, E. P. (2003). *Connections between childhood and adulthood*. In J. Mortimer & M Shanahan (Eds.), *Handbook of the life course* (pp. 391-411). New York, US: Springer.

McLeod, J. D., & Kaiser, K. (2004). Childhood emotional and behavioral problems and educational attainment. *American sociological review, 69*(5), 636-658.

McLeod, J. D., & Shanahan, M. J. (1993). Poverty, parenting, and children's mental health. *American sociological review*, 351-366.

McLeod, J. D., & Shanahan, M. J. (1996). Trajectories of poverty and children's mental health. *Journal of Health and Social Behavior*, 207-220.

Meadows, S. O., McLanahan, S., & Knab, J. T. (2009). Economic Trajectories in Non-Traditional Families with Children. *RAND Working Paper Series WR- 701.*

Melhuish, E. C. (2011). Preschool matters. *Science, 333*(6040), 299-300.

Mihalic, S., Irwin, K., Elliott, D., Fagan, A., & Hansen, D. (2004). *Blueprints for violence prevention.* Paper presented at the for the Study and Prevention of Violence, University of Colorado. National Scientific Council on the Developing Child.

Moffitt, & Scott, S. (2009). Conduct disorders of childhood and adolescence. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), *Rutter's child and adolescent psychiatry* (5th ed., pp. 543-564). Oxford, UK: Blackwell.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Ross, S. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*(7), 2693-2698.

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology, 14*(01), 179-207.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review, 26*(1), 67-82.

Murali, V., & Oyebode, F. (2004). Poverty, social inequality and mental health. *Advances in psychiatric treatment, 10*(3), 216-224.

Muthén, B., & Muthén, L. (2012). Mplus version 7 [Computer Programme]. Los Angeles, CA, US: Muthén & Muthén.

National Academies of Sciences, E., & Medicine. (2017). *Communities in Action: Pathways to Health Equity*. Washington, DC, US: National Academies Press.

O'Connor, E. E., Dearing, E., & Collins, B. A. (2011). Teacher-child relationship and behavior problem trajectories in elementary school. *American educational research Journal, 48*(1), 120-162.

Ogden, T., & Fixsen, D. L. (2014). Implementation Science. *Zeitschrift für Psychologie, 222*(1), 4-11.

Ogden, T., Forgatch, M. S., Askeland, E., Patterson, G. R., & Bullock, B. M. (2005). Implementation of parent management training at the national level: The case of Norway. *Journal of Social Work Practice, 19*(3), 317-329.

Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of Parent Management Training in Norway: a randomized controlled trial of children with conduct problems. *Journal of consulting and clinical psychology, 76*(4), 607-621.

Parke, R. D., Coltrane, S., Duffy, S., Buriel, R., Dennis, J., Powers, J., . . . Widaman, K. F. (2004). Economic stress, parenting, and child adjustment in Mexican American and European American families. *Child development, 75*(6), 1632-1656.

Patterson, G. R. (1996). Some characteristics of a developmental theory for early-onset delinquency. *Frontiers of developmental psychopathology*, 81-124.

Patterson, G. R. (2002). The early development of coercive family process. In J. B. Reid, G. R. Patterson, & J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 25-44). Washington, DC, US: American Psychological Association

Patterson, G. R., Forgatch, M. S., & DeGarmo, D. S. (2010). Cascading effects following intervention. *Development and Psychopathology, 22*(04), 949-970.

Pescosolido, B. A., Boyer, C. A., & Medina, T. R. (2013). The social dynamics of responding to mental health problems. In C.S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health* (pp. 505-524): New York, US: Springer Science and Business Media.

Pettersson, C., Lindén-Boström, M., & Eriksson, C. (2009). Reasons for non-participation in a parental program concerning underage drinking: a mixed-method study. *BMC public health, 9*(1), 1-19.

Piotrowska, P. J., Stride, C. B., Croft, S. E., & Rowe, R. (2015). Socioeconomic status and antisocial behaviour among children and adolescents: A systematic review and meta-analysis. *Clinical psychology review, 35*, 47-55.

Reedtz, C., Martinussen, M., Wang Jørgensen, F., Helge Handegård, B., & Mørch, W.-T. (2011). Parents seeking help in child rearing: who are they and how do their children behave? *Journal of Children's Services, 6*(4), 264-274.

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems – a meta - analytic review. *Journal of child psychology and psychiatry, 47*(1), 99-111.

Rhemtulla, M. (2016). Population performance of SEM parceling strategies under measurement and structural model misspecification. *Psychological Methods, 21*(3), 348-368. doi:10.1037/met0000072

Rishel, C. W. (2012). Pathways to prevention for children of depressed mothers: a review of the literature and recommendations for practice. *Depression research and treatment,* 1-11.

Robins, L. N. (1966). DEVIANT CHILDREN GROWN UP: A sociological and psychiatric study of sociopathic personality. California, US: Williams & Wilkins.

Roe, B. E., & Just, D. R. (2009). Internal and external validity in economics research: Tradeoffs between experiments, field experiments, natural experiments, and field data. *American Journal of Agricultural Economics, 91*(5), 1266-1271.

Sagatun, Å., Heyerdahl, S., Wentzel-Larsen, T., & Lien, L. (2014). Mental health problems in the 10 th grade and non-completion of upper secondary school: the mediating role of grades in a population-based longitudinal study. *BMC public health, 14*(1), 1-13.

Sameroff, A. J., Bartko, W. T., Baldwin, A., Baldwin, C., & Seifer, R. (1998). Family and social influences on the development of child competence. In M. Lewis & C. Feiring (Eds.), *Families, risk, and competence* (pp. 161-185). Mahwah, NJ: Lawrence Erlbaum.

Satcher, D. (2000). Report of the surgeon general's conference on children's mental health: A national action agenda. *American Journal of Health and Education,* 32(3)*,* 179-182.

Sayal, K., Washbrook, E., & Propper, C. (2015). Childhood Behavior Problems and Academic Outcomes in Adolescence: Longitudinal Population-Based Study. *Journal of the American Academy of Child & Adolescent Psychiatry,* 54 (5), 360-368.

Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow up study of antisocial children into adulthood. *Bmj,* 323(7306), 1-5.

Seabrook, J. A., & Avison, W. R. (2012). Socioeconomic status and cumulative disadvantage processes across the life course: implications for health outcomes. *Canadian Review of Sociology/Revue canadienne de sociologie, 49*(1), 50-68.

Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*(2), 171-186.

Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., McGuinn, L., . . . Care, D. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics, 129*(1), e232-e246.

Skogen, J., & Torvik, A. (2013). Atferdsforstyrrelser blant barn og unge i Norge. Beregnet forekomst og bruk av hjelpetiltak. Report 2013/4. Oslo, Norway: Folkehelseinstituttet.

Smith, K. P., & Christakis, N. A. (2008). Social networks and health. *Annu. Rev. Sociol, 34*, 405-429.

Solantaus, T., Leinonen, J., & Punamäki, R.-L. (2004). Children's mental health in times of economic recession: replication and extension of the family economic stress model in Finland. *Developmental psychology, 40*(3), 412-429.

Solholm, R., Kjøbli, J., & Christiansen, T. (2013). Early Initiatives for Children at Risk—Development of a Program for the Prevention and Treatment of Behavior Problems in Primary Services. *Prevention Science, 14*(6), 535-544. doi:10.1007/s11121-012-0334-x

Solholm, R., Ogden, T., & Jakobsen, R. (2014). Conduct problems in children. Characteristics of families recruited for a clinical outcome trial as compared to families in an implementation study. *Children and Youth Services Review, 42*, 1-9.

Spencer, K. L., & Grace, M. (2016). Social foundations of healthcare inequality and treatment bias. *Annual Review of Sociology, 42*, 101-120. doi:10.1146/annurev-soc-081715-074226

Sroufe, L. A., Egeland, B., Carlson, E. A., & Collins, W. A. (2009). *The development of the person: The Minnesota study of risk and adaptation from birth to adulthood*: New York, US: Guilford Press.

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics. New York, US: Pearson.

Thoits, P. A. (2011). Mechanisms linking social ties and support to physical and mental health. *Journal of Health and Social Behavior, 52*(2), 145-161.

Trentacosta, C. J., Hyde, L. W., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. (2008). The relations among cumulative risk, parenting, and behavior problems during early childhood. *Journal of child psychology and psychiatry, 49*(11), 1211-1219.

Trochim, W. (2006). Introduction to Validity. Retrieved from: https://www.socialresearchmethods.net/kb/introval.php

Turner, R. J., Wheaton, B., & Lloyd, D. A. (1995). The epidemiology of social stress. *American sociological review*, 104-125.

Vogt, K. C. (2017). Vår utålmodighet med ungdom. *Tidsskrift for samfunnsforskning, 58*(01), 105-119.

Waldfogel, J., Craigie, T.-A., & Brooks-Gunn, J. (2010). Fragile families and child wellbeing. *The Future of children/Center for the Future of Children, the David and Lucile Packard Foundation, 20*(2), 87-112.

Weininger, E. B., & Lareau, A. (2003). Translating Bourdieu into the American context: the question of social class and family-school relations. *Poetics, 31*(5), 375-402.

Weisz, J. R., & Kazdin, A. E. (2010). *Evidence-based psychotherapies for children and adolescents*. New York, US : Guilford Press.

Weisz, J. R., Kuppens, S., Eckshtain, D., Ugueto, A. M., Hawley, K. M., & Jensen-Doss, A. (2013). Performance of evidence-based youth psychotherapies compared with usual clinical care: a multilevel meta-analysis. *JAMA psychiatry, 70*(7), 750-761.

Welsh, B. C., Sullivan, C. J., & Olds, D. L. (2010). When early crime prevention goes to scale: a new look at the evidence. *Prev Sci, 11*(2), 115-125. doi:10.1007/s11121-009-0159-4

Wheaton, B. (1994). Sampling the stress universe. In W. R. Avison & I. H. Gotlib, *Stress and mental health* (pp. 77-114). New York, US: Springer.

Wichstrøm, L., Berg‐Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E., & Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers. *Journal of child psychology and psychiatry, 53*(6), 695-705.

Zilberstein, K. (2016). Soapbox: Class matters in parenting interventions. *Clinical Child Psychology and Psychiatry, 21*(3), 359-367.

I

II

J Child Fam Stud
DOI 10.1007/s10826-017-0777-6

ORIGINAL PAPER

Family Resources and Effects on Child Behavior Problem Interventions: A Cumulative Risk Approach

Truls Tømmerås[1] · **John Kjøbli**[2]

Abstract Family resources have been associated with health care inequality in general and with social gradients in treatment outcomes for children with behavior problems. However, there is limited evidence concerning cumulative risk—the accumulation of social and economic disadvantages in a family—and whether cumulative risk moderates the outcomes of evidence-based parent training interventions. We used data from two randomized controlled trials evaluating high-intensity ($n = 137$) and low-intensity ($n = 216$) versions of Parent Management Training—Oregon (PMTO) with a 50:50 allocation between participants receiving PMTO interventions or regular care. A nine-item family cumulative risk index tapping socioeconomic resources and parental health was constructed to assess the family's exposure to risk. Autoregressive structured equation models (SEM) were run to investigate whether cumulative risk moderated child behaviors at post-treatment and follow-up (6 months). Our results showed opposite social gradients for the treatment conditions: the children exposed to cumulative risk in a pooled sample of both PMTO groups displayed lower levels of behavior problems, whereas children with identical risk exposures who received regular care experienced more problems. Furthermore, our results indicated that the social gradients differed between PMTO interventions: children exposed to cumulative risk in the low-intensity (five sessions) Brief Parent Training fared equally well as their high-resource counterparts, whereas children exposed to cumulative risk in the high-intensity PMTO (12 sessions) experienced vastly better treatment effects. Providing evidence-based parent training seem to be an effective way to counteract health care inequality, and the more intensive PMTO treatment seemed to be a particularly effective way to help families with cumulative risk.

Keywords Family resources · Social risk · Cumulative risk · Behavior problems · Health care inequality · Evidence-based parent training interventions

Introduction

It is well established that behavioral problems in childhood (i.e., conduct problems, oppositional behaviors, and inattentive problems) negatively impact children's long-term well-being through an association with school problems, work problems, social exclusion, and poor health (Maughan et al. 1985; Rutter et al. 1970; Sroufe et al. 2009). As with many other mental health-related problems, a social gradient has been established for behavior problems. Specifically, a family's lack of social and economic resources has been found to be a social risk factor for the development and prevalence of such problems (Bøe et al. 2012; Mazza et al. 2016; Piotrowska et al. 2015; Sameroff et al. 1998). Furthermore, it has been recognized that there is a social gradient in the outcomes of mental health care services in general (i.e. health care inequality) and in the parent training

✉ Truls Tømmerås
trulstom@atferdssenteret.no

John Kjøbli
john.kjobli@r-bup.no

[1] Norwegian Center for Child Behavioral Development, P.O. Box 7053, Majorstuen 0306 Oslo, Norway

[2] Regional Center for Child and Adolescent Mental Health, Eastern and Southern Norway, P.O. Box 4623, Nydalen 0405 Oslo, Norway

Published online: 18 May 2017

interventions targeting behavior problems (Leijten et al. 2013; Lundahl et al. 2006; Pescosolido et al. 2013). Knowing that behavioral problems are one of the most frequent reasons for referrals to mental health services (Storvoll 1997; World Health Organization 2003), health care inequality in the services provided to children with behavior problems implies that we will fail to help high-risk children and that the effectiveness of these services will be reduced.

Health inequality remains a major societal challenge, and extensive research has examined how health care systems exacerbate these disparities (Spencer and Grace 2016). Service use, patient adherence, and service outcomes have been acknowledged as elements that are important for understanding health care inequality (Alegría et al. 2011), and lack of family resources has been an important focus of health care inequality research (Muntaner et al. 2013). In caring for behavior problems, social gradient approaches have commonly focused on family resources in the form of socioeconomic status (SES), typically assessed in terms of parental income and education level (Leijten et al. 2013). SES has been proposed as a "fundamental cause" of health inequality, structuring (un-) favorable mechanisms across contexts and diseases (Link and Phelan 1995; Muntaner et al. 2013). This implies that high SES families enjoy a vast number of flexible assets that they can use to their advantage to implement protective strategies and produce favorable treatment outcomes.

A more finely graded family resource approach may involve measurement of a wide array of social and economic resources, including parental mental and somatic health, that are associated with SES, (i.e., cumulative risk). Cumulative risk denotes a situation in which several social risk factors operate together. To give an example, comprehending a risk factor often involves envisioning the circumstances that accompany it. For instance, is having a single mother a risk if she has good health, a good income, and only one child? Probably not. However, if the single mother is poor, undereducated, and has three children, the picture is different. This example supports the fact that human beings often contend with constellations of risk factors rather than isolated instances of adverse circumstances (Seifer et al. 1992). In parent training, parents are the agents of change in their children (Forgatch and Patterson 2010). Thus, it is likely that limitations in parents' access to resources create strain and stress on family life which in turn may create health care inequality.

Social gradients and thus health care inequality have been identified in interventions for behavior problems (Lundahl et al. 2006). In a meta-analysis, Reyno and McGrath (2006) found that single parent status, family size, low income, low education, and parental mental health issues diminished the effects of parent training. In another meta-analysis, Lundahl et al. (2006) found that low SES families benefitted less from parent training, particularly when the mode of treatment was group therapy. However, findings regarding family resources and the outcomes of parent training are inconsistent (Deković et al. 2011). For instance, low educational level, low marital satisfaction, maternal depression, and a lack of psychological resources have been found to enhance the benefits of treatment (Berlin et al. 1998; Gardner et al. 2009; Lundahl et al. 2006).

These conflicting findings regarding separate family risks have been taken as evidence that it is not the quality but the quantity of resources that is relevant to family functioning and children's behavior (Rutter 2000; Sameroff et al. 1993; see Stolk et al. 2008, p. 57). However, we know of only two previous studies that have investigated how cumulative risk influences the effects of parent training interventions for addressing child behavior problems. In their study of a program aimed at reducing behavioral problems among children aged 1 to 3 years, Stolk et al. (2008) found no associations between cumulative risk and treatment effects. In studying another parenting intervention intended to promote cognitive development among low-birthweight infants, Liaw and Brooks-Gunn (1994) found that cumulative risk did not moderate treatment effects on behavior problems.

The social gradients in parent training outcomes are likely grounded in rings of social influence ranging from societal macro-level factors to micro-level factors such as individual characteristics and client-practitioner interactions (Spencer and Grace 2016). Hence, family resources and cumulative risk are proxies for different change mechanisms that operate at different levels of health care inequality and account for more complex situational decisions, rationalizations, and reasons for actions following parent training intervention. Given the findings of the above review, there is no consensus on whether family resources moderate the benefits of parent training interventions. Several change mechanisms are likely to vary according to family resources and thus impact treatment outcomes for behavior problems.

Following intervention, social network is one group of mechanisms that likely affects social gradients in treatment outcomes. According to Thoits (2011), socially graded network mechanisms may affect parents' coping strategies. Low-resource families have less access to secondary networks of significant others who can offer various types of beneficial support, such as information and advice on interventions, encouragement, social influence, and role modeling based on past experiences (Smith and Christakis 2008; Thoits 2011). Sociocultural mechanisms may be another group of mechanisms that can reduce intervention benefits in socially graded patterns. Low-resource parents may hold more negative attitudes toward treatments and professional advice. Their norms, practices, and values may

adhere less strictly to the parenting practices recommended in parent training (Gillies 2006), as parenting interventions may more closely match the resources and realities of middle-class households (Zilberstein 2016). Skill acquisition in parent training may also be less suited to low-resource parents' modes of learning and loci of control (Pescosolido et al. 2013; Zilberstein 2016). Moreover, low-resource parents have been found to behave in a less self-assertive way when interacting with professionals, leading to differential outcomes that may disfavor their children (Gengler 2014; Weininger and Lareau 2003). Finally, different practical mechanisms, in the form of social and economic stressors (Conger et al. 1992), may reduce low-resource families' potential to gain from interventions. Parents with cumulative risk are less likely to live in traditional two-parent families, and they might have poorer health and less money to pay for transportation expenses and child care, all of which could limit their access to the practical and social resources needed to participate in intervention and integrate the parenting strategies learned into their daily lives.

Although there is a general consensus that parent training interventions work better for high-resource families, some scholars have found compensatory effects (i.e., more positive effects) of parent training for the low-resource families (Leijten et al. 2013). Thus, several mechanisms are likely to be involved in beneficial compensatory patterns. Initial problem severity is a factor that could impact the benefits of treatment (Leijten et al. 2013) in that more severely troubled children may have more room for improvement. Similarly, the severity of a child's problem might impact parental motivation and readiness to change (Baydar et al. 2003; DiClemente and Velasquez 2002). Features of the interventions could create compensatory effects for children from families with cumulative risk. Evidence-based interventions are based on a curriculum and follow a structured progression. Thus, there might be less room for high-resource parents to influence the treatment situation in an evidence-based intervention, where the teaching of core components is somewhat fixed. In that regard, evidence-based interventions might promote equality of care by ensuring that effective practices are provided to both rich and poor families (Cochrane 2004; Kristiansen and Mooney 2004). Moreover, the parent training interventions in focus are based on the Structured Interaction Learning model (SIL) of behavioral change (Forgatch and Patterson 2010). According to the SIL, family resources, and thus cumulative risk, affect the development of behavior problems by disrupting parenting style. Hence, children exposed to cumulative risk may have behavior problems that are more strongly induced by social risk and disrupted parenting. This implies that the SIL-based PMTO interventions might work better for families with cumulative risk, as their children's behavior problem etiology might be more influenced by social risk environments and thus more in line with the PMTO curriculum. Relatedly, low-resource parents have more often been found to adhere to negative parenting styles (Elstad and Stefansen 2014). Hence, the parenting focus in interventions may be better suited to low-resource parents at the pre-intervention stage if they, to a greater degree than high-resource parents, lack parenting skills and abilities. If so, the practical focus on and rehearsal of parenting practices, particularly in the high-intensive intervention (explained in more detail below), could add a compensatory mechanism to the treatment experience of cumulative risk families.

In this study, we examined two evidence-based interventions differing in intensity (i.e., dosage and scope): the high-intensity PMTO Parent Group (hereafter called PMTO) and the low-intensity PMTO short form Brief Parent Training (BPT; when discussed together hereafter, these are called PMTO interventions). We focused on exposure to environments that were characterized by a lack of family social and economic resources in which we assessed the quantity of family resources, i.e., cumulative risk. The primary question raised was whether cumulative risk moderated the treatment effects of the PMTO interventions. Thus we elaborated on the conditions under which quantitative aspects of family resources exacerbate or ameliorate health care inequalities in (1) parent training vs. regular care and (2) a two-case comparison of the low-intensity BPT intervention and the high-intensity PMTO intervention.

Method

Participants

We used data from two randomized experiments evaluating PMTO and BPT interventions. They were designed as pretest (T1), posttest (T2; 8 weeks and 12 weeks after pretest for BPT and PMTO, respectively) and follow-up (T3; 6 months after post-test) parallel-group randomized trials with a 50:50 allocation ratio for the intervention and comparison groups. This implies that children and their parents were randomized to either one of the PMTO intervention groups or to the comparison groups receiving the regular care offered in the Norwegian health care system for children with behavior problems. The participants were randomized after completing the pretest questionnaire. Data collection occurred from 2007 to 2008 for BPT and from 2008 to 2009 for PMTO. Importantly, the recruitment and data-collection procedures were similar for both the BPT and PMTO groups. The participating children and families came from all five Norwegian health regions. The families had contacted the services themselves or had been referred by a primary care agency (e.g., child health clinics, child welfare agencies, schools or

kindergartens). To mirror the regular referral procedures used in Norwegian health care services at that time, no formal screenings were used in this study. Thus, the inclusion of children and their families was based on practitioners' clinical opinions after they had consulted with eligible parents. The participants were the parents (or caretakers) of children between the ages of 3 and 12 years.

Participant characteristics in the PMTO and BPT samples have been examined in a study by Tømmeraas 2016. Both samples overall contained participants that had lower economic and social resources compared to the Norwegian population of families with children. Regarding baseline differences between the two samples, the participant characteristics differed between the BPT and the PMTO sample. The descriptive statistics and baseline differences between the samples are described in Table 1. As Table 1 shows, the PMTO participants were clearly different from the BPT sample as they had more family risks and higher levels of behavior problems. Finally, we examined baseline differences related to treatment conditions in the two samples in terms of demographic and child behavior differences. In the PMTO sample there were no significant baseline differences between those receiving PMTO or the alternative of regular care. In the BPT sample, one difference emerged. Parent in the regular care group had on average higher education levels compared to the BPT group, $t(185) = 2.47$, $p = 0.1$.

Procedure

Both BPT and PMTO are part of a comprehensive evidence-based intervention program called TIBIR (Norwegian acronym; Early Initiatives for Children at Risk), which was developed to prevent and treat behavior problems in children (Solholm et al. 2013). PMTO targets children at moderate to high risk and is an intensive intervention consisting of 12 weekly sessions of 2.5 h each. The reduction of negative parenting and teaching and the rehearsal of positive parenting skills are central to PMTO; parents practice their parenting skills through role-play and participate in discussions. PMTO focuses on the following parenting skills: positive involvement, skill encouragement, family problem solving, monitoring, and effective discipline, which includes mild contingent sanctioning through ignoring and time-outs (or cool-downs). Moreover, much emphasis is placed on parent emotional control to reduce coercive interaction cycles between parents and children.

BPT is a low-intensity intervention that targets children between low and moderate risk and consists of up to five 1-h sessions. In these sessions, parents are taught only the most exigent parenting practices, much less time is devoted to skill rehearsal, and positive involvement and effective discipline are emphasized. Important differences between the two interventions are difference in dosage and comprehensiveness and the fact that PMTO is a group therapy. Both BPT and PMTO group and individual therapies have been tested in randomized effectiveness trials and have been shown to be more effective for reducing behavior problems than the practices regularly used in the Norwegian health care system for children with behavior problems (Kjøbli et al. 2013; Kjøbli and Ogden 2012; Ogden and Hagen 2008).

Regular care consisted of the following approaches: 63 families (35%) received no treatment, 25 families received

Table 1 Descriptive statistics (means, standard deviations) and baseline group differences (chi-square and *t*-tests)

	PMTO *M* (SD)	BPT *M* (SD)	*t*	*p*	*Contrasts*
Demographics					
Household income[a]	63,7 (43.5)	67,4 (41.0)	0.6	0.543	ns.
Parent education[b]	2.2 (0.7)	2.4 (0.8)	3.1	0.002**	PMTO < BPT
Parent age	37.4 (6.3)	35.3 (6.1)	3.1	0.002**	PMTO > BPT
Cumulative risk	2.1 (1.6)	1.7 (1.8)	2.2	0.030**	PMTO > BPT
Child characteristics					
ECBI[c]	124.9 (27.9)	134.9 (31.2)	4.1	0.000***	PMTO > BPT
Child age	8.6 (2.4)	7.3 (2.3)	4.7	0.000***	PMTO > BPT
Dichotomized demographics	Percent (%)	Percent (%)			*Chi-square*
Single parents	32.8%	31.9%			ns.
Non-western[b]	8%	6%			ns.
Child gender	64% (boys)	68% (boys)			ns.
N	137	216			

[a] Household income in USD divided by 1000. [b] Non-Western immigrant

[b] Parent education level scale ranging from 1 (elementary school) to 4 (higher university degree)

[c] Eyberg Child Behavior Inventory—intensity scale (ECBI) 36 item version (raw scores)

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

help from school-based psychological services, 21 families received counseling from public health nurses, 22 families received counseling from public social workers in Norway's welfare services, 2 families received behavioral counseling, 2 families received Marte Meo (Aarts 2000), and 4 families received other treatments (Kjøbli et al. 2013; Kjøbli and Ogden 2012). Overall, the regular care given to the comparison group varied in its content and intensity, and none of the participants in this group received other evidence-based treatments.

Measures

Children's externalizing behavior problems were measured with a 22-item version of the Eyberg Child Behavior Intensity scale (ECBI). We used this abbreviated version of the ECBI because previous studies have indicated that brief versions of the ECBI have better psychometric properties than the original scale (Hukkelberg et al. 2016). We created a child behavior problem latent construct, ECBI, based on three parcels of sum scores taken from the ECBI scale: inattentive behavior, 4 items ("Easily gets distracted", "Has problems with concentration"); oppositional behavior, 10 items ("Does not follow rules without threat of punishment", "Argues about rules"); and conduct problems, 8 items ("argues with similarly aged friends", "destroys things"). Because of sample power issues and to maintain accurate identification, we used parcels in our measurement model, which is considered a better alternative than using the observed sum scores as outcomes (Rhemtulla 2016).

To measure the quantity of exposure to family risk, we constructed a cumulative risk index combining nine different family social and economic resources with parental health, as shown in Table 2. These resource indicators have been previously acknowledged as risk factors for the development of behavior problems and have been shown to affect health care outcomes (Alegría et al. 2011; Kjeldsen et al. 2014; Moffitt and Scott 2009; Narayanan and Nærde 2016; Piotrowska et al. 2015; Waldfogel et al. 2010; Zilberstein 2016). Our cumulative risk index had indicators similar to those used in the Sameroff cumulative risk index (Sameroff et al. 1987). Cumulative risk indexes are usually calculated by summing the number of dichotomized risk factors (Evans et al. 2013; Trentacosta et al. 2008). We computed our cumulative risk index using the nine dichotomized indicators shown in Table 2. Parents scored an indicator as "1" if it was present and "0" if it was absent. Cut-offs were based on previously established limits and/or corresponded to previous cumulative risk research or population-validated numbers (for mental health; Evans et al. 2013; Reedtz et al. 2008; Trentacosta et al. 2008). The cumulative risk indicator was based on parent-reported information assessed before treatment.

Table 2 Cumulative risk indicators, definitions, and percentages complying with sample criteria

Indicator:	Description of criteria	%
OECD poor[a]	OECD 50% of median income	25.8
Low education	Did not finish upper secondary school	23.5
Unemployed	Financial unemployment support	6.2
Non-Western immigrant	From Eastern Europe or south of the equator	6.8
Single parent	One caregiver in the family	32.3
Young caregiver	Parent ≤ 21 years of age	10.2
Caregiver ratio	Ratio ≤ 0.5 adults per child	28.3
Somatic health[b]	Cut off at ≥3	18.1
Mental health	Average score cut off ≥ 2	35.4

[a] Organization for Economic Co-operation and Development (OECD) equivalence poverty scale

[b] One-item scale ranging from 1 (excellent health status) to 5 (poor health status)

[c] Adjusted SCL-5 scores ranging between 1 and 5, a higher score indicates more anxiety and mental distress

The poverty measure, OECD poor, was based on the OECD equivalence scale (Organisation for Economic Co-operation and Development 2016); families with less than 50% of the median net income were coded as poor. Thus, the different poor cut-offs were calculated as a function of family constellation and the 2008 average population median net income of approximately 27,500 USD. The low education variable was a dichotomized variable derived from a categorical education level variable counting; 1 = elementary school; 2 = upper secondary school; 3 lower university degree; and 4 higher university degree (>4 years). Parents that scored 1 on the education level variable were coded as 1 in the low education cumulative risk indicator. Parents' mental health was measured with the Symptom Checklist 5 (SCL-5), which measures anxiety and depression. SCL-5 is a (very) short-form of the SCL-25 mental health index which is derived from the SCL-90 psychopathology rating scale (Derogatis 1992; Strand et al. 2003). In a Norwegian population-based sample, Strand et al. (2003) found that the correlation between SCL-5 and SCL-25 was 0.91. SCL-5 risk cut-offs used were based on numbers validated and normed in a Norwegian study (Tambs and Moum 1993). Cronbach's alpha for the SCL-5 was 0.88. The variable measuring parents' somatic status was a one-item and non-validated scale ranging from 1 ("Excellent health status") to 5 ("Very poor health status"). Table 3 displays the bivariate correlations among the cumulative risk indicators.

Risk seemed to cluster in our sample. Hence, being poor was significantly correlated with all the other nine risk factors, as shown in Table 3. Moreover, the cumulative risk index was significantly correlated with higher baseline

Table 3 Bivariate correlations between family factors and child behavior problems

Variable:	1	2	3	4	5	6	7	8	9
1. OECD poor[a]	1								
2. Low education	0.25***	1							
3. Unemployed	0.22***	0.13*	1						
4. Non-Western	0.15**	−0.04	0.07	1					
5. Single parent	0.44***	0.15**	0.17***	0.05	1				
6. Caregiver ratio	0.14**	0.26***	−0.01	0.06	0.11*	1			
7. Young parent	0.41***	0.17***	0.18***	0.01	0.60***	0.02	1		
8. Somatic health	0.16**	0.18***	0.03	0.14*	0.13*	0.01	0.05	1	
9. Mental health	0.20***	0.13*	0.10	0.11*	0.07	0.04	0.09	0.34***	1

[a] Organization for economic Co-operation and development (OECD) equivalence poverty scale

levels of behavior problems ($r = 0.17$, $p = 0.002$), meaning that children who were exposed to cumulative risk in their families had, on average, higher levels of behavior problems before treatment.

Data Analyses

Children were the unit of analysis in this study. Social gradients in the outcomes of PMTO, namely, whether cumulative risk moderates treatment effects, were examined with autoregressive SEM analysis using Mplus 7 (Muthén and Muthén 2012). We ran the models as intent-to-treat analyses to examine intervention effects across treatment conditions using the three time-points: T1, T2, and T3. The outcomes in our models showed changes in the children's behavior problems from baseline levels. The SEM models are displayed in Fig. 1 and Fig. 2. The treatment variable shows the treatment effects for families that scored 0 on the cumulative risk variable, the cumulative risk variable displays the effects of exposure to cumulative risk for the regular care group, and the interaction term cumulative risk * treatment shows the treatment effects for the families who were exposed to cumulative risk in the PMTO groups.

The measurement model, displayed in Figs. 1 and 2, shows the standardized factor loadings and correlated error terms for the three parcels constituting our latent outcome variable (ECBI). The error terms in the three parcels were correlated across the time points. These unanalyzed associations represent shared sources of variability over and above the latent factors. To investigate whether there were any sample-specific differences due to differences in cumulative risk and the intensity of treatment (i.e., dosage and comprehensiveness), we ran auto-regressive multi-group SEM models analyzing the BPT and the PMTO interventions separately. We used several goodness-of-fit indexes to evaluate our theoretical model fit: chi-square statistics, the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root-mean-square error of approximation (RMSEA; see table and figure notes). Moreover, we checked our data for potential outlier observations through a visual inspection of residual plots and the estimation of a five percent trimmed mean in the outcome variable. Neither of these procedures indicated that the effects of outliers biased our results.

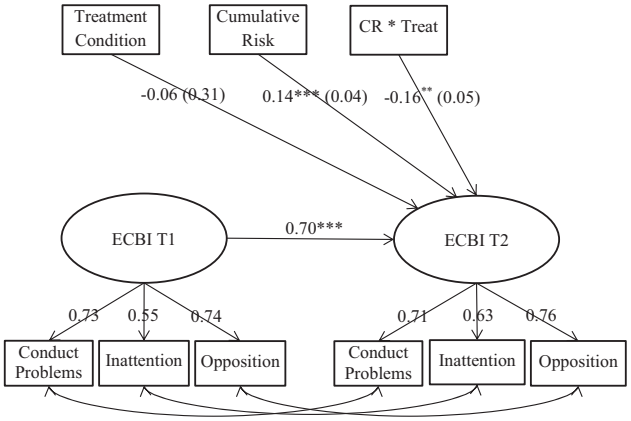

Fig. 1 Cumulative risk and child behavior change in PMTO interventions vs. regular care. Autoregressive SEM analysis, posttreatment (T2) regressed on pretreatment behavior (T1). *Note:* Eyberg Child Behavior Inventory—intensity scale (ECBI). Interaction variable Cumulative Risk multiplied by Treatment Condition (CR * Treat). Coefficients were standardized on *Y* (equals Cohen's *d*), *standard error* is displayed in parentheses. Model fit information: $X^2(\text{df}) = 27.0$ (20), CFI = 0.99 TLI = 0.99 RMSEA = 0.06. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Additionally, we performed several sensitivity tests to inspect functional forms in our data and to address potential rival conclusions. We also included child age and gender as covariates in our analyses and tested non-linear patterns in our data. Moreover, we evaluated the family risk factors in the cumulative risk index using independent-additive models to investigate the unique effects of each cumulative risk indicator. Finally, we partialled out the control group families that received no treatment to determine whether the cumulative risk comparison group estimates

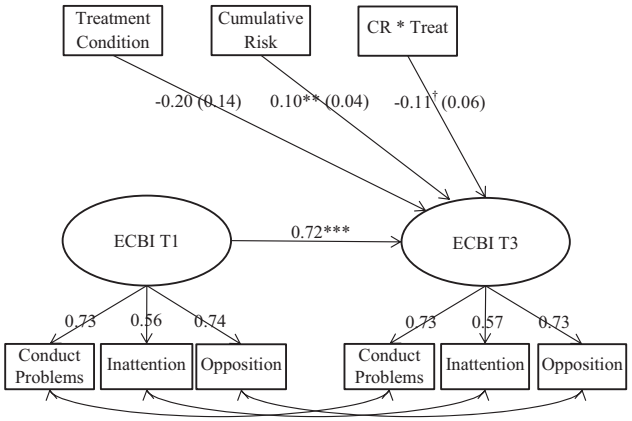

Fig. 2 Cumulative risk and child behavior change in PMTO interventions vs. regular care. Autoregressive SEM analysis, follow-up (T3) regressed on pretreatment behavior (T1). *Note:* Eyberg Child Behavior Inventory—intensity scale (ECBI). Interaction variable Cumulative Risk multiplied by Treatment Condition (CR * Treat). Coefficients were standardized on *Y* (equals Cohen's *d*); standard errors are displayed in parentheses. Model fit information: X^{2}(df) = 36.3 (20), CFI = 0.98 TLI = 0.97 RMSEA = 0.05. ***$p < 0.001$; ** $p < 0.01$; *$p < 0.05$; † = 0.059

were influenced by receiving active treatment or no treatment.

There was little attrition in our sample. Of the 353 participating families, 301 (85%) completed T2, and 275 (78%) completed T3. When comparing the attrition group with the completers in each trial, few differences in intake characteristics emerged (for more details see Kjøbli et al. 2013; Kjøbli and Ogden 2012). A missing data analysis, Little's MCAR test, indicated that the missing data were missing completely at random. Thus, we modeled the data using full-information maximum likelihood, which uses all the available information from the observed data to handle missing data (Wothke 2000).

Results

The post treatment effects (T2) for the pooled sample of PMTO interventions are displayed in Fig. 1. The results showed that the children in the PMTO group from families with one additional cumulative risk generally experienced more benefit from treatment; in T2, behavior problems were reduced by an average of 16% of a standard deviation for each accumulated risk ($\beta = -0.16$, $p < 0.01$; results were standardized on *Y* only). Conversely, for the regular care group, scoring higher on cumulative risk was significantly associated with lower treatment benefits; this group displayed increased levels of problem behavior in T2 ($\beta =$ 0.14, $p = < 0.001$). This implies that one additional cumulative risk entailed an increase of 14% of a standard deviation in children's behavior problems in T2. In Fig. 2, the pooled PMTO group results were not significantly replicated at T3 ($\beta = -0.11$, $p = 0.06$); however, the coefficient had a considerable size in a similar direction as in T2. For the regular care group, the T2 results were replicated at T3, ($\beta = 0.104$, $p < 0.01$). Overall, the model fit was within an acceptable range for the fit indexes in all the estimated models (Hu and Bentler 1999); see table notes. The factor loadings from the parcels in all the SEM measurement models were >0.50; see Figs. 1 and 2. Moreover, for the Fig. 1 results, we computed the simple slopes and calculated the regions of significance; see Fig. 3 (Preacher et al. 2006). Differences between the groups were significant for cumulative risk scores above 0.9, meaning that group differences between those who received parent training and those who received regular care were significant for families with one or more risks.

Next, we examined whether there were sample-specific differences in the associations between cumulative risk and changes in behavior problems and whether the treatment effects differed for the families with cumulative risks according to treatment intensity. The path coefficients revealed such differences; see Table 4. The PMTO intervention seemed to be particularly effective for children from families with cumulative risks at both T2 ($\beta = -0.33$, $p = < 0.001$) and T3 ($\beta = -0.30$, $p = < 0.001$). The low-intensity BPT intervention results did not reveal any significant changes in the treatment effects for the families with cumulative risk. In both samples, the children who received regular care experienced significant increases in behavior problems at all time points; see Table 4. We computed the region of significance for the path T1 → T2 in the PMTO sample. We found that group difference between the PMTO group and comparison group was significant for cumulative risk scores of 1.7 and higher (Fig. 4; Preacher et al. 2006).

Additionally, we performed several sensitivity tests to gauge the robustness of our conclusions. First, we examined

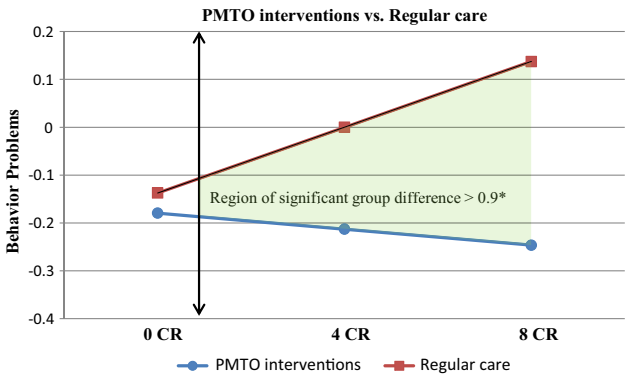

Fig. 3 Simple slopes and region of significance for the interaction between treatment conditions and cumulative risk, child behavior change at T2

Table 4 Autoregressive multi-group SEM analysis displaying separate path coefficients for the BPT and the PMTO samples

Parameter	BPT		PMTO	
	β	SE	*β*	SE
Model 1[a]: ECBI T2				
ECBI T1	0.69***	0.05	0.69***	0.07
Treatment	−0.22	0.16	0.26	0.24
Cumulative risk	0.12**	0.04	0.19**	0.07
Treat * CR	0.08[ns]	0.07	−0.33***	0.09
Model2[b]: ECBI T3				
ECBI T1	0.72***	0.05	0.68***	0.07
Treatment	−0.29	0.16	−0.03	0.25
Cumulative Risk	0.08	0.04	0.18*	0.08
Treat * CR	0.00	0.07	−0.30**	0.10

Eyberg Child Behavior Inventory—intensity scale (ECBI). Interaction variable Cumulative Risk multiplied by Treatment Condition (CR * Treat). [a] Model 1 ECBI T2 regressed on T1, model fit information: X^2(df) = 55.9 (48), CFI = 0.99 TLI = 0.99 RMSEA = 0.05

[b] Model 2 ECBI T3 regressed on T1, model fit information: $X^{2/}$(df) = 59.0 (48), CFI = 0.99 TLI = 0.98 RMSEA = 0.04

Coefficients were standardized on *Y* (equals Cohen's *d*)

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

whether there were non-linear patterns in the cumulative risk associations with changes in behavior problems. We tested both a cubic parameterization of cumulative risk and threshold effects for families with between 2 and 5 cumulative risks. No threshold or non-linear patterns were significantly different from 0 (results available upon request). Moreover, we tested whether cumulative risk had unequal effects according to the child's gender and age. Both variables were entered into the analysis as covariates and into a three-way interaction term with treatment and cumulative risk (results available upon request). No significant effects of gender or age emerged. Furthermore, we tested the nine cumulative risk factors singularly in independent-additive models to test for unique predictive validity. We found that no significant results emerged from these analyses. Finally, 35% of our control group cases received no treatment. We suspected that these children and families biased our estimates, and we ran additional analyses without these 63 families. The results were similar to those of the original models in terms of both the coefficient sizes and significance levels, and the full sample was thus included in our final analysis.

Discussion

In this study, we extended the literature on health care inequality in behavior problem interventions by examining the relationships between family cumulative risk and

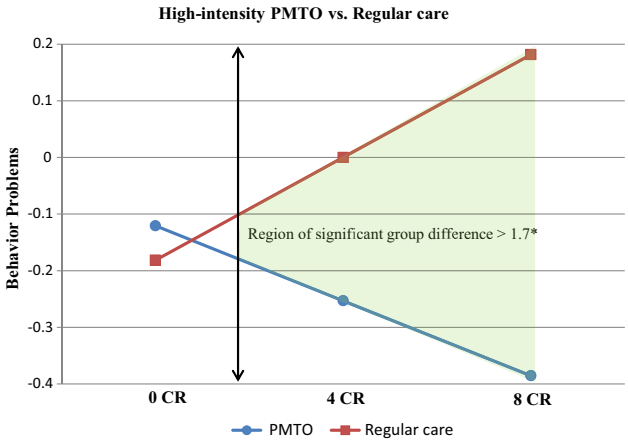

Fig. 4 Simple slopes and region of significance for the interaction between treatment condition and cumulative risk for the high-intensity PMTO sample, child behavior change at T2

treatment outcomes in evidence-based parent training and regular care. We also examined cumulative risk associations in a case comparison between low-intensity and high-intensity PMTO. First, we found that exposure to cumulative risk differentially moderated the treatment effects of PMTO interventions and regular care, as reflected by the opposite social gradients of the changes in the children's behavior problems. The children who received PMTO interventions and were exposed to one or more cumulative risks experienced compensatory effects, meaning that the children from families with low amounts of resources experienced greater reductions in their behavior problems than the children from high-resource families. Conversely, the regular care group exposed to equal levels of risk experienced more behavior problems over time, indicating that the children from low-resource families had poorer treatment outcomes with regular care than the children from high-resource families. Second, we found that the families with cumulative risk benefitted differently according to the intensity of the PMTO treatments; the children who were exposed to cumulative risks experienced vast improvement with high-intensity PMTO. Thus, cumulative risk produced social gradients in treatment effects according to both the treatment conditions and the treatment intensity.

The effects of cumulative risk seem to be linear, and we conclude that it is the sheer number of risk factors that changes the treatment effects rather than differences in treatment effects below and above a certain threshold. Previous research regarding cumulative risk and the effects of parent training is both limited and inconsistent (Liaw and Brooks-Gunn 1994; Stolk et al. 2008). In this sample of at-risk children, our results revealed opposite social gradients, indicating that children from families with cumulative risk were highly receptive to the type of care provided by the Norwegian health care system.

The finding of opposite social gradients for the treatment effects, related to the PMTO interventions and regular care, implies that type of treatment may either create or reduce health care inequality among low-resource families receiving help for their child's behavior problem. Hence, there seem to be different mechanisms related to changes in health care outcomes for those who receive parent training and those who receive regular care. Unfortunately, we have limited knowledge about the contents of regular care. However, we know that the differences between PMTO interventions and regular care are rooted in the differences between structured, curriculum-grounded, evidence-based parent training and the more unstructured parent counseling provided in regular care. Hence, the mechanisms related to inequality in regular mental health care and parent training interventions, such as network mechanisms, sociocultural mechanisms and practical mechanisms (Alegría et al. 2011; Spencer and Grace 2016; Zilberstein 2016), applied more to the regular care group in our study. It might be that these mechanisms come into play more when the mode of treatment is less structured and does not explicitly target parenting style. Moreover, this could indicate that there was more room for high-resource parents to influence treatment content—and thus their children's outcomes—under the less structured health care conditions.

Conversely, the compensatory effects of PMTO interventions for low-resource families indicate that other beneficial mechanisms were operating within these structured treatment conditions. It might be that the children from low-resource families were more exposed to disrupted parenting practices and that systematic parent training was more adapted to their pre-intervention skills and family climate. Moreover, separate analysis of our preventive BPT and high-risk PMTO samples revealed that the compensatory effects were more prominent under the latter treatment condition. In the BPT sample, the families with both low and high cumulative risk experienced positive changes in their children, whereas in the more intensive PMTO treatment group, the families with high cumulative risk experienced a vast improvement. The reduction in their children's behavior problem levels had the effect size of approximately 30% of a standard deviation change per level increase in cumulative risk, which underpins this argument. It seems that providing more intensive treatment to the more troubled families exaggerates the compensatory effect mechanisms. Thus, there is probably interplay between compensatory mechanisms, such as parent's pre-intervention parenting skills, the etiology of child behavior problems, and readiness for change, that produces favorable outcomes for the families with cumulative risk in the PMTO treatment. However, more research is needed to reveal the mediational relationships behind these compensatory patterns.

Behavior problems in childhood contribute to social gradients in child well-being, but behavior problems are also partly the products of social disparities. Low-resource backgrounds have been found to increase the risk that a child will experience behavior problems (Piotrowska et al. 2015), and behavior problems themselves have negative long-term developmental impacts, as children are exposed to multiple threats to their well-being later in life (Moffitt et al. 2002). From a mental health care perspective, this underscores the need to provide effective care for this vulnerable group of children and their families. If we fail to do this, mental health care interventions aimed at behavior problems will certainly produce health care inequality and exacerbate existing health inequality (and ultimately social inequality) among low-resource populations. The opposing social gradients we found in care for behavior problems support this argument; the type of treatment provided can either produce or reduce health care inequality.

The steady negative development displayed by children from low-resource backgrounds in the regular care group is thus consistent with the theory of family resources as a "fundamental cause" structuring flexible assets when coping with children with behavior problems (Link and Phelan 1995; Muntaner et al. 2013). This negative development is also consistent with behavior problems risk theory, which postulates that social risk, in the form of family resources, intensifies the development of behavior problems (Mazza et al. 2016). When helping children at risk for such negative development, evidence-based parent training interventions seem to be an efficient strategy for counteracting health care inequality and the lack of family resources as a "fundamental cause" and thus for effectively preventing and altering negative developmental trajectories for children from low-resource backgrounds. It has been postulated that evidence-based treatments can promote equality in care (Cochrane 2004; Kristiansen and Mooney 2004). Regarding the outcomes of treatment, our results support this postulation.

Limitations

In this study, we had the advantage of using experimental data gathered in the Norwegian regular care system for children with behavior problems. However, several limitations should be considered. Admittedly, combining risk factors into a cumulative risk index will, to some extent, obscure the etiology of social risk. We also applied an equal weights assumption when we pooled the dichotomized risk factors, and we combined risk factors from different domains, such as family demographics and parental health. However, we can offer insights into how different amounts of risk exposure affect health care outcomes for children with behavior problems. The predictive validity of our

cumulative risk index supports this notion. Moreover, although we adhered to previously established cut-offs in cumulative risk research, one may still argue that the process of dichotomization inflicts arbitrary limits on risk factors. Nevertheless, evidence is very limited regarding cumulative risks and the outcomes of care for behavior problems. Thus, the effects of cumulative risk must be replicated in other samples and contexts. Moreover, the RCT design evaluating treatment intervention packages, such as PMTO, did not allow us to address which treatment components that produced the compensatory effects of PMTO (Collins 2014). Other approaches, such as factorial designs, might be more appropriate for elaborating further on the change mechanisms related to cumulative risk that is at work in parent training.

Acknowledgements Asgeir R. Olseth, Bjørn Arild Kristiansen, Elisabeth Askeland, Henrik Dae Zachrisson, John Kjøbli, Kristian Heggebø, Silje Hukkelberg, Terje Ogden, Terje Christiansen, Torkild Hovde Lyngstad, and all you in the Department of research and the Department of child program development at the Norwegian center for child behavior development.

Authors' Contributions T.T.: wrote the manuscript and performed the data analyses. J.K.: designed the study, administered the data collection, and commented on the manuscript.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no competing interests.

Ethical Approval All procedures performed in this study were in accordance with the ethical standards of an ethical review board; The Norwegian National Committee for Research Ethics, Region South; The Norwegian Social Science Data Services; and with the 1964 Helsinki declaration and its later amendments.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Aarts, M. (2000). *Marte Meo: Grundbog [Marte Meo: Basic Manual]*. Copenhagen, Denmark: Aarts Production.

Alegría, M., Pescosolido, B. A., Williams, S., & Canino, G. (2011). Culture, race/ethnicity and disparities: Fleshing out the sociocultural framework for health services disparities. In B. A. Pescosolido, J. K. Martin, J. D. McLeod, A. Rogers (Eds.), *Handbook of the sociology of health, illness, and healing* (pp. 363–382). New York, US: Springer Science and Business Media.

Baydar, N., Reid, M. J., & Webster-Stratton, C. (2003). The role of mental health factors and program engagement in the effectiveness of a preventive parenting program for Head Start mothers. *Child development*, *74*(5), 1433–1453. doi:10.1111/1467-8624.00616.

Berlin, L. J., Brooks-Gunn, J., McCarton, C., & McCormick, M. C. (1998). The effectiveness of early intervention: Examining risk factors and pathways to enhanced development. *Preventive Medicine*, *27*(2), 238–245.

Bøe, T., Øverland, S., Lundervold, A. J., & Hysing, M. (2012). Socioeconomic status and children's mental health: Results from the Bergen Child Study. *Social Psychiatry and Psychiatric Epidemiology*, *47*(10), 1557–1566. doi:10.1007/s00127-011-0462-9.

Cochrane, A. L. (2004). *Effectiveness and efficiency: Random reflections on health services*. London, UK: Royal Society of Medicine.

Collins, L. M. (2014). Optimizing family intervention programs: The multiphase optimization strategy (MOST). In S. M. McHale, P. Amato, A. Booth (Ed), *Emerging methods in family research*. 4th edition (pp. 231–244). Switzerland: National Symposium on Family Issues.

Conger, R. D., Conger, K. J., Elder, G. H., Lorenz, F. O., Simons, R. L., & Whitbeck, L. B. (1992). A family process model of economic hardship and adjustment of early adolescent boys. *Child Development*, *63*(3), 526–541. doi:10.1111/j.1467-8624.1992.tb01644.x.

Deković, M., Slagt, M. I., Asscher, J. J., Boendermaker, L., Eichelsheim, V. I., & Prinzie, P. (2011). Effects of early prevention programs on adult criminal offending: A meta-analysis. *Clinical Psychology Review*, *31*(4), 532–544.

Derogatis, L. R. (1992). *SCL-90-R: Administration, scoring and procedures manual for the R (evised) version and other instruments of the psychopathology rating scale series*. 2nd ed. Towson Md: Clinical Psychometric Research Inc.

DiClemente, C. C., & Velasquez, M. M. (2002). Motivational interviewing and the stages of change. In W. R. Miller, S. Rollnick (Eds.), *Motivational interviewing: Preparing people for change*. 2nd ed (pp. 201–216). New York, US: Guilford Press.

Elstad, J., & Stefansen, K. (2014). Social variations in perceived parenting styles among Norwegian adolescents. *Child Indicators Research*, *7*(3), 649–670. doi:10.1007/s12187-014-9239-5.

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin*, *139*(6), 1342–1396.

Forgatch, M. S., & Patterson, G. R. (2010). Parent Management Training—Oregon Model: An intervention for antisocial behavior in children and adolescents. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* 2nd ed. (pp. 159–178). New York, NY: Guilford Press.

Gardner, F., Connell, A., Trentacosta, C. J., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2009). Moderators of outcome in a brief family-centered intervention for preventing early problem behavior. *Journal of Consulting and Clinical Psychology*, *77*(3), 543–553.

Gengler, A. M. (2014). "I want you to save my kid!" illness management strategies, access, and inequality at an elite university research hospital. *Journal of Health and Social Behavior*, *55*(3), 342–359.

Gillies, V. (2006). *Marginalised mothers: Exploring working class experiences of parenting*. London, UK: Routledge.

Hu, Lt, & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria vs. new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Hukkelberg, S. S., Reedtz, C., & Kjøbli, J. (2016). Construct validity of the eyberg child behavior inventory (ECBI). *European Journal*

of Psychological Assessment. Published online, 1–8. doi:10.1027/1015-5759/a000316.

Kjeldsen, A., Janson, H., Stoolmiller, M., Torgersen, L., & Mathiesen, K. S. (2014). Externalising behaviour from infancy to mid-adolescence: Latent profiles and early predictors. *Journal of Applied Developmental Psychology*, *35*(1), 25–34. doi:10.1016/j.appdev.2013.11.003.

Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: reducing child conduct problems in real-world settings. *Behaviour Research and Therapy*, *51*(3), 113–121.

Kjøbli, J., & Ogden, T. (2012). A randomized effectiveness trial of brief parent training in primary care settings. *Prevention Science*, *13*(6), 616–626.

Kristiansen, I. S., & Mooney, G. (2004). *Evidence-based medicine: In its place*. Abdington, UK: Routledge.

Leijten, P., Raaijmakers, M. A. J., de Castro, B. O., & Matthys, W. (2013). Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. *Journal of Clinical Child & Adolescent Psychology*, *42*(3), 384–392. doi:10.1080/15374416.2013.769169.

Liaw, F.-R, & Brooks-Gunn, J. (1994). Cumulative familial risks and low-birthweight children's cognitive and behavioral development. *Journal of Clinical Child Psychology*, *23*(4), 360–272.

Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*, *35*, 80–94.

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, *26*(1), 86–104. doi:10.1016/j.cpr.2005.07.004.

Maughan, B., Gray, G., & Rutter, M. (1985). Reading retardation and antisocial behaviour: A follow-up into employment. *Journal of Child Psychology and Psychiatry*, *26*(5), 741–758. doi:10.1111/j.1469-7610.1985.tb00588.x.

Mazza, J. R. S. E., Boivin, M., Tremblay, R. E., Michel, G., Salla, J., Lambert, J., & Côté, S. M. (2016). Poverty and behavior problems trajectories from 1.5 to 8 years of age: Is the gap widening between poor and non-poor children? *Social Psychiatry and Psychiatric Epidemiology*, *51*(8), 1083–1092. doi:10.1007/s00127-016-1252-1.

Moffitt, T. E., & Scott, S. (2009). Conduct disorders of childhood and adolescence. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, A. Thapar (Eds.), *Rutter's child and adolescent psychiatry*. 5th ed. (pp. 543–564). Oxford, UK: Blackwell.

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology*, *14*(01), 179–207.

Muntaner, C., Ng, E., Vanroelen, C., Christ, S., & Eaton, W. W. (2013). Social stratification, social closure, and social class as determinants of mental health disparities. In S. Aneshensel, J. C. Phelan (Eds), *Handbook of the sociology of mental health* (pp. 205–227). New York, US: Kluwer Academic/ Plenum Publishers.

Muthén, B. O., & Muthén, L. K. (2012). Mplus version 7. Retrieved from https://www.statmodel.com/.

Narayanan, M. K., & Nærde, A. (2016). Associations between maternal and paternal depressive symptoms and early child behavior problems: Testing a mutually adjusted prospective longitudinal model. *Journal of Affective Disorders*, *196*, 181–189. doi:10.1016/j.jad.2016.02.020.

Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of Parent Management Training in Norway: a randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology*, *76*(4), 607–621.

Organisation for Economic Co-operation and Development. (2016). What are equivalence scales? http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf.

Pescosolido, B. A., Boyer, C. A., & Medina, T. R. (2013). The social dynamics of responding to mental health problems. In S. Aneshensel, J. C. Phelan (Eds), *Handbook of the sociology of mental health* (pp. 505–524). New York, US: Kluwer Academic/ Plenum Publishers.

Piotrowska, P. J., Stride, C. B., Croft, S. E., & Rowe, R. (2015). Socioeconomic status and antisocial behaviour among children and adolescents: A systematic review and meta-analysis. *Clinical Psychology Review*, *35*, 47–55. doi:10.1016/j.cpr.2014.11.003.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437–448.

Reedtz, C., Bertelsen, B., Lurie, J. I. M., Handegård, B. H., Clifford, G., & Mørch, W. T. (2008). Eyberg child behavior inventory (ECBI): Norwegian norms to identify conduct problems in children. *Scandinavian Journal of Psychology*, *49*(1), 31–38. doi:10.1111/j.1467-9450.2007.00621.x.

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems-a meta-analytic review. *Journal of Child Psychology and Psychiatry*, *47*(1), 99–111. doi:10.1111/j.1469-7610.2005.01544.x.

Rhemtulla, M. (2016). Population performance of SEM parceling strategies under measurement and structural model misspecification. *Psychological Methods*, *21*(3), 348–368. doi:10.1037/met0000072.

Rutter, M. (2000). Psychosocial influences: Critiques, findings, and research needs. *Development and Psychopathology*, *12*(03), 375–405.

Rutter, M., Tizard, J., & Withmore, K. (1970). *Education, health and behavior*. London, UK: Longman.

Sameroff, A. J., Seifer, R., Baldwin, A., & Baldwin, C. (1993). Stability of intelligence from preschool to adolescence: The influence of social and family risk factors. *Child development*, *64*(1), 80–97.

Sameroff, A. J., Seifer, R., Zax, M., & Barocas, R. (1987). Early indicators of developmental risk: Rochester Longitudinal Study. *Schizophrenia Bulletin*, *13*(3), 383–394.

Sameroff, A. J., Bartko, W. T., Baldwin, A., Baldwin, C., & Seifer, R. (1998). Family and social influences on the development of child competence. In M. Lewis, C. Feiring (Eds.), *Families, risk, and competence* (pp. 161–185). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Seifer, R., Sameroff, A. J., Baldwin, C. P., & Baldwin, A. (1992). Child and family factors that ameliorate risk between 4 and 13 years of age. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*(5), 893–903.

Smith, K. P., & Christakis, N. A. (2008). Social networks and health. *Annual Review of Sociology*, *34*, 405–429. doi:10.1146/annurev.soc.34.040507.134601.

Solholm, R., Kjøbli, J., & Christiansen, T. (2013). Early initiatives for children at risk—development of a program for the prevention and treatment of behavior problems in primary services. *Prevention Science*, *14*(6), 535–544. doi:10.1007/s11121-012-0334-x.

Spencer, K. L., & Grace, M. (2016). Social foundations of healthcare inequality and treatment bias. *Annual Review of Sociology*, *42*, 101–120. doi:10.1146/annurev-soc-081715-074226.

Sroufe, L. A., Egeland, B., Carlson, E. A., & Collins, W. A. (2009). *The development of the person: The Minnesota study of risk and adaptation from birth to adulthood*. New York, US: Guilford Press.

Stolk, M. N., Mesman, J., van Zeijl, J., Alink, L. R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Koot, H. M. (2008). Early parenting intervention: Family risk and first-time parenting related to intervention effectiveness. *Journal of Child and Family Studies*, *17*(1), 55–83.

Storvoll, E. E. (1997). *Barn og unge med alvorlige atferdsvansker: Hvem er de, og hvilken hjelp blir de tilbudt* [Children and youth with serious behavior difficulties: Who are they, and what kind of help are they offered?]. Oslo, Norway: Norwegian Institute for Research on Development, Welfare, and Aging.

Strand, B. H., Dalgard, O. S., Tambs, K., & Rognerud, M. (2003). Measuring the mental health status of the Norwegian population: a comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). *Nordic Journal of Psychiatry*, *57*(2), 113–118.

Tambs, K., & Moum, T. (1993). How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatrica Scandinavica*, *87*(5), 364–367.

Thoits, P. A. (2011). Mechanisms linking social ties and support to physical and mental health. *Journal of Health and Social Behavior*, *52*(2), 145–161.

Trentacosta, C. J., Hyde, L. W., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. (2008). The relations among cumulative risk, parenting, and behavior problems during early childhood. *Journal of Child Psychology and Psychiatry*, *49*(11), 1211–1219.

Tømmeraas, T. (2016). Social gradients and participant characteristics in child behavior problem interventions. *Children and Youth Services Review*, *70*, 57–64.

Waldfogel, J., Craigie, T.-A., & Brooks-Gunn, J. (2010). Fragile families and child wellbeing. *The Future of Children*, *20*(2), 87–112.

Weininger, E. B., & Lareau, A. (2003). Translating Bourdieu into the American context: the question of social class and family-school relations. *Poetics*, *31*(5), 375–402.

World Health Organization. (2003). *Caring for children and adolescents with mental disorders: Setting WHO directions*. Geneva, Switzerland: WHO.

Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. U. Schnabel, J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* 269–281, (219–240). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Zilberstein, K. (2016). Soapbox: Class matters in parenting interventions. *Clinical Child Psychology and Psychiatry*, *21*(3), 359–367.

III

IV

Adm Policy Ment Health (2017) 44:203–216
DOI 10.1007/s10488-015-0712-3

ORIGINAL ARTICLE

Is There a Scale-up Penalty? Testing Behavioral Change in the Scaling up of Parent Management Training in Norway

Truls Tommeraas[1] · **Terje Ogden**[1]

Published online: 29 December 2015

Abstract In the present study, the scaling up of Parent Management Training, Oregon Model (PMTO) in Norway was examined by investigating how large-scale dissemination affected the composition of the target group and the service providers by comparing child behavioral outcomes in the effectiveness and dissemination phases of implementation. Despite the larger heterogeneity of the service providers and the intake characteristics of the target group, which are contrary to the expectations that were derived from the literature, no attenuation of program effects was detected when scaling up PMTO. In Norway, a long-term-funded centralized center, combined with an active implementation strategy, seems to have affected the quality of PMTO delivered system-wide in services for children with behavior problems.

Keywords Implementation · Large-scale dissemination · Testing evidence-based interventions

Introduction

Recently, many family-focused prevention and treatment programs have been scaled up and introduced in new settings. However, many of these programs have a limited impact because the implementation quality is lacking or it is not sustained over time (I.O.M, 2014). Moreover, research regarding programs that are disseminated on a large scale is limited (Elliott and Mihalic 2004; McHugh and Barlow 2010; Ogden and Fixsen 2014; Ogden et al. 2005). A substantial number of parenting programs have been tested in efficacy or effectiveness studies, but the outcomes of large-scale dissemination have rarely been studied systematically. However, it is a widely-held view that the positive effects of evidence-based parenting programs attenuate when they are scaled up from the effectiveness phase to the broader dissemination phase (Dodge 2001; Kellam and Langevin 2003; Welsh et al. 2010). In going to scale, effective programs are assigned scale-up penalties due to challenges in the implementation process (Welsh et al. 2010), although this assumption has rarely been empirically tested. In the present study, we conceptualized the scale-up penalty as a reduction of behavioral changes in large-scale dissemination, and we examined potential scale-up penalties when PMTO was scaled up in Norway.

Previous Research

When a program reaches the phase of large-scale dissemination, the implementation process increases in complexity (Dodge 2001; Kellam and Langevin 2003). Welsh et al. (2010) pinpoint these challenges: "With the program expanded beyond its tightly controlled environs and no longer under the immediate control of its chief architects and well trained clinical staff, how can critical implementation and process issues that underlie the program's successful delivery be maintained?" The concept of the scale-up penalty has been used to describe decreases in program effects when programs move from the effectiveness phase to the large-scale dissemination phase (Welsh et al. 2010). Across three cost-benefit studies, parent-

✉ Truls Tommeraas
trulstom@atferdssenteret.no

Terje Ogden
terje.ogden@atferdssenteret.no

[1] Norwegian Center for Child Behavior Development, P.O. Box 7053, Majorstuen, 0306 Oslo, Norway

training programs were found to be beneficial even if they were assigned scale-up-penalties (Aos et al. 2001; Donohue and Siegleman 1998; Greenwood 1998). Greenwood (1998) assigned a scale-up-penalty of 40 % to a PMTO program, which is the program of focus in this study. Based on the three studies, Welsh et al. (2010) expected an attenuation of effects to occur, and they reported how scale-up penalties in parent-training programs ranged from a low of 25 % and a high of 50 %.

Many of the challenges in sustaining the program effects in large-scale dissemination are related to the barriers or obstacles that are described in the implementation literature (Ogden and Fixsen 2014). These moderators of program effects in large-scale dissemination may be categorized as (1) implementation factors, (2) the heterogeneity of service providers, and (3) the heterogeneity in target populations. *First*, the challenges regarding *implementation factors* may be related to the entire range of implementation drivers in the framework set forth by Fixsen et al (2005), such as an insufficient service infrastructure, insufficient training or supervision, a lack of technical support, and generally poor implementation (Dodge 2001; Elliott and Mihalic 2004; Kellam and Langevin 2003; Lipsey 2009; Mihalic and Irwin 2003). For instance, there may be insufficient community resources that are needed to fund the large-scale training, supervision and other expenditures that are related to sustained, system-wide implementation (Welsh et al. 2010). Furthermore, modifications due to demands for the local adaptation of programs may lead to a loss of treatment fidelity and hence to the attenuation of program effects (Elliott and Mihalic 2004; Ogden and Fixsen 2014). *Second*, the increased *heterogeneity of program or service providers* may affect the level of treatment integrity and treatment outcomes; this includes more diverse background training, motivation, clinical skills and experience among the practitioners, along with variations in the time that is set aside to practice the program (Forgatch et al. 2013; Kellam and Langevin 2003; Mihalic and Irwin 2003; Welsh et al. 2010). Other challenges to service provider systems may be the need for competent leadership by administrators who buy into the program, the management of staff turnover, and the securing of funding and organizational support (Elliott and Mihalic 2004; Welsh et al. 2010). *Third*, increased *heterogeneity in target populations* may be related to moving from homogenous populations in the efficacy and effectiveness phases to more heterogeneous target populations with less problem behavior to treat in the large-scale dissemination phase (Bonta and Andrews 2007; Dodge 2001; Kellam and Langevin 2003). There may be greater variations in the motivation of families, more comorbidity, and increased rates of non-consenting parents who do not show up for or who drop out of treatment (Welsh et al. 2010).

Based on the literature review, it seems relevant to hypothesize a scale-up penalty as a function of challenges from these three categorized levels' interactions with the local context. However, these relationships have rarely been empirically tested. Therefore, we wanted to empirically test whether there was a scale-up penalty in the process of implementing PMTO in Norway.

PMTO and Norwegian Research Findings

PMTO is a curriculum based parent-training intervention that is anchored in Patterson and colleagues' social interaction learning theory and draws on ecological and transactional principles (Dishion and Patterson 2006; Forgatch and Patterson 2010). It provides prevention and treatment for families and children with externalizing behavior problems (Forgatch and Patterson 2010). The aim of this parent-training intervention is to promote effective parenting skills to reduce and prevent the further escalation of child problem behavior. The central aims of PMTO are to target coercive transactional communication processes in the family and to teach and practice the parenting skills; positive involvement, effective discipline, problem solving, skill encouragement, and monitoring. Furthermore, in PMTO there is an emphasis on individual adaptation of session contents and progression, typically provided over 25 one-hour sessions.

In Norway, PMTO has been tested in two RCTs, both of which revealed more positive outcomes for PMTO than for usual treatment in the Norwegian services system (Kjøbli et al. 2013) Moreover, sustaining program fidelity is one of the acknowledged challenges in the process of scaling-up programs. Forgatch and DeGarmo (2011) investigated PMTO fidelity in terms of adherence to program factors across three generations of therapists (G1, G2, and G3), which correspond to the therapists in the present study. Their study showed a small drop in fidelity from G1 to G2, but the G3 therapists maintained the same high levels of fidelity as the G1 therapists. The participants in the studies that were reported by Ogden and Hagen (2008) and Forgatch and DeGarmo (2011) were included in the present study to compare changes in child problem behavior following PMTO across effectiveness and large-scale dissemination conditions.

Implementation of PMTO in Norway

Sociopolitically, the Norwegian implementation of PMTO was put forward in a social democratic welfare state that offers free public health care to all citizens. There are three separate service systems for youth with behavior problems: the child mental health service system (e.g., psychiatric or specialist services), the child welfare system, and the

school system, which includes educational and psychological counseling services. Candidates for PMTO training were recruited from all three service systems. Hereafter, when we refer to the child welfare system, we include educational and psychological counseling services in this category.

As part of the implementation plan that was introduced by the Ministry of Child and Family Affairs, representatives of all 19 county health directors in Norway were invited by the government to participate in the testing and the subsequent implementation of PMTO. All county municipalities accepted and decided to take part in the nationwide implementation project (Ogden et al. 2009). The implementation plan for PMTO was designed corresponding to what Fixsen and others have described as an active implementation approach (Fixsen et al. 2009; Fixsen et al. 2005) This framework underlines the importance of describing (1) the intervention (e.g., handbooks that describe treatment principles and procedures), (2) how the intervention is supported in practice (e.g., recruitment, leadership, training, supervision, fidelity assessment), and (3) who implements the program (individuals or teams of purveyors; Ogden and Fixsen 2014). Consequently, great effort was invested in the establishment of a comprehensive infrastructure to support the PMTO implementation (Ogden et al. 2005). Following a five-year project phase at the University of Oslo, a non-commercial, self-sustained national center for implementation and research was established on a more permanent basis: the Norwegian Center for Child Behavioral Development (NCCBD). The center is fully owned by the University of Oslo but is funded by several Norwegian Ministries, particularly the Ministry of Child and Family Affairs and the Ministry of Health. The aim of the center is to establish an implementation infrastructure for several evidence-based programs and to recruit candidates for PMTO training, which is relevant to this study. NCCBD employees further organized and supported the PMTO implementation.

Central to the implementation infrastructure was the establishment of a National Implementation Team (NIT), which was recruited from the first group of trainees in Norway and is often referred to as generation one (G1). G1 essentially had a background in specialist psychiatric services, and G1 was trained by PMTO founders Dr. Marion Forgatch and her colleagues at the Oregon Social Learning Center. Together with NCCBD employees, some of the therapists in G1 became members of NIT, training and supervising subsequent generations of PMTO therapists (e.g., G2 and G3). The NIT conducted numerous implementation support activities. PMTO-candidates had to undergo an 18-month training period to become a therapist. Regional groups of four candidates met one workday every second week throughout the 18-month period. Moreover, after becoming a PMTO therapist, onsite coaching and supervision were performed in regional groups with up to eight therapists, where therapists shared experiences and polished clinical skills (Ogden et al. 2005). PMTO therapists were obliged to attend 85 % of the supervising groups to attain or retain certification. There was regular monitoring of fidelity, and therapists had to provide between two and eight videotaped therapy sessions each year to maintain certification as PMTO therapists. The therapists' local agencies had to agree to provide resources, such as money, and time to engage in training and quality assurance activities. Together with the provision of technical support, the activities mentioned serve as examples of the central quality assurance implementation support tasks that were performed by the NIT. Importantly, by offering continuous training in PMTO, the NCCBD staff prevented the negative effects of turnover among therapists and local agency leaders. Thus, an important part of this study involves the service providers, i.e., the generations of therapists in the Norwegian dissemination. In this study, the first three generations of PMTO therapists represented the service providers in the transition of PMTO from regional specialist services to generalists in the municipal welfare system. Following the county health directors' consent to participate in PMTO-implementation, therapists were recruited through their local leaders and agencies throughout Norway. Motivated candidates signed up for PMTO-training voluntarily. Thus, all three generations of therapists who delivered cases in this study were likely to be highly motivated to practice PMTO. Today, there are six generations of PMTO therapists in Norway.

Regarding the challenges of large-scale dissemination and the conceptualized implementation factors, the NCCBD and the NIT team comprised the service infrastructure that supported the implementation process (e.g., recruitment, training, recertification, and supervision) from effectiveness to large-scale dissemination. Therefore, implementation factors were more or less a constant in our study.

Aims

In the present study, we aimed to investigate the potential scale-up penalties in the implementation of PMTO by focusing on child behavioral change across two phases of implementation, the effectiveness phase and the large-scale dissemination phase. In this evaluation of the dissemination of PMTO, we relate the primary outcome of child behavior to scale-up penalties to participants' benefiting less from PMTO that is delivered in the large-scale dissemination phase than in the initial effectiveness phase. In that vein, we define the scale-up penalty as the reduction in child behavioral change when children and families are treated in

the dissemination phase of implementation. When we speak of child behavioral change, we refer to a reduced amount of positive change regarding externalizing, internalizing, and social behavior problems. Similarly, when we speak of the attenuation of program effects, we refer to the decline of child behavioral change across phases of implementation (not to be confused with the reduction of long-term or follow-up effects in individuals). First, we ask: *Is there a scale-up penalty in the Norwegian large-scale dissemination of PMTO?*

Although our main objective was to study scale-up penalties, we additionally focused on how the composition of the service providers, or practitioners, and the target group were affected by the dissemination process. In our review of the challenges in large-scale dissemination, we have reported on how programs that are taken to scale often face increasing challenges regarding larger heterogeneity both in the target group and among the service providers. Therefore, we aimed to investigate the composition of the participant group and service providers in the large-scale dissemination of PMTO in Norway. Secondly, we ask: *How is the composition of the target group and the service providers affected by the scale-up process?*

Moreover, partly due to changes that were found in the composition of the two groups, we wanted to examine rival hypotheses in our results. We therefore included additional analyses in our results section, including analysis in which we matched the target groups and regrouped the service providers.

Method

Participating PMTO Therapists

The data that were used in the present study were collected in two interconnected studies, the effectiveness study and the dissemination study. The PMTO therapists were recruited from three generations of therapists who work at different levels of the Norwegian service system. In Norway, there are three separate service systems for children and youth with behavior problems: the child mental health service system (e.g., psychiatric or specialist services), the child welfare system, and the school system, which includes educational and psychological counseling services. Candidates for PMTO training were recruited from all three of the service systems. Hereafter, when we refer to the child welfare system, we include educational and psychological counseling services in this category. Parallel to the effectiveness study, the dissemination study was initiated to study the implementation process when implementing PMTO nationwide in Norway. The latter study was sponsored by the US National Institute of Drug Abuse. NCCBD and program developers from OSLC organized a meeting to recruit all three generations of PMTO therapists to deliver cases to the dissemination study, wherein open invitations to partake in the study were sent to all Norwegian PMTO therapists. Most of the therapists agreed to participate; however, not all of them delivered cases to the study, see Table 1. The effectiveness and the scale-up phases partly overlapped, and a relevant issue is how the two phases differed from one another. We aim to show the differences first by focusing on how the three generations of PMTO therapists differed and second by describing differences in how the three generations supplied cases in the two phases of implementation.

First, the differences between the three generations of practitioners are summed up in Table 1. Table 1 displays a shift in the therapists' background training from G1 to the subsequent cohorts, G2 and G3. In *category 1,* PMTO therapists had a minimum of six years of training in psychology, psychiatry or education in addition to extended relevant clinical practice. In *category 2*, therapists had a three-year college education primarily in child welfare, social work, teaching or nursing. In G1, 70 % of the candidates had category 1 levels of background training. Regarding G2 and G3 therapists, the percentages of category 1 level were markedly lower at 27 and 19 %, respectively. This change in educational background was an intended aspect of the implementation plan to transition PMTO from mental health specialist services to generalists in the child welfare services. Furthermore, Table 1 displays how G1 therapists were largely recruited from specialist services (71 %), and it also shows that G2 therapists were recruited evenly from specialists and welfare services. G3 therapists were almost exclusively recruited from generalist welfare services (94 %).

Second, the three generations of therapists supplied an unequal proportion of cases (children and families) to the EG and DG, see Table 1. Of the EG cases, 73 % were delivered by G1 therapists, whereas the remaining cases came from G2. The DG largely consisted of cases that were treated by G3 (33 %) and G2 (58 %), and only 9 % were supplied by G1. Furthermore, the PMTO therapists in the DG were scattered across all of the Norwegian health regions, and they were situated essentially in all of the service level organizations that were intended to deliver PMTO in Norway (9; see Table 1). The 263 cases in the DG were extracted from these 9 organizations and do not represent all of the cases that received PMTO during the data collection period. Of the 187 educated therapists in the data collection period, 134 (72 %) delivered cases to this study, see Table 1. In 2014, approximately 2500 families received PMTO in Norway, and a total of approximately 10,000 children and families had received PMTO through these services from its beginning to 2014. Moreover, the

Table 1 Descriptive statistics of PMTO therapists and phases of implementation

	Education level		Workplace		Therapists (total T.)[a]	Effect group	Dissemination group	FIMP[b]
	Category 1	Category 2	Psychiatric services	Child welfare				
Generation 1	70 % (18)	30 % (7)	80 % (20)	20 % (5)	25 (34)	73 %	9 %	6.94
Generation 2	27 % (15)	73 % (42)	53 % (30)	47 % (27)	57 (84)	27 %	58 %	6.34
Generation 3	19 % (10)	81 % (42)	8 % (4)	92 % (48)	52 (69)	0 %	33 %	6.94
Therapist ratio						1.8	2.2	
Organizations[c]						2	9	

Category 1 education level: a minimum of 6 years of higher education matching a degree as a clinical psychologist. Category 2 education level: a minimum of 3 years of higher education matching a degree in social work or teaching

[a] Total of number of therapists by each generation

[b] FIMP is a PMTO fidelity measure, numbers taken from Forgatch and DeGarmo (2011)

[c] Number of overarching service organizations where therapists worked (not to be confused with total number of institutions)

therapist ratio was low in both groups, 1.8 in the EG and 2.2 in the DG.

To summarize, the initial plan was to first roll out PMTO in the mental health specialist services and then to therapists in the primary welfare services. Thus, the DG contained therapists from multiple service institutions and across all service levels who were intended to deliver PMTO. Furthermore, the DG therapists had more diverse background training than the EG therapists. Therefore, differences between the DG and the EG, and thereby differences in the phases of implementation, were marked by disparities in the workplace and the background training of the three generations of therapists and by their differentiated delivery of cases to the EG and the DG. Regarding our second hypothesis that concerns the composition of the service providers and the conceptualized challenges in large-scale dissemination, the DG is clearly hallmarked by an increasing heterogeneity among the PMTO service providers.

Inclusion Criteria and Recruitment of Families

The participants in this combined study were 322 children and their parents, out of whom 263 families belonged to the DG and 59 belonged to the EG. The data collection period was from approximately 2001–2005 for the EG and from 2003 to 2005 for the DG. The children and families who were enrolled in both studies were recruited through the PMTO therapists' regular services. The EG children were mostly recruited in the county specialist services, and thus, they were mostly children who were referred from primary municipal welfare services. The children and families in the DG were essentially recruited in the municipal welfare services, see Table 1. Prior to the inclusion of families in the studies, a screening was performed based on clinical opinion in accordance with the regular procedures that were used in the agencies (Kjøbli and Ogden 2009; Ogden and Hagen 2008). In contrast to the more formal screening that was grounded in diagnostic criteria, clinical opinions were based on therapists' judgements after consulting with the parents of children with various externalizing behavior problems (e.g., conduct problems, disruptive behavior, antisocial behavior, and oppositional behavior). Thus, the participants who were included into the two studies were recruited from the pool of clients in the 134 PMTO therapists' regular practices, and the recruitment process matched the inclusion procedures that were routinely used in PMTO treatment in Norwegian services. Importantly, children were included in the studies before pre-assessment, and both pre- and post-assessment were administered to the families by a local therapist. However, there was one important difference in the recruitment process. In the EG, the participants had to accept the possibility that they could be randomly assigned to PMTO or to the usual treatment. Thus, the control group in the effectiveness study was not included in the present study. In the DG, all of the participants knew they would be assigned to PMTO. The eligible families were informed about the study, invited to participate, and accepted by signing a written informed consent.

Measures

The effectiveness and dissemination studies had identical measures, which allowed for direct comparisons of child behavioral change. The measures of child behavior had previously been translated and used in Norwegian studies, and both parents and teachers performed assessments.

The child behavior checklist (CBCL) and *Teacher Report Form (TRF)* are widely used instruments for assessing children's adjustment and behavior (Achenbach 1991). Both instruments have been standardized and validated for Norwegian studies (Nøvik 1999; Ogden and Hagen 2008). Both

externalizing and internalizing problem behavior scales were used in this study. The tests are comprised of 3-point Likert-scale items to which the respondents answered "0" (never/seldom true of the child), "1" (sometimes or somewhat true), or "2" (often or always true). A higher score indicates more problem behavior.

The social skills rating system (SSRS; Gresham and Elliott 1990) is a multi-rater instrument that assesses social skills in children. The parent and teacher versions were used, and both versions were previously found to be reliable and valid for Norwegian studies. The original 3-point Likert scale was modified to a 4-point version (Ogden 2003). The SSRS parent scale has 38 items, and the SSRS teacher scale has 30 items. A higher score indicates higher social competence.

Overall, the internal consistencies (Cronbach's alphas) for all of the child behavior instruments ranged from .86 to .96 and were all within an acceptable range.

Children's age and gender, parents' demographic background factors, and organizational levels were used as covariates in the analytic models. To measure family economic resources and to compare them with population statistics, an income-to-poverty ratio (*OECD poor*) was computed based on the OECD equivalent measure. Congruent with the OECD measure, a conservative poor cut-off was computed as 50 % of the median net income. *Parental education* was computed in 6 categories, (1) 7-year elementary, (2) junior high school, (3) high school vocational (<11 years), 4) high school general sciences (< 11 years), 5) college and some university courses, and 6) university degree or professional college. *Non-Western ethnicity* was computed as a dichotomized variable between non-Western immigrants (which includes Eastern Europeans, Asians, and people south of the equator) and other participants. Single parents were computed as a dichotomized variable. *Parental mental distress* (anxiety and depression) was measured with the Symptom Check List 5 (SCL-5; e.g., "feeling fearful"). The SCL-5 is a short form of the SCL-25 that measures anxiety and depression and that had previously been validated and normed in a Norwegian study (Tambs and Moum 1993). The Cronbach's alpha was .88 for the SCL-5. In addition, *parent age* was used as a covariate. Organizational level was measured with a dichotomous variable where municipal child welfare was coded 0, and county specialist services were coded 1. Moreover, parent age was also included as a covariate in the main analyses.

Analytic Procedures

Missing Data and Outliers

Missing data were inspected, and a missing value analysis was performed using SPSS version 22. The outcome variables were investigated for missing completely at random test (MCAR). Tests showed that the outcome data missing were MCAR, and a single imputation method based on an expectation maximization procedure (EM) was performed. EM is an imputation method that is based on an iterative procedure to fit the most unbiased values (Tabachnick and Fidell 2001). Imputation was performed on *missing items only*, thereby leaving out cases in which the entire instruments were missing. Therefore, children with missing values on all post-outcome variables were not a part of the analysis. Additionally, we also performed a multiple imputation (MI) procedure on the dataset and ran outcome analyses in regression models to test the robustness of our results without missing values, see results section. Unfortunately, SPSS does not support multiple imputation and multiple analysis of covariance (MANCOVA), which were used in our main analysis. Therefore, we kept to the original analytic procedure, see the next section.

Outliers were identified and inspected to ensure that these values were within the range of scores that were defined by the minimum and maximum values of the scales. The 5 % trimmed mean was compared to the original mean. In all of the cases, the differences were marginal, which indicates that the outliers had little effect on the original means. Therefore, the outliers were not modified.

All of the scales were examined in terms of normal distribution and were found to be within an acceptable range of skewness and kurtosis (+/−2; (Frankfort-Nachmias and Nachmias 1996). Consequently, no transformations of variables were performed.

Analyses of Children's Behavioral Change

Children's behavioral change and group differences were investigated in a pre-post design using a within-subject factorial MANCOVA. Two MANCOVA models that contained parent- and teacher-reported outcomes were run using composite variables that were both empirically and conceptually related. All of the variables within each composite were significantly correlated, ranging from .197 to 420 and .312 to .504, for parent-reported and teacher-reported outcomes, respectively. The parent-reported composite outcomes consisted of the CBCL externalizing and internalizing problem scales and the SSRS parent scale. The teacher-reported composite variable contained the TRF externalizing and internalizing scales and the SSRS teacher scale. MANCOVA models were run with composite measures of the main outcome to reduce the probability of type 1 errors. However, to further explore group differences, significant post hoc analyses (simple contrasts) are displayed in the text. MANCOVA models

were run using the SPSS multivariate general linear modeling procedure. Due to unequal sample sizes, type 1 sums of squares were used in the MANCOVA analyses (Tabachnick and Fidell 2001). Furthermore, due to possible problems of bias in an unbalanced design, a nonequivalent group analysis was performed in MANCOVA models. Separate pre-score measurement errors and Cronbach's alphas were adjusted in both the EG and in the large-scale DG by computing new adjusted pre-scores (Trochim and Donnelly 2007). The results of the nonequivalent group analysis displayed similar results as in the original MANCOVA results (table not shown). Therefore, non-adjusted MANCOVA models are displayed in the results section. We also considered running nested models. Several authors have indicated that one should consider multilevel models for design effects >2.0 (see, Peugh 2010, pp. 90–91). We calculated intra correlation coefficients and then design effects for families clustered within therapists. Our design effects ranged from 1.02 and 1.2. Therefore, we did not run nested multilevel models.

Covariates were entered into the analysis separately and were removed if they were non-significant and/or did not influence the error variance that was accounted for by the model (SS_E). (P-score child behavior outcome variables were included in all of the models. The background factors that concerned family and parental demographics (e.g., total family income, parental education, marital status, and parent age), parental mental distress (SCL-5), organizational level, and child characteristics (e.g., age and sex), were tested in the models. However, all of the variables were non-significant and were thus removed from the final models. To test for homogeneity in the regression slopes, scatterplots and simple slopes were inspected, and statistical interaction variables were computed for all covariates and run separately in the GLM models. None of the interaction variables were significant, which indicates that the assumption of homogeneity in the regression slopes was not violated. Partial eta squared was used as an effect size measure. This variance-based effects size measure shows a percentage of variance explained that is non-related to covariates in the model (Field 2013).

Results

Attrition

The pre-assessment included 322 families, and 238 (74 %) completed outcome instruments at post-assessment. As mentioned, dropout from treatment is one of the acknowledged challenges in large-scale dissemination (Welsh et al. 2010). As it turned out, the dropout rate from the study was unevenly distributed across the phases of implementation, DG 32.7 % (89) and EG 6.8 % (9). There were likely numerous reasons for drop from the study groups. Questionnaires were mailed to families who did not show up for assessment. Furthermore, some families chose not to answer or answered only parts of the assessment battery. Some of the families that showed up for assessment did not have the time to fill out all of the measures, nor did they mail them to the researchers afterwards. Additionally, we do not know whether the dropouts from the study also dropped out of treatment. Attrition was dummy coded to test for potential differences between the families who completed the study and families who were lost before the post-assessment. The results revealed that there were no significant differences in the attrition rates due to pre-score child outcome variables, but regarding background covariates, a higher parent age was significantly associated with drop-out before post-assessment $t(221) = -2.57$, $p < .05$. Moreover, there was also significant attrition that was related to organizational level $t(329) = -2.09$, $p < .05$, which indicates that there was a higher likelihood of drop-out for children who were treated in the municipal child welfare services compared to the county specialist services. Furthermore, we tested whether there was statistical interaction between study condition, child behavior, and covariates, regressed on whether data were missing post treatment. Analyses revealed that there was no significant attrition related to differences in study conditions (DG and EG).

The Heterogeneity of Service Providers

Regarding our second hypotheses, concerning the composition of the service providers and the conceptualized challenges in large-scale dissemination, the results in Table 1 show that DG is clearly hallmarked by an increasing heterogeneity among the PMTO service providers. Increasing diversity according to work place, background training, and the number of service organizations were PMTO was given in the DG, back up this notion.

Participant Characteristics and Baseline Differences

In general, the participating families across the two studies represented a midrange Norwegian income level, with an annual gross income of 415.000 NOK (Statistics Norway 2014). The proportion of single parents (divorced, separated or never married) in our study was markedly higher than that of the Norwegian population: 37.5 % compared to 20.3 %, respectively. The participants had a slightly higher education level than the Norwegian population: 29.9 % of the parents reported having a college or higher university degree, and 18 % reported having completed high school or elementary school (population numbers, 24.4 % college/

higher degree, 44 % elementary or high school (Statistics Norway 2014). In terms of ethnicity, 94 % of parents reported to be of Norwegian origin compared to 93 % in the Norwegian population (Statistics Norway 2014).

Baseline differences between the DG and the EG are summarized in Table 2. An analysis of variance (ANOVA) was used for continuous variables, and Chi square tests were used for dichotomous variables. According to the parents, the children in the DG had significantly lower levels of externalizing problem behavior (M = 23.33) than the children in the EG ($M = 26.05$) $F(1313) = 3.97$, $p = .047$. Moreover, the children in the DG scored marginally higher on parent-reported social skills than the children in the EG ($M = 86.30$) $F(1/305) = 3.53$, $p = .061$. The baseline differences regarding teacher-reported data displayed that the children from the DG had less externalizing problem behavior than the children in the EG ($M = 25.41$) $F(1/277) = 4.93$, $p = .027$. Teachers also reported children's social skills scores to be significantly higher in the DG ($M = 70.14$) than in the EG ($M = 65.82$) $F(1/270) = 7.50$, $p = .007$. Concerning parent characteristics, there were two significant baseline differences between the groups (see Table 2). Parents in the DG ($M = 38.0$ years) were slightly older than EG parents ($M = 35.9$ years) $F(1/221) = 3.54$, $p = .061$, and the former group of parents reported significantly lower levels of mental distress $F(1/288) = 5.28$, $p = .022$.

As to our second hypotheses, regarding the composition of the target group and the conceptualized heterogeneity in the target population, we operationalized it as a function of child behavior at the baseline means and standard deviation (SD) in the outcome measures. As shown in Table 2, the DG displayed a lower problem level than the EG on four out of six child behavior outcomes. However, regarding differences between the groups in terms of SD, the numbers indicated that the variation around the baseline mean outcome scores was relatively equally distributed between the DG and EG (see Table 2). Nevertheless, based on the DG's lower problem levels in four out of six outcomes and thus with potentially less problem behavior to treat, we conclude that there was an increasing heterogeneity among the target population displayed in the DG.

Child Behavioral Outcomes

To investigate our first question, i.e., whether there was a scale-up penalty, two MANCOVA models were run for parent- and teacher-reported outcomes to investigate the

Table 2 Means, standard deviations, Chi square, and significance tests (ANOVA, F-tests & Pearson's r) of group differences (effect group & dissemination group) at baseline (pre-score)

Variables	Dissemination group (DG) M (SD)	Effect group (EG) M (SD)	*F*	*p*	Contrasts
Parent-reported outcome					
CBCL ext	23.33 (9.21)	26.05 (10.43)	3.97	.047*	DG < EG
CBCL int	13.10 (8.06)	13.59 (9.07)	.167	.683	
SSRS parent[a]	89.47 (11.66)	86.30 (11.18)	3.527	.061†	DG > EG
Teacher-reported outcome					
TRF ext	20.28 (15.35)	25.41 (14.09)	4.93	.027*	DG < EG
TRF int	8.88 (6.73)	10.46 (7.96)	2.20	.139	
SSRS teacher[a]	70.14 (10.53)	65.82 (9.76)	7.50	.007**	DG > EG
Covariates					
Salary	412[b] (220[b])	403[b] (189[b])	.086	.769	
Parent education	3.72 (1.21)	3.53 (1.23)	1.14	.287	
Parent age	38.0 (6.5)	35.9 (5.2)	3.54	.061†	DG > EG
Parent mental distress	1.77 (.83)	2.11 (.88)	5.28	.022*	DG < EG
Child age	8.6 (2.19)	8.9 (1.92)	1.018	.314	
Dichotomized covariates	Percent (%)	Percent (%)	χ^2 (p)		
Single parents	33.7 %		1 versus 2 ns. (.523)		
Child sex	71 % (boys)	81 % (boys)	1 versus 2 ns. (.112)		
N	263	59			

CBCL Child behavior check list, *ext* externalizing behavior problems, *internalizing behavior problems*, *SSRS* social skills rating scale, *TRF* Teacher Report Form

[a] A higher score indicates more social skills

[b] Means salary divided on 1000

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .010$

differences in children's behavioral changes in the EG and the DG. Table 3 presents the means and standard deviations of the pre-treatment and post-treatment scores, and an omnibus F-test for the composite parent- and teacher-reported scale-up penalty. The F-test indicates group differences between the DG and EG, and the partial eta squared displays effect size differences between the groups.

As displayed in Table 3, no significant scale-up penalties were detected in either of the composite outcome measures. Nevertheless, regarding parent-reported outcomes, children in the DG displayed 2.9 % (n_p^2 .029) more behavioral change than children in the EG. This behavioral change difference was not statistically significant ($p = .125$), but the significance level was in a range that indicated possible statistical significance in post hoc tests. The post hoc tests revealed that there was a significant difference between the DG and the EG regarding SSRS, $t(201) = -1.97$, $p = 0.50$, DG > EG, meaning that DG children displayed more positive change in social skills after PMTO treatment. The teacher-reported outcome revealed no significant differences between the DG and the EG.

In addition, we wanted to examine alternative explanations to our results by addressing heterogeneity issues in the large-scale dissemination study. First, we investigated the issue of participant heterogeneity by matching the participants in the EG and the DG on the CBCL externalizing problem behavior scale. We excluded children in both EG and the DG who scored below the 90 percentile, a clinical range (DG N = 197, EG N = 50), to make the target groups more similar according to problem behavior. Together with externalizing behavior, matching the groups resulted in parent reported social skills baseline differences that were also non-significant. These matched group results replicated the results from our original MANCOVA models. The DG group displayed slightly more positive behavioral change than the EG, but this effect size difference was not in a statistically significant continuum, see Appendix Table 4. Furthermore, we addressed the heterogeneity among the service providers by analyzing child behavioral outcomes in MANCOVA models for separate generations of PMTO therapists, G1, G2, and G3 (see Appendix Table 6). With regard to both parent reports and teacher reports, these analyses revealed a similar pattern as that which was displayed in Table 3 between the phases of implementation. The results indicated no significant differences between the generations of therapists. Although small and not significant, both of these analyses favored G2 and G3 over G1 regarding child behavioral change, see Appendix Table 6. Furthermore, we wanted to test whether attrition and missing data biased our results. Thus, we created a MI dataset, where missing data were handled by creating five different datasets based on the EM algorithm, and where the results of these five imputed datasets were pooled in the outcome analyses. The results of these analyses revealed results that were similar to the original MANCOVA analyses that contained missing cases. For example, with regard to the parent reported outcomes, the significance levels were all non- significant, ranging between p = .064 for internalizing behavior and p = .289 for externalizing behavior. The effect sizes (R^2) were in the range <1 % that favored the DG over the EG (not shown).

Table 3 Means, standard deviations, effect sizes, and analyses of covariance (MANCOVA) on child behavior by treatment group

Variable	Dissemination group (DG)		Effect group (EG)		Scale-up penalty		Contrasts	Effect size
	Pre-treatment M (SD)	Post-treatment M (SD)	Pre-treatment M (SD)	Post-treatment M (SD)	*F*	*p*		n_p^2
Parent reports					1.94	.125	DG > EG	.029
CBCL EXT	23.33 (9.21)	16.27 (8.72)	26.05 (10.43)	18.92 (11.86)				
CBCL INT	13.10 (8.06)	9.57 (7.45)	13.59 (9.07)	11.80 (9.71)				
SSRS[a]	89.47 (11.66)	95.01 (12.97)	86.30 (11.18)	89.67 (10.98)				
Teacher reports					.513	.674	DG < EG	.009
TRF EXT	20.28 (15.35)	19.02 (15.60)	25.41 (14.09)	18.80 (14.36)				
TRF INT	8.88 (6.73)	8.44 (7.03)	10.46 (7.96)	8.93 (8.06)				
SSRS[a]	70.14 (10.53)	70.47 (11.14)	65.82 (9.76)	68.88 (9.21)				

CBCL child behavior check list, *EXT* externalizing behavior problems, *INT* internalizing behavior problems, *SSRS* social skills rating scale, *TRF* Teacher Report Form

Parent reports DG N = 149, and EG N = 52. *Teacher reports* DG N = 133, and EG N = 48

[a] A higher score indicates more social skills. Children's behavior pre-treatment scores were used as covariates in all models. (All were significant at p < .001)

Discussion

The main purpose and first hypothesis in our study was to examine whether there was a scale-up penalty in PMTO implementation by comparing child behavior outcomes between the effectiveness phase and the large-scale dissemination phase. Contrary to previously reported scale-up penalties, no scale-up penalties were found in the Norwegian large-scale dissemination of PMTO. None of the two composite outcomes, representing home and school environments, displayed significant results. This is an indication that there were no differences regarding child behavioral change between the EG and the DG. Despite indications of a larger heterogeneity among both the service providers and the target population, the program was at least as effective in the large-scale dissemination phase as in the effectiveness phase, as measured by the amount of child behavioral change. Therefore, we suggest a scale-up penalty of 0 % in the Norwegian large-scale dissemination of PMTO.

In the second hypothesis in our study, we addressed whether scaling up affected the composition of the service providers and the target group. In that vein, we conceptualized three categories of challenges in sustaining program effects in large-scale dissemination: (1) implementation factors, (2) the heterogeneity of service providers, and (3) heterogeneity in target populations. Coupled by the fact that inclusion criteria were similar in both phases of implementation, the larger heterogeneity that we found in the DG target population might be caused by the fact that the Norwegian welfare service agencies traditionally target children with more differential risk levels compared to the specialist services. However, we cannot rule out that other and more informal inclusion criteria were at play in different parts of the service system and thus contributed to the heterogeneity of target populations. The larger heterogeneity among service providers reflects the transition of PMTO first to therapists in the psychiatric specialist service system and second to generalists in the relevant child welfare services. The implementation factors were held "constant" in our study because it was essentially the same organization (NCCBD) and the same purveyor team (NIT) that implemented PMTO from the effectiveness to the large-scale dissemination phases. Therefore, it is plausible to relate the absence of a scale-up penalty to Norwegian implementation factors using an active implementation approach (Fixsen et al. 2013) and the establishment of a sustainable implementation infrastructure. This stable infrastructure could not have been established without long-term governmental funding. Moreover, the active implementation approach and the absence of a scale-up penalty should be considered within the Norwegian context, along with the fact that child welfare and specialists services in Norway are essentially public and funded by the state. We may speculate as to whether an active implementation approach in which resources are needed for recertification and other fidelity-maintaining activities might be more feasible in a public service system than in private services. Another reasonable explanation for the absence of a scale-up penalty may be related to a program's maturation effects in the implementation organization that supports PMTO, i.e., the NCCBD and NIT. The program maturation effects have been defined as improvements in treatment outcomes due to increased experience and competence over time among therapists and in the implementation teams (Leschied and Cunningham, 2002; Ogden et al. 2007). Maturation effects could have outperformed the potential negative effects from the challenges in going to scale, here in the form of increasing heterogeneity in the target population and service providers.

To test rival (heterogeneity) hypotheses, additional analyses were conducted. First, a test was performed to see whether the lack of scale-up penalty was a result of the program maturation effects among the G1 therapists who delivered the cases to the DG, but in separate analyses of the generations of therapists, no maturation effects among G1 therapists were supported by our data; G1 did not outperform G2 or G3 in terms of child behavioral change. Moreover, in the DG, the average therapist ratio was 2.2. Therefore, it was most likely not program maturation among the G1 in the DG that biased our results and the absence to detect a scale-up penalty. In other words, in support of our implementation factor explanation above, a possible maturation effect could be related to the service infrastructure that supports PMTO in Norway. Another competing hypothesis was that children with less pervasive and serious problem behaviors benefitted more from PMTO therapy. This issue was addressed by matching the participants in terms of problem behavior in both of the PMTO groups. However, these analyses did not support the notion that increased heterogeneity in the target population could explain the absence of a scale-up penalty in our data.

Finally, our findings were supported by previous studies that demonstrated the sustainability of fidelity ratings across generations of therapists (Forgatch and DeGarmo 2011) and over time (Hukkelberg and Ogden 2013). The results indicate that the close monitoring of PMTO fidelity by NCCBD and NIT employees affected both program fidelity and child behavioral outcomes. Moreover, recent data from the NCCBD replicates the high fidelity levels that were displayed by Forgatch and DeGarmo (Forgatch and DeGarmo 2011) in subsequent generations of PMTO therapists, from generation 3 to generation 6 (Ogden and Fixsen 2014). The high fidelity scores in subsequent generations of therapists support our explanation that an active centralized implementation strategy may have affected program sustainability in terms of both behavioral outcomes and fidelity.

Limitations

Although this study has the advantage of using a multi-informant approach that was measured before and after PMTO treatment in two phases of implementation, it also has some limitations. The study did not allow for the randomization of participants to different phases of implementation, so we cannot claim any causal relationships between the implementation phase and stable child behavioral outcomes. *Moreover, we have to bear in mind that the additional analyses that were performed did not eliminate heterogeneity issues in our data. Clearly, the children and families in the two phases of implementation were different. Thus, all of the measured child and parental characteristics that differed at baseline were addressed in analyses and entered as covariates. However, we cannot rule out that other unmeasured parental and child confounders might have caused the effects in our results.* Relating this issue, we related the lack of scale-up penalties to implementation factors in the discussion. We do not know whether the lack of detection of a scale-up penalty might be related to other unmeasured implementation factors. Although there are many similarities in design, comparing two different studies might have resulted in unknown dissimilarities between the studies that could have biased our results. Moreover, there was a difference in the recruitment conditions in the EG and the DG: the participants in the EG had to accept the possibility of being randomized to usual treatment, whereas all of the participants in the DG knew that they would receive PMTO. We do not know, however, if this influenced the recruitment to the studies and hence the generalizability of the results. Moreover, an explanation of our results may be related directly to features in the PMTO intervention. For example, the PMTO intervention may be a very teachable and trainable program that is especially suited to large-scale dissemination. However, we do not know if these findings can be replicated and extended beyond PMTO to less curriculum-based and more complex clinical interventions. Furthermore, attrition cannot be dismissed as a potential influence in our results. Although attrition analysis indicated no systematic influence on the baseline outcome variables, we cannot completely rule out other hypotheses, e.g., that client satisfaction affected dropout in our study. Even so, dropout is a potential penalty in large-scale dissemination. Statistical power, a type II error, is another limitation to regarding low N in the EG. This may have resulted in false negative result; i.e., we statistically failed to detect an existing scale-up penalty. However, overall, our results indicate that the DG profited more than the EG; therefore, a scale-up bonus is more adjacent in our results than a scale-up penalty.

Conclusions

Despite the limitations, the outcomes of this study rather consistently demonstrate how the emphasis on implementation factors could have an impact on program effects in the large-scale dissemination of model programs. Moreover, this study has showed that the PMTO intervention is well suited for dissemination across service systems when it is delivered under different conditions. More research is needed to confirm whether a centralized, comprehensive and long-term active approach to implementation may prevent the dilution of program effects in the face of increased heterogeneity in service providers and client populations. From an applied point of view, the findings underline the importance of having a central organization that can establish a comprehensive implementation infrastructure that may sustain a high program implementation quality and a high level of treatment adherence over time across an increasing number of therapists and clients. Such an infrastructure may maintain program effects on child behavior by supporting core implementation components at the competency level (e.g., recruitment, training, supervision and practice/fidelity assessment) and at the organizational level (e.g., data decision support data systems, technical support and evaluation). Long-term funding is an important prerequisite for such organizations, but their success is also dependent on having an infrastructure for scaling up empirically supported interventions and the ability to strike a good balance between program integrity and local adaptations, as well as to monitor and evaluate clinical outcomes.

Acknowledgments Thanks to: Anett Apeland, Asgeir R. Olseth, Bjørn Arild Kristiansen, Elisabeth Askeland, John Kjøbli, Kristian Heggebø, Kristin Nordahl, Mari-Anne Sørlie, Marion Forgatch, Terje Christiansen, Torkild Hovde Lyngstad, all you in the research department at the Norwegian Center for Child Behavior Development, and the National Implementation Team.

Authors' Contributions TT designed the study, wrote parts of the manuscript, conducted the statistical analyses, and interpreted the results. TO designed the study, organized the data collection, wrote parts of the manuscript, and interpreted the results.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no competing interests.

Appendix

See Tables 4, 5 and 6

Table 4 Means, standard deviations, effect sizes, and analyses of covariance (MANCOVA) on child behavior by treatment groups matched on child behavior

Variable	Dissemination group (DG)		Effect group (EG)		Scale-up penalty		Contrasts	Effect size
	Pre-treatment M (SD)	Post-treatment M (SD)	Pre-treatment M (SD)	Post-treatment M (SD)	*F*	*p*		n_p^2
Parent reports					1.70	.169	DG > EG	.032
CBCL EXT	26.79 (7.42)	18.02 (8.68)	28.62 (9.09)	20.28 (12.02)				
CBCL INT	14.59 (8.07)	10.23 (7.74)	15.10 (8.07)	12.55 (9.86)				
SSRS[a]	87.87 (11.34)	94.11 (13.50)	86.35 (11.67)	89.63 (11.35)				
Teacher reports				1.04	.378		DG < EG	.022
TRF EXT	21.49 (15.32)	19.52 (15.51)	26.33 (14.36)	19.36 (14.62)				
TRF INT	9.07 (6.55)	8.83 (6.69)	10.87 (7.77)	9.26 (8.54)				
SSRS[a]	69.63 (10.42)	69.65 (10.67)	65.66 (9.96)	70.03 (8.89)				

CBCL child behavior check list, *EXT* externalizing behavior problems, *INT* internalizing behavior problems, *SSRS* social skills rating scale, *TRF* Teacher Report Form

Parent reports DG N = 115, and EG N = 44. *Teacher reports* DG N = 103, and EG N = 41

[a] A higher score indicates more social skills. Children's behavior pre-treatment scores were used as covariates in all models. (All were significant at $p < .001$)

Table 5 Means, standard deviations, effect sizes, and multivariate analyses of covariance (MANCOVA) on child behavior by treatment groups using nonequivalent group analysis (alpha adjusted pre-scores)

Variable	Dissemination group (DG)		Effect group (EG)		Scale-up penalty		Contrasts	Effect size
	Pre-treatment M (SD)	Post-treatment M (SD)	Pre-treatment M (SD)	Post-treatment M (SD)	*F*	*p*		n_p^2
Parent reports					1.40	.244	DG > EG	.024
CBCL EXT	23.33 (7.98)	16.21 (9.09)	26.05 (9.51)	19.32(12.23)				
CBCL INT	13.10 (6.81)	9.37 (7.56)	13.59 (8.08)	11.41 (9.56)				
SSRS[a]	89.47 (10.16)	94.64 (13.10)	86.30 (9.33)	88.81 (11.08)				
Teacher reports					.768	.513	DG < EG	.013
TRF EXT	20.28 (14.75)	19.01 (15.47)	25.41 (13.35)	18.91 (14.44)				
TRF INT	8.88 (5.69)	8.65 (7.04)	10.46 (6.76)	8.81 (8.22)				
SSRS[a]	70.14 (9.07)	70.36 (11.20)	65.82 (8.29)	69.37 (9.10)				

CBCL child behavior check list, *EXT* externalizing behavior problems, *INT* internalizing behavior problems, *SSRS* social skills rating scale, *TRF* Teacher Report Form

Parent reports DG N = 132, and EG N = 49. *Teacher reports* DG N = 133, and EG N = 48

[a] A higher score indicates more social skills. Children's behavior pre-treatment scores were used as covariates in all models. (All were significant at $p < .001$)

Table 6 Means, standard deviations, effect sizes, and multivariate analyses of covariance (MANCOVA) on child behavior by generation of therapists

Groups	G1		G2		G3		Omnibus test	
Outcomes	Pre-treatment M (SD)	Post-treatment M (SD)	Pre-treatment M (SD)	Post-treatment M (SD)	Intake M (SD)	Post-treatment M (SD)	*F*	*p*
Parent reports							.846	.535
CBCL EXT	26.91 (9.21)	18.83 (12.37)	26.05(10.43)	16.84 (9.14)	20.55 (8.67)	14.72 (7.35)		
CBCL INT	14.23 (9.62)	11.25 (7.35)	13.55 (7.63)	10.13 (7.42)	11.75 (8.15)	8.21 (5.97)		
SSRS	87.00 (11.34)	90.66 (12.71)	89.37 (11.69)	95.26 (13.39)	89.36 (11.70)	92.77 (11.39)		
Teacher reports							.391	.885
TRF EXT	25.82 (14.36)	21.52 13.57)	20.81 (15.54)	18.56 (15.94)	18.53 (14.64)	17.13 (15.02)		
TRF INT	8.58 (6.71)	8.20 (7.87)	9.89 (7.15)	9.04 (6.93)	8.22 (6.81)	8.46 (7.81)		
SSRS TEA	66.78 (10.40)	69.17 (.73)	70.15 (9.90)	70.90 (10.74)	69.60 (11.47)	69.76 (11.14)		

CBCL child behavior check list, *EXT* externalizing behavior problems, *INT* internalizing behavior problems, *SSRS* social skills rating scale, *TRF* Teacher Report Form

Parent reports G1 N = 49, G2 N = 107, and G3 N = 45. *Teacher reports* G1 N = 46, G2 N = 93, and G3 N = 42

[a] A higher score indicates more social skills. Children's behavior pre-treatment scores were used as covariates in all models. (All were significant at $p < .001$)

References

Achenbach, T. (1991). *Child behavior checklist/4-18*. Burlington: University of Vermont.

Aos, S., Phipps, P., Barnoski, R., & Lieb, R. (2001). *The comparative costs and benefits of programs to reduce crime. Version 4.0*. Washington: Olympia.

Bonta, J., & Andrews, D. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. *Rehabilitation, 6*, 1–22.

Dishion, T. J., & Patterson, G. R. (2006). The development and ecology of antisocial behavior in children and adolescents. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol 3: risk, disorder, and adaptation* (2nd ed) (Vol 3, pp. 503–541). Hoboken, NJ, US: Wiley, xvi, 944 pp

Dodge, K. A. (2001). The science of youth violence prevention: Progressing from developmental epidemiology to efficacy to effectiveness to public policy. *American Journal of Preventive Medicine, 20*(1), 63–70.

Donohue, J., & Siegleman, P. (1998). Allocating resources among prisons and social programs in the battle against crime. *The Journal of Legal Studies, 27*(1), 1–43.

Elliott, D. S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science, 5*(1), 47–53.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Los Angeles: Sage.

Fixsen, D., Blasé, K., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children, 79*(2), 213–230.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*(5), 531–540.

Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. M. (2005). Implementation research: A synthesis of the literature.

Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining fidelity following the nationwide PMTO™ implementation in Norway. *Prevention Science, 12*(3), 235–246.

Forgatch, M. S., & Patterson, G. R. (2010). Parent management training—Oregon Model: An intervention for antisocial behavior in children and adolescents. *Evidence-Based Psychotherapies for Children and Adolescents, 2*, 159–178.

Forgatch, M. S., Patterson, G. R., & Gewirtz, A. H. (2013). Looking forward the promise of widespread implementation of parent training programs. *Perspectives on Psychological Science, 8*(6), 682–694.

Frankfort-Nachmias, C., & Nachmias, D. (1996). Research Methods in the Social Sciences. 1996. *Arnold, London*.

Greenwood, P. W. (1998). *Diverting children from a life of crime: Measuring costs and benefits*. Santa Monica: Rand Corporation.

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system (SSRS)*. Circle Pines: American Guidance Service.

Hukkelberg, S. S., & Ogden, T. (2013). Working alliance and treatment fidelity as predictors of externalizing problem behaviors in parent management training. *Journal of Consulting and Clinical Psychology, 81*(6), 1010.

Kellam, S. G., & Langevin, D. J. (2003). A framework for understanding "evidence" in prevention research and programs. *Prevention Science, 4*(3), 137–153.

Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: reducing child conduct problems in real-world settings. *Behaviour Research and Therapy, 51*(3), 113–121.

Kjøbli, J., & Ogden, T. (2009). Gender differences in intake characteristics and behavior change among children in families receiving parent management training. *Children and Youth Services Review, 31*(8), 823–830.

Leschied, A., & Cunningham, A. (2002). Seeking effective interventions for young offenders: interim results of a four year randomized study of multisystemic therapy in Ontario, Canada. London, ON: Centre for Children and Families in the Justice System. *London.* Retrieved 8 Oct, 2013.

Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders, 4*(2), 124–147.

I.O.M, N. R. C. (2014). Institue of Medicine and National Research Council. Strategies for scaling effective family-focused preventive interventions to promote children's cognitive, affective, and behavioral health: workshop summary Washington DC: The National Academic Press.

McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: a review of current efforts. *American Psychologist, 65*(2), 73.

Mihalic, S. F., & Irwin, K. (2003). Blueprints for Violence prevention from research to real-world settings—factors influencing the successful replication of model programs. *Youth Violence and Juvenile Justice, 1*(4), 307–329.

Nøvik, T. (1999). Validity of the child behaviour checklist in a Norwegian sample. *European Child and Adolescent Psychiatry, 8*(4), 247–254.

Ogden, T. (2003). The validity of teacher ratings of adolescents' social skills. *Scandinavian Journal of Educational Research, 47*(1), 63–76.

Ogden, T., & Fixsen, D. L. (2014). Implementation Science. *Zeitschrift für Psychologie, 222*(1), 4–11.

Ogden, T., Forgatch, M. S., Askeland, E., Patterson, G. R., & Bullock, B. M. (2005). Implementation of parent management training at the national level: The case of Norway. *Journal of Social Work Practice, 19*(3), 317–329.

Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of Parent Management Training in Norway: a randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology, 76*(4), 607.

Ogden, T., Hagen, K., & Andersen, O. (2007). Sustainability of the effectiveness of a programme of multisystemic treatment (MST) across participant groups in the second year of operation. *Journal of Children's Services, 2*(3), 4–14.

Ogden, T., Hagen, K. A., Askeland, E., & Christensen, B. (2009). Implementing and evaluating evidence-based treatments of conduct problems in children and youth in Norway. *Research on Social Work Practice, 19*(5), 582–591.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48*(1), 85–112.

Statistics Norway (2014). Retrieved from http://ssb.no/en/arbeid-og-lonn

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Pearson.

Tambs, K., & Moum, T. (1993). How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatrica Scandinavica, 87*(5), 364-367.

Trochim, W., & Donnelly, J. (2007). *The research methods knowledge base*. Thomson: Custom Publishing.

Welsh, B. C., Sullivan, C. J., & Olds, D. L. (2010). When early crime prevention goes to scale: a new look at the evidence. *Prevention Science, 11*(2), 115–125. doi:10.1007/s11121-009-0159-4.