JYU DISSERTATIONS 571

Elina Aaltio

Effectiveness of the Finnish Systemic Practice Model for Children's Social Care

A Realist Evaluation

UNIVERSITY OF JYVÄSKYLÄ
FACULTY OF HUMANITIES AND
SOCIAL SCIENCES

JYU DISSERTATIONS 571

Elina Aaltio

Effectiveness of the Finnish Systemic Practice Model for Children's Social Care

A Realist Evaluation

Esitetään Jyväskylän yliopiston humanistis-yhteiskuntatieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Musica-rakennuksen salissa M103
joulukuun 3. päivänä 2022 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities and Social Sciences of the University of Jyväskylä,
in building Musica, auditorium M103, on December 3, 2022 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2022

Editors
Olli-Pekka Moisio
Department of Social Sciences and Philosophy, University of Jyväskylä
Ville Korkiakangas
Open Science Centre, University of Jyväskylä

Copyright © 2022, by University of Jyväskylä

ISBN 978-951-39-9224-8 (PDF)
URN:ISBN:978-951-39-9224-8
ISSN 2489-9003

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-9224-8

ABSTRACT

Aaltio, Elina
Effectiveness of the Finnish Systemic Practice Model for children's social care. A realist evaluation.
Jyväskylä: University of Jyväskylä, 2022, 131 p.
(JYU Dissertations
ISSN 2489-9003; 571)
ISBN 978-951-39-9224-8 (PDF)

The aim of this dissertation was to study the effectiveness of the *Systemic Practice Model* (SPM) for children's social care within a realist evaluation framework. The SPM is a Finnish adaptation of Reclaiming Social Work (RSW), a practice model for child and family statutory social work developed in the London borough of Hackney. Previous research has shown that RSW improved the quality of practice and work-related well-being in Hackney. While implementation in new sites in the UK has met various challenges, practitioners have primarily reported positive experiences of RSW. However, evidence of its effectiveness remains limited.

This thesis is based on four sub-studies and three distinct study designs and data. Two of the sub-studies, which compared family-level outcomes between a SPM group (n=56 cases) and service-as-usual group (n=56 cases), using repeated measures data, were quasi-experimental. The third sub-study, which assessed implementation fidelity and the factors influencing this, applied a concurrent transformative mixed-methods design, i.e., simultaneous quantitative and qualitative data collection, and mutual interpretation of the findings. The fourth sub-study was a research-led process aimed at formulating a programme theory for the SPM based on discussions generated in workshops and expert interviews.

The key finding was that during its initial implementation the SPM did not outperform service as usual in relation either to the selected outcome measures or service-user feedback, although the volume of meetings in the SPM group was higher. The implementation fidelity analysis revealed several barriers that inhibited practitioners from adopting the SPM and applying systemic practice as intended: the core components of the SPM were unclear, training was insufficient, and organisational factors (i.e., high caseloads, high staff turnover, inadequate support from leaders) hindered its implementation. Finally, to aid future implementation and research, a programme theory, including a detailed description of the SPM's core components, was formulated together with key informants involved in the national development and dissemination of the SPM.

Keywords: Systemic Practice Model, children's social care, child protection, realist evaluation, effectiveness, outcomes, implementation, programme theory

TIIVISTELMÄ

Aaltio, Elina
Suomalaisen systeemisen lastensuojelun toimintamallin vaikuttavuus. Realistinen arviointi.
Jyväskylä: University of Jyväskylä, 2022, 131 p.
(JYU Dissertations
ISSN 2489-9003; 571)
ISBN 978-951-39-9224-8 (PDF)

Tässä väitöstutkimuksessa arvioidaan *systeemisen lastensuojelun toimintamallin* vaikuttavuutta realistisen arvioinnin viitekehyksessä. Systeeminen lastensuojelun toimintamalli on suomalainen sovellus *Reclaiming Social Work* -mallista (RSW), joka kehitettiin alun perin Hackneyn sosiaalitoimistossa Lontoossa lasten ja perheiden sosiaalityöhön. Aiemmassa tutkimuksessa RSW-mallin on todettu parantaneen asiakastyön laatua ja työntekijöiden työhyvinvointia Hackneyssä. Mallin implementointi uusilla paikkakunnilla on osoittautunut haastavaksi, mutta työntekijöiden kokemukset ovat kuitenkin olleet pääosin myönteisiä. Mallin vaikuttavuudesta ei toistaiseksi ole juurikaan tutkimusnäyttöä.

Väitöstutkimus koostuu neljästä osatutkimuksesta, joissa on käytetty kolmea eri tutkimusasetelmaa ja osa-aineistoa. Kahdessa kvasikokeellisessa osatutkimuksessa vertailtiin asiakastason tuloksia systeemisissä (asiakastapauksia n=56) ja tavanomaisissa (n=56) tiimeissä toistomittausaineistolla. Kolmannessa osatutkimuksessa arvioitiin mallin implementoinnin fideliteettiä ja siihen vaikuttaneita tekijöitä monimenetelmällisellä tutkimuksella, jossa laadullista ja määrällistä aineistoa kerättiin rinnakkain ja tuloksia tulkittiin yhdessä. Neljännessä osatutkimuksessa mallille muodostettiin ohjelmateoria tutkijoiden vetämien työpajakeskustelujen ja asiantuntijahaastattelujen pohjalta.

Tutkimuksen päätulos on, ettei systeeminen työskentely ollut toimintamallin pilottivaiheessa tavanomaista asiakastyöskentelyä vaikuttavampaa tutkimuksessa käytetyillä mittareilla tai asiakastyytyväisyydellä mitattuna siitä huolimatta, että systeemisissä tiimeissä asiakastapaamisten määrä oli tavanomaista korkeampi. Mallin implementointia tarkastelleessa tutkimuksessa paikannettiin useita tekijöitä, jotka estivät työntekijöitä soveltamasta toimintamallia ja systeemistä asiakastyötä tarkoitetulla tavalla. Näitä olivat systeemisen mallin epäselvyys, koulutuksen riittämättömyys ja organisaatioon liittyvät tekijät, kuten korkeat asiakasmäärät, työntekijöiden suuri vaihtuvuus ja johdon tarjoaman tuen puute. Mallin tulevaa käyttöä ja tutkimusta varten mallille muodostettiin ohjelmateoria ja kuvaus mallin ydinelementeistä yhdessä mallin kansalliseen kehittämiseen ja levittämiseen osallistuneiden avainhenkilöiden kanssa.

Avainsanat: systeeminen lastensuojelun toimintamalli, lasten ja perheiden sosiaalityö, lastensuojelu, realistinen arviointi, vaikuttavuus, tulokset, implementointi, ohjelmateoria

Author's address

Elina Aaltio
Department of Social Sciences and Philosophy
University of Jyväskylä
elina.m.aaltio@jyu.fi
ORCID 0000-0003-4834-5473

Supervisors

Johanna Kiili
Department of Social Sciences and Philosophy
University of Jyväskylä

Mikko Mäntysaari
Department of Social Sciences and Philosophy
University of Jyväskylä

Kaisa Aunola
Department of Psychology
University of Jyväskylä

Sirpa Kannasoja
Department of Social Sciences and Philosophy
University of Jyväskylä

Reviewers

Riitta Vornanen
Department of Social Sciences
University of Eastern Finland

David Wilkins
School of Social Sciences
University of Cardiff

Opponent

Louise Caffrey
School of Social Work and Social Policy
Trinity College Dublin

FOREWORD

I am grateful to many people who have given me their time and support, advice, expertise and wisdom, and always valuable critical reflections on this project. Producing research knowledge requires many hours sitting alone at one's desk (wherever it may have been located during these turbulent years). However, I seldom felt lonely. As I was reading, writing or analysing the data, I always had an imaginary companion to debate with, spur me on, or ask me to refine my arguments. And quite often, thanks to today's highly digitalised communication systems, real-world reflections were only an email, phone call, Teams meeting or Messenger message away despite the physical distance between my home and the university, or COVID-19 restrictions. But, as an old-fashioned type of person, I have enjoyed the most my many face-to-face meetings in various congenial locations, such as Jyväskylä, Helsinki, and Cardiff, with my supervisors and with experienced researchers, methodologists, other doctoral students, and practitioners and experts working in the field.

My sincerest thanks go to my supervisors, Mikko Mäntysaari, Johanna Kiili, Kaisa Aunola, and Sirpa Kannasoja at the University of Jyväskylä. You have provided me advice and wisdom from multiple perspectives. I highly appreciate my lengthy discussions with Mikko on the philosophical underpinnings of social work research, his encouragement during the various phases of this project, and hilarious anecdotes I have enjoyed both during supervision sessions and over conference dinner tables. I am grateful to Johanna for her expertise in child protection research and her uncompromisingly honest style of supervision. Johanna has the wonderful talent of being able to pay attention both to detail and to the big picture at the same time. I am also grateful for having Kaisa, from the Department of Psychology, as one of my supervisors. Our original idea was to collect additional outcome data and analyse this together, but the pandemic intervened and I could use only the data I had collected during the initial stage of SPM implementation. Nevertheless, Kaisa's advice on the best strategy for statistical analyses and on research reporting standards have been of great importance. Finally, I thank Sirpa for near-term support in applying new statistical measures and interpreting the findings. Sirpa is a careful reader and, even more careful than me, in refining arguments and conclusions. In addition to my official supervisors, I have received invaluable guidance on statistical analyses from Professor Asko Tolvanen in the Department of Psychology. I am very grateful for our Teams meetings. I also wish to thank Tuija Kotiranta, who with Mikko was my supervisor at the beginning of my PhD project. And finally, I had the excellent opportunity of receiving face-to-face supervision from Professor Donald Forrester during my visit to the University of Cardiff in spring 2020. Unfortunately, this visit was cut short due to the pandemic, but I am grateful to all the members of CASCADE for the discussions I was able to participate in.

I warmly thank the reviewers of my PhD manuscript, Professor Riitta Vornanen at the University of Eastern Finland and Reader David Wilkins at the Uni-

versity of Cardiff. Your insightful comments have aided me to refine the manuscript in this final stage. In addition, I thank Assistant Professor Louise Caffrey for acting as my opponent.

I am grateful for the opportunity I have had to work full-time from March 2019 to December 2021 and part-time during spring 2022 as a doctoral student in the University of Jyväskylä. During these years I lived first in Helsinki and then in Espoo. Hence, I was not, unfortunately, able to meet my co-workers very often. Luckily, trains to Jyväskylä became faster and cheaper during these years and I very much enjoyed my trips and special routines when visiting Jyväskylä either for a seminar or teaching. I wish to thank Professor Marjo Kuronen, Head of the Department of Social Sciences and Philosophy, for her support and readiness to give advice even on smaller issues. I thank Marjo and Professor Heli Valokivi for leading the doctoral seminar and commenting on my work. Similarly, I wish to thank all my fellow doctoral researchers in our seminar for commenting on my manuscripts. I wish to thank Maija Mänttäri-van der Kuip for innovating a research project on practitioner's work-related well-being (which we were unable to execute in the end) and for advice in running the course on quantitative research methods for social work students. I thank Professor Kati Närhi and Heli for their advice and support as head of discipline. In addition, I wish to thank all my co-workers in social work for scholarly discussions and recreational activities during these past years. I also wish to thank Professor Anna Rönkä and Professor Kimmo Jokinen for running a multidisciplinary doctoral seminar for Child and Family Research, in which I had an additional opportunity to present my papers. Finally, many thanks to Michael Freeman for language editing throughout my PhD process.

I am most grateful to all local organisations, practitioners, children and parents who participated in this research. I wish to thank the National Institute for Health and Welfare (THL) for the opportunity to work there as a researcher from March 2017 to February 2019. During this period, I had the opportunity to collect the data for my PhD and work in a team expert in child protection. Special thanks to the members of the national evaluation steering group, Tarja Heino, Pia Eriksson, Päivi Petrelius, Johanna Hietamäki, Minna Kivipelto and Päivi Santalahti. I wish to thank Katarina Fagerström and Pekka Aarninsalo for their insights on systemic practice, and the members of the SPM training programme steering group for their comments and discussion during the national evaluation. I would also like to thank Jukka Mäkelä for commenting on my research forms. Warm thanks to my co-workers in the unit of Children, Young people, and Families and its director Johanna Lammi-Taskula. Finally, a special thanks to Johanna Hietamäki, who suggested over lunch in THL cantine that I should contact Professor Mäntysaari in Jyväskylä and ask for supervision if I were interested in evaluation research.

My closest companion during this project has been doctoral researcher Nanne Isokuortti at the University of Helsinki. Without knowing each other's plans, we had simultaneously decided to explore SPM in our dissertations. When this came up, I was first worried that our projects would overlap. I very soon

realized that our projects should indeed overlap, and that this coincidence was a great resource for both of us as doctoral students and THL, which gained additional data and analysis for the national evaluation. The THL report was our first joint publication. We have since published four peer-reviewed articles, commentaries, and a working paper, and have had a number of presentations both in Finland and in international conferences. Nanne is a brilliantly well-organized and hardworking researcher who shares my passion for finding new and better ways to conduct evaluation research. We have had endless reflective discussions about SPM and research in general, and I have always been able to rely on her critical eye and empathetic support. I also wish to thank Nanne's supervisors Professor Mirja Satka and Professor Nelli Hankonen for commenting our work. In addition, many thanks to my co-authors, Taina Laajasalo, Jane Barlow, and Niko Flink, for papers closely connected to but not included in this thesis.

During the final year of my PhD project, I was able to work as a researcher in the ETU project in the Center of Excellence on Social Welfare in the Metropolitan Area (Socca). I wish to thank Professor Tarja Pösö for leading this research project and the whole research team comprising Noora Aarnio, Maija Haapala, Eveliina Heino, Anne-Mari Jaakola, Tuuli Lamponen, Jenni Repo, and Minna Veistilä. I also wish to thank Pirjo Marjamäki, Kaisa Pasanen and Miia Pitkänen for their support, and all my other co-workers in Socca.

Finally, I wish to thank my friends and family for their support, interesting debates, and leisure-time activities. Most especially I wish to thank my godfather Jukka for offering my family a place to live when we suddenly had to leave Cardiff, and my friend Elisa for arranging me a place to work and my mother Sirkka for accommodating us when our house was being renovated. Most of all, I am grateful to my husband Jussi for taking care of the household and for being patient while I concentrated on my research. The beginning of this project was shadowed by a great sorrow in my childhood family - thank you, Jussi, for being there for me and my whole family. We have also had great fun and memorable adventures during this period, especially in Southampton, Cardiff, and Viitapohja, and finally some lovely (but less ecological) experiences of spending time in the countryside all year round. I also want to thank you for sharing my passion for trying to understand the causal relations of our social world and means of explore these. It was rather useful to have such company at hand in our home office during the period of social distancing, but also after office hours, when the greatest insights seem to come about. Finally, I thank Anton for being the light of my life and reminding me each day of what is the most important in life.

I dedicate this thesis to my mother Sirkka and to the memory of my beloved grandmothers, Martta and Anni. From you I have inherited the determination and perseverance that have taken me this far.

Espoo 1.11.2022

Author

FIGURES

TABLES

CONTENTS

ORIGINAL PAPERS

1. Aaltio, E. (accepted). Evaluating the effectiveness of the Systemic Practice Model of children's social care – A pilot study on child- and family-level outcomes. *Children and Youth Services Review*.
2. Aaltio, E. & Kannasoja, S. (accepted) Achieving agreement on service needs in child protection. Comparing children's, mothers', and practitioners' views over time and between approaches. *Journal of Children's Services*.
3. Isokuortti, N. & Aaltio, E. (2020) Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland. *Children and Youth Services Review, 119*, Article 105647.
4. Aaltio, E. & Isokuortti, N. (2021) Developing a programme theory for the Systemic Practice Model in children's social care: key informants' perspectives. *Child and Family Social Work*, 27(3), 444-453.

1 INTRODUCTION

1.1 Background

Is child protection effective? When I began my first research project on this issue ten years ago, this question was not as topical as it is today. The effectiveness of child protection, or for that matter social work and social services in general, was unquestionable for many practitioners and researchers. It was generally assumed that if practitioners only had enough time for face-to-face meetings with service users, social work would automatically be effective. However, this was not provable, as it was not possible to measure the relevant aspects of this work. I was often told that the very essence of social work, the unique needs of each individual or a family and the tailored responses to those needs cannot be measured: they are not comparable across clients, and standardised instruments cannot detect important nuances in human life. As suggested by Gambrill (2001), social work leaned more on authority-based than evidentiary criteria, and thus was an authority-based field of practice. Gambrill argued that social work professionals had considered good intentions as evidence of good outcomes or, based on their professional status, simply declared that social work was effective. Similarly, Sundell et al. (2010) claimed that Swedish social care services were dependent on opinion-based rather than evidence-based interventions. Although Gambrill's analysis focused on the situation in the United States, Mäntysaari (2003) suggested it was similar in Finland. To resolve these problems, Gambrill and Sundell et al. argued for evidence-based practice, while Mäntysaari proposed doing realist research on social work interventions and evaluating their outcomes. From the perspective of administrative science, Rousu (2007) analysed child protection organisations' practices in monitoring and assessing their services. Rousu found that these organisations were primarily guided by their 'social instinct', i.e., experiential knowledge and reflective discussions, rather than formal performance assessment protocols. The data used for assessment was often qualitative, narrow and random, and indicated outputs rather than outcomes. Rousu suggested that organisations would benefit from more systematic gathering and analysis of data

regarding service users and service outcomes. However, up to the present, no national guidelines exist that would aid local child protection organizations to produce outcome data or monitor the change in key indicators, e.g., the primary reasons for social services involvement. In the context of social work research, the idea of evaluating social work effectiveness received a cautiously optimistic response as Pohjola et al. (2012) published a compilation that included analysis of the concept and some empirical examples of evaluation studies. However, Pohjola (2012) argued that the general understanding of what research methods and 'evidence' were considered acceptable was too narrow to be applied in the field of social work. Similarly, the editors of a compilation on Finnish practice research regarded evidence-based practice too simplifying while practice research was found to be more contextual, critical, and reflective (Satka et al., 2016).

It seems that attitudes to evidence-based practice, outcome measures, and evaluation research have shifted in a more positive direction over the past few years. For example, several research projects that received funding from the new research instrument administered by the Ministry of Social Affairs and Health focus on evaluating social work interventions and their effectiveness (Ministry of Social affairs and Health, 2020, 2021, 2022). Yet, the kind of research needed to compare evidence on the effectiveness and usefulness of one social work intervention over another is still very limited. While development projects focusing on improving the quality of social care services may provide reports and testimony on the usefulness of innovations, these are rarely backed by evidence gathered using scientific methods, and hence are at a high risk for bias in various respects. At the same time, experimental study designs that would provide robust evidence of intervention outcomes have rarely been used in social work research either in Finland or elsewhere (Fisher, 1973; Holosko, 2010; Isokuortti et al., 2020; Olsson & Sundell, 2016; Rosen et al. 1999; Sheehan et al., 2018). Thus, it seemed to me when beginning this project that the instruments, methods and motivation to provide scientific evidence on the effectiveness of social work were lacking. The aim of this dissertation research was to find a meaningful way to explore the effectiveness of a specific child protection intervention by applying a quasi-experimental study design in the context of statutory social work in Finland.

While social work research has been reluctant to operationalise social work practice as something that can be measured and compared, the arrival of New Public Management (NPM) changed the way public servants and decision-makers plan and monitor public services and expenditure, including in social work organisations (Morago, 2006). In his grand plan for Finland, "The Social Policy for the Sixties" (1964 [60-luvun sosiaalipolitiikka, 1961]), Pekka Kuusi argued that social care is an investment in human capital and accelerates economic growth. In the 1990s, after nearly three decades of expansion, the ethos supporting the Finnish welfare state changed with the advent of neoliberal ideas, concrete NPM reforms, and a severe economic crisis (Julkunen, 2001). Now, a large public sector was no longer seen as underpinning but as hindering the economic growth and competitiveness of the Finnish economy. Increased productivity of

the public sector was called for. In practice, this meant budget cuts, lay-offs, and the marketisation of welfare services (Koskiaho, 2008; Patomäki, 2007). Economists were now interested in measuring the productivity of these services. Productivity refers to the ratio of input to output. In social work, production inputs are labour and capital, for example social workers and the cost of their wages, office space and equipment. Outputs can be measured as the number of service users or administrative actions. For many economists, productivity is the key source of economic growth, and hence also a prerequisite for a welfare state. For social work, however, measuring productivity is potentially harmful. Social work is social interaction between a practitioner and service user. Social work interventions are primarily different forms of communicating, discussing, and reflecting with service users. Increasing the productivity of social work means increasing the number of service users per practitioner. This in turn means decreasing the amount of time a practitioner can spend in interaction with each service user and in planning this interaction or related interaction with other professionals and family members in the service user's network. High caseloads and lack of time have already caused social workers moral distress and low work-related wellbeing (Mänttäri-van der Kuip, 2015). From this perspective, it is understandable that the field has shown reluctance towards the idea of quantitative measurement (Kotiranta & Mäntysaari, 2017).

Although it is tempting to disengage in harsh and tedious debates with mainstream economists, I was interested in finding ways to fight fire with fire – in this case, numbers with numbers. To have more nuanced discussion on the effects of cuts in services or in investment budgets on the wellbeing of children and families, we need empirical data on the relationship between productivity and effectiveness on the service-user level. In this way we can contribute the kind of evidence and arguments to the public debate that are also comprehensible to economists and economic-driven decision-making. Currently, two discourses are competing in the public arena: the need for budget cuts to increase the productivity of public services, and the need to ensure adequate resources for public services to be effective. In the absence of concrete knowledge on the effectiveness or unit cost of specific social work interventions this debate is based solely on unsupported claims.

Even without the need to prove anything to economists, social work has an ethical responsibility to ensure that its methods and daily practices are valid. We cannot simply take for granted that if only we use enough time, individuals or families will get the help they need. We have no evidence to support the argument that any interaction whatsoever between a practitioner and a service user is effective or even to some extent beneficial. In health care, decisions on the safety and health of a patient are based on empirical evidence. According to the Finnish Health Care Act (1326/2010), "[t]he provision of health care shall be based on evidence and recognised treatment and operational practices." In social care, the base for decision-making is different. The Finnish Social Welfare Act (30.12.2014/1301) states that the service user has a right to receive good quality

social care from the service provider. Decisions should be based on "the best interest" of the service user. The law does not mention "evidence". However, the idea of evidence-based practice seems to be embedded in the definition of social work given in chapter 3, section 15: "Social work refers to practice and expertise that builds a unity of support and services that respond to the needs of an individual, a family, or a community, integrates this unity with the support provided by other services, and guides and monitors its execution and effectiveness." [Translation EA] Instead of using research evidence as the basis for providing effective and good quality services, social workers are expected, on their own, to evaluate the effectiveness of the services they and their colleagues provide. This task is quite a challenge. To date, no consensus exists on social work outcomes or standardised instruments to measure them in either social work practice or research. Hence, practitioners have no comparable data on changes in the safety, wellbeing and service needs of service users. Moreover, monitoring one's own cases is not considered to yield reliable evidence on the effectiveness of one's practice. It can be useful for both the practitioner and service user as tool for observing change, but it does not constitute scientific evidence.

Despite the lack of scientific knowledge on social work outcomes, the need to improve its practice is widely accepted. Local-level authorities, i.e., the municipalities responsible for providing social services, have implemented various projects to develop the processes, tools, and methods of social work practice. Similarly, the Finnish government has launched various large-scale projects to reform, among others, the social services provided for children and families (e.g., Kaste Programmes 2008-2011 and 2012-2015, LAPE Programmes 2016-2018, 2019 and 2020-2022). However, instead of implementing evidence-based interventions or creating new ones based on research knowledge, stakeholders have based their decisions more often on values, ideas, and experiences collected from practitioners, experts, and service users. Swedish scholars have labelled this opinion-based as opposed to evidence-based social work (Sundell et al., 2010). In 2016, the Finnish government launched LAPE, a national programme aimed at reforming child and family services. Among other projects, the programme included the implementation of a new framework for child protection, the Systemic Practice Model (SPM). The SPM was a Finnish adaptation of Reclaiming Social Work (RSW), a practice model for child and family statutory social work, that had been developed in London borough of Hackney in the first decade of 21st century. Two evaluation reports (Cross et al. 2010; Forrester et al. 2013) concluded that RSW had improved the quality of practice and work-related well-being in Hackney. In Finland, some ideas drawn from RSW had been applied in one child protection team in the small municipality of Mäntsälä. The positive experiences of the practitioners in Mäntsälä and in Hackney were of value to experts and academics participating in workshops set up to prepare proposals for the projects to be included in the LAPE programme.

My own and the SPM's paths crossed by chance. I had earlier had an opportunity to ponder how to evaluate the effectiveness of child protection in a re-

search project which attempted to develop a structural assessment tool for statutory child protection (Aaltio, 2015). The idea was that if we could find valid instruments to measure change in the factors that are associated with the need for child protection, we could also use them to compare the effectiveness of child protection units, service providers or specific interventions. After nearly three years of development, trial and error, I planned to proceed with analysing the psychometric properties of the instrument. However, the organisation responsible for the project hesitated over its funding for so long that I decided to leave. When the Finnish Institution for Health and Welfare (THL) invited applications for the post of a researcher to evaluate the national implementation of the SPM, I applied for the job. I had been looking for a case, and suddenly here it was – an innovation that promised to improve child protection outcomes. I then worked for THL for two years and conducted the initial evaluation of the SPM.

1.2 Objectives and research questions

The aim of this dissertation was to study the effectiveness of the *Systemic Practice Model* (SPM) for child protection within a realist evaluation framework. Here, effectiveness refers to desired effects of a program or policy under real-world conditions (Flay et al., 2005). The aim of realist inquiry is to understand what works, for whom, under what circumstances, and why. In the case of the SPM, the programme was disseminated and developed simultaneously. This had implications for my research. I had begun by exploring the outcomes in accordance with the initial logic model, but soon realised that it was not clear with what resources and mechanisms the SPM was expected to achieve these outcomes. Hence, formulating the theory underlying the programme and identifying its various components was needed before the workings of the mechanisms could be explored. My study thus had multiple interconnected objectives designed to accumulate research knowledge of the SPM from various perspectives. The research questions were as follows:

1. Does the SPM work? Does the SPM produce the desired outcomes more effectively than service as usual?
2. In what context does the SPM work or not work, and why? What kind of mechanisms were activated or un-activated during the implementation, and how did this lead to the observed outcomes?
3. Why should the SPM be effective? What are the components and mechanisms that make the SPM different and more beneficial compared to service as usual, and in what context these can work as intended?

This dissertation comprises four complementary articles that report the findings of four sub-studies on the SPM accumulated during the research process. In this summary, I synthesise the findings of the four individual sub-studies. The four sub-studies and the research questions they address are presented in Table 1.

The first sub-study reports a quasi-experimental study evaluating change in the well-being and service needs of children and families in children's social care. The second sub-study explores agreement between child, mother and practitioner over time and between approaches, and the relationship between agreement and child protection outcomes. Both sub-studies evaluate the outcomes – the improvement in well-being and shared understanding between the practitioner and the family – of the SPM. The third sub-study evaluates the implementation of the SPM by assessing the fidelity and factors that hinder or support quality implementation. The fourth sub-study presents a programme theory for the SPM.

TABLE 1 Individual sub-studies and research questions.

	Sub-studies	Research questions
A1	Aaltio, E. (accepted). Evaluating the effectiveness of the Systemic Practice Model of children's social care – A pilot study on child- and family-level outcomes	Does the SPM produce the desired outcomes more effectively than service-as-usual with respect to a child's wellbeing and safety, the family's dynamics, the need for child protection, and the quantity and quality of practice?
A2	Aaltio, E. & Kannasoja, S. (accepted) Achieving agreement on service needs in child protection. Comparing children's, mothers', and practitioners' views over time and between approaches	Does the SPM produce better mutual understanding between the social worker, child and parent compared to service-as-usual? Is agreement associated with outcomes?
A3	Isokuortti, N. & Aaltio, E. (2020) Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland	In what context does the SPM work? Was the SPM implemented as intended? What were the barriers and facilitators that were connected with the success or failure of the SPM implementation?
A4	Aaltio, E. & Isokuortti, N. (2021) Developing a programme theory for the Systemic Practice Model in children's social care. Thematic analysis of key informants' perspectives	Why should the SPM be effective? What are the core components and mechanisms that make the SPM more effective and different compared to service-as-usual?

The aim of this summary is to introduce the Finnish context in which the SPM was implemented, describe the original RSW model and the initial ideas of the SPM, give an overview of previous research on effectiveness in the context of child protection, specify the ontological and methodological premises of this dissertation project, present the data and methods applied in the articles, synthesise the findings, sum up the conclusions, and, based on the findings, make recommendations for further research and practice.

2 THE FINNISH CONTEXT AND THE SYSTEMIC PRACTICE MODEL

2.1 The ideal and current challenges of the universal welfare state in Finland

As suggested by Lorenz (1994), the distinctions between welfare state models and structural factors explain the distinctions in the tasks of social work and its forms or practice in different welfare states. In the literature on welfare state models, Finland, along with the other Nordic countries, has been categorised as an institutional, socio-democratic or universal welfare state.

In Titmuss's classification (1974), an institutional welfare state is based on the idea of achieving greater equality through comprehensive social provision that guarantees benefits to all citizens, whereas the residual welfare state model provides only temporary aid if other channels of welfare provision fail, and the industrial-achievement welfare state model complements the market economy by providing a minimum standard of social security. According to Titmuss, the Anglo-Saxon countries were seen as examples of the residual welfare state model, the Central European countries as examples of the industrial-achievement model, and the UK and Scandinavia as examples of the institutional model.

Based on comparative empirical research, Esping-Andersen (1990) identified three diverse regime types: 'conservative', 'liberal' and 'social democratic'. In a conservative welfare regime, social insurance typically excludes non-working wives, family benefits encourage motherhood, and family services are underdeveloped. A liberal regime provides means-tested assistance or modest universal transfers to low-income citizens with strict entitlement rules. A social democratic regime seeks to promote the highest equality by providing upgraded services and benefits to all and granting equal rights to all socio-economic classes. A universal welfare state grants transfers directly to children and provides care for them and other vulnerable citizens. In Esping-Andersen's classification, the United States, Canada and Australia are archetypical examples of the liberal

model. Austria, France, Germany and Italy are examples of the conservative model, and the Scandinavian countries of the social democratic model.

The universal system has been associated with improved societal outcomes. For example, Korpi and Palme (1998) found that providing high-income earners with earning-related benefits can reduce inequality and poverty more efficiently than flat-rate benefits or benefits targeted to low-income earners. The researchers suggest that this encompassing model generates incentives to work, avoids poverty gaps, and provides citizens with significant benefits in return for their taxes.

The expansion of the Finnish welfare state began in the 1960s. In 1961, Kuusi published a comprehensive social policy programme "60-luvun sosiaalipolitiikka" ["A Social Policy for the Sixties"]. At the time, this programme was the key reference point in social policy debates and decision-making (Anttonen Sipilä, 2000; Riihinen 1992). Although many of the policy measures suggested in this programme had already been under debate or preparation, scholars saw social policy as too fragmented. Hence, the Social Policy Association in Finland invited Kuusi to write a comprehensive research-based programme for social policy (Tuomioja, 1996). This was a broad programme covering unemployment benefits and jobs-creating investments, housing, family benefits, health care, insurance for the elderly and disabled, and social care, mainly in the form of poor relief. The emphasis of the programme was on cash benefits and economic growth. Kuusi was inspired by the Keynesian idea of the state ensuring a sufficient level of aggregate demand, i.e., consumption and investment expenditure, along with demand management with the aim of reaching full employment (Keynes, 1936). In addition, Kuusi applied Myrdal's theory of cumulative growth (Myrdal, 1957), according to which accelerating economic growth in under-developed regions simultaneously accelerates human and social growth. Through social transfers, social policy activates passive members of the population to take the initiative.

In the mid-1960s, the Finnish social policy debate was complemented with the idea of universal social services. In the UK, Titmuss (1968) had argued that a welfare state should provide universal social services and selective services aiming to discriminate positively "in favour of those whose needs are the greatest." However, Titmuss's writings had not been discussed among key figures in Finnish social policy (Mäntysaari, 2013). According to Rauhala (1996), the idea of universal social services was derived from the Beveridge report (1942).

The main driver of this turn was the increasing number of women joining the labour market (Rauhala, 1996). This meant that the care of children and the elderly had to be re-organised. Previously, these services had been targeted to the poor, but now it was necessary to provide them universally to all citizens. In the following decades, the provision of public services expanded considerably (Anttonen & Sipilä, 2000). The quality of services and the level of education and skills of personnel were increased. The middle-class nature of the services increased the demand for them, expanding the service system and employment opportunities for women. Both the political left and the agrarian parties sup-

ported this change, as the services created greater employment and equality between regions. Most of all, these services were essential to ensure women's independence and gender equality (Hellsten, 1993).

Nevertheless, the gradual adoption of the tenets of the New Public Management from the 1980s onwards, the severe economic crisis of the early 1990s, and austerity following the global financial crisis of the 2000s have greatly altered the welfare state paradigm and the national and local context in which social services have been carried out (Julkunen, 2001; Jutila, 2011; Kantola & Kananen, 2013). The funding of public services has been squeezed and services have been gradually weakened (Jutila, 2011). The financialisation of care services has transformed them into site of profit extraction and financial engineering, a shift which has made service production less transparent and democratic (Hoppania et al., 2022). In child welfare services, around 80 percent of residential care units were run by for-profit companies in 2018 as compared to only 23 percent in 1988 (Shanks et al., 2021). Although the insufficient financing of social services and their privatisation has been met with concerns over the well-being of service users and practitioners (Hoppania et al., 2022; Shanks et al., 2021; Sihto & Van Aerschot, 2021), attempts to decrease the 'fiscal sustainability gap' continues to dominate the public debate and decision-making (Sorsa, 2017). The fiscal sustainability gap is an arbitrary calculation of the ratio between the increase in service needs due to population ageing and long-term estimates of public-sector income. The size of this gap in euros is calculated in the Ministry of Finance and based on simplistic models and simplifying assumptions on the long-term development of the economy and public finances (see Ministry of Finance, 2020). However, in the domains of political economy and social policy, understanding of what constitutes a 'sustainable' level of public expenditure in the public debate seems to have narrowed down to this one, highly debatable, indicator at the expense of indicators more relevant to social and ecological sustainability (Sorsa, 2017). Partial optimisation, aimed at achieving short-term cutbacks, has led to discontinuity and ineffectiveness in the social security domain, including in child protection services (Metteri, 2012; Sipilä & Österbacka, 2013). High caseloads, insufficient resources, lack of organisational support, high personnel turnover, and demanding job descriptions have caused for workers in social services to experience negative stress (Junnonen et al., 2019; Van Aerschot et al., 2021). More especially, owing to insufficient resources, social workers suffer from moral distress, which has impaired their well-being and motivation to remain in the sector (Mänttäri-van der Kuip, 2016). Policy makers have acknowledged the need to improve services, although this is expected to happen via temporary projects rather than as a permanent increase in resources.

2.2 The Finnish child protection system – key features and identified problems

In Finland, child protection refers to both preventative services provided for the whole population in primary services, and child protection services provided for children and their families if the child has been assessed as in need of protection. According to the Child Welfare Act 2007, preventative child protection seeks to promote and ensure a child's growth, development, well-being and parenthood. While "child and family specific" child protection shares these goals, it also obliges authorities to take action in the "best interest of the child". Child protection services include open care (i.e., services provided for children living at home with their parents), emergency placements, care orders, out-of-home care, and aftercare. The legislation states that the authorities should favor open care unless the child's best interests demand otherwise (Child Welfare Act 2007, 4 §).

In 2019, 4.0 percent of Finland's 0- to 17-year-olds were open care clients of child protection and 1.3 percent placed outside the home. A child welfare report had been filed for 7.9 percent of the same age group (Lastensuojelu 2020). On the national level, no statistics exist on the reasons for initiating child welfare reports or on the child- and family-level factors of cases that have entered the child protection system. A case study (Heino, 2007), based on practitioner reports, exploring the characteristics of families (n=330) entering child protection in nine Finnish municipalities during a one-year period found that most typical parent-related factors were parental fatigue (37%), family conflicts (29%), and inadequate parenthood (28%). The most common child-related factors were conflicts with parents (22%), problems at school (20%), and poor mental health (14%). A study examining the risk factors for out-of-home care (Kestilä et al., 2012), conducted by analysing various registers, found that in a cohort of children born in Finland in 1987 (n=60 069), 3 percent had been placed outside of home at some point. On average, the first placement took place at the age of 10.7. Key risk factors were low parental education, mother's psychiatric inpatient care, mother's single parenthood at the time of birth, and financial difficulties in the family.

In an earlier comparative international study (Gilbert, 1997), the Finnish child protection system was identified as family service-oriented. The study analysed and compared social policies and professional practices pertaining to child abuse in nine Western countries. Two distinctive orientations were identified: a child protective orientation and a family service orientation. The former emphasises legalistic interventions and the latter therapeutic interventions. Finland, along with two other Nordic countries (Sweden and Denmark) participating in the study, was grouped into the family service category. England, the US and Canada were categorised as having a child protection orientation. According to Pösö's (1997) analysis on the Finnish system, child protection cases were generally treated in an undifferentiated manner as 'family problems.' Psychological abuse was not treated as a distinctive problem, nor unless medically diagnosed, was physical abuse. The only exception was sexual abuse, which was recognised

as a specific problem demanding immediate and specialised treatment. In general, the Finnish system has preferred family-centred practice and cooperation with parents. However, a demand for more individual and child-centred practice had already been expressed.

In a subsequent study (Pösö, 2011), the Finnish child protection system had shifted away from a family-centred policy towards a more child-centred orientation, as the new Child Welfare Act of 2007 had strengthened the role of children and their participation. Nevertheless, the terms child abuse and neglect were still not used. Practitioners assess 'the need for child welfare' instead of specific problems. The emphasis continued to be on supportive measures, prevention and open care, but now the focus was on "the child's best interest." This was expected to guide all decisions. However, the definition of the child's best interest in the legislation has left plenty of room for local interpretation. While the Finnish system had moved from a family service system to child-oriented system, the English system had sought to emphasise a more family service orientation, aimed at prioritising prevention and early intervention (Parton & Berridge, 2011). The findings of a comparative analysis (Gilbert et al., 2011) suggested that all countries previously identified as child protection-orientated had taken on some elements of the family service orientation. In addition, an emerging approach, i.e., a child-focused orientation, was identified. This orientation concentrated on the child as an individual with independent needs and rights. Common to all countries participating in the study was the rapid pace of organisational, policy, and legislative change. In addition, public demands to reform the child protection system had been expressed in many of countries after fatal cases of the abuse of children already known to the child protection authorities (Gilbert et al., 2011).

In Finland, the high-profile death of an 8-year-old girl under a child protection plan in 2012 resulted in several reports analysing the state of Finland's child protection services and making recommendations to reform child and family services. Two reports commissioned by the Finnish government (Kananoja et al., 2013; Sipilä & Österbacka, 2013), along with other reports (Alhanen, 2014; Sinko & Muuronen, 2013) identified problems such as impractical bureaucracy, high staff turnover, poor management, high caseloads and a lack of several critical factors such as service-user participation, early interventions, service coordination, continuity, and a comprehensive understanding of families' needs. In response, the Finnish government launched a comprehensive large-scale project to reform child and family services (LAPE) to be implemented during the period 2016-2019. One of the many ideas included in the programme was to develop and disseminate a new systemic practice model to guide statutory child protection. The new model was inspired by the *Reclaiming Social Work* model (RSW) developed in England, which had already inspired reform in one small municipality in Finland (Yle Uutiset, 28.9.2017).

While child protection orientations and legislation differ between Finland and England, the RSW model was expected to show a good fit with the Finnish system (Lahtinen et al., 2017). In terms of decision-making, the Finnish organisational culture already involved features that supported team-based and vertical

processes compared to the typical organisational culture in England. First, the English system involves more national regulation while the Finnish system is a deregulated system that allows more professional discretion (Berrick et al., 2015). Second, Finnish social workers are more used to involving co-workers or a team in the decision-making process whereas practitioners in England use more vertical institutional support structures and rely more on managers to authorise their decisions (Berrick et al., 2016). Finnish practitioners typically have a partner, e.g., a family practitioner with a lower qualification, to reflect with on cases, while English practitioners would typically turn to their manager for advice (Falconer & Shardlow, 2018). Nevertheless, in general Finnish social workers have a caseload over twice that of their English colleagues, and they have been found to experience higher work pressure (Berrick et al., 2016).

Although Finnish social workers have a wider support-network compared to their English counterparts, decisions are primarily made by the social worker responsible for the case. Decisions are made by other agents in certain cases. For example, decisions on voluntary care orders are made by the manager or team leader, whereas decisions on involuntary care orders (i.e., care orders in which one custodian or a child of 12 years of age or more opposes the care-order proposal) are made by the regional administrative court (see Pösö et al., 2016). In England, the 1989 Children Act and accompanying statutory guidance expect social workers to 'decide the nature and level of child's needs', while social work managers are expected to 'challenge the social worker's assumptions as part of this process' (Falconer & Shardlow, 2018).

2.3 The Systemic Practice Model (SPM)

2.3.1 Reclaiming Social Work (RSW)

The original RSW model was pioneered in the London Borough of Hackney starting from 2007. The model had been developed by Isabelle Trowler and Steve Goodman, who were social work directors in Hackney at the time. The model developers had not produced a manual but had published a description of the general ideas, values, and theories of systemic practice (Goodman & Trowler, 2012). For dissemination of the model, the co-founders had established a social enterprise, Morning Lane Associates (MLA), to provide consultation and training for local authorities interested in RSW implementation. At present, this work is being continued by Collective Space (https://www.collectivespace.org.uk/).

RSW sought to improve direct practice with families by building relationships, co-constructing problems and enabling change (Schiltroth, 2012). While relationships are, in theory, in the core of social work, child protection practice had not been successful in relationship building due to lack of continuity and a focus on bureaucracy. Hence, attention to the whole system was necessary to create a structure that supported systemic practice. In Hackney (Trowler & Goodman,

2012), this involved establishing a shared value base (e.g., collaboration and respect), delivering a set of evidence-based skills and interventions, creating structures and systems (e.g., procedures and routines) that facilitate effective working with families, improving leadership, selecting competent staff, developing a training programme to enhance the skills and knowledge base of the staff, and introducing small multi-disciplinary units. Previous research has found systemic family therapy effective in a variety of forms and contexts (Stratton, 2005). Thus, the developers of RSW considered that child protection and systemic family therapy would fit well together, although the considerable pressure on services could hinder workers from applying this new approach in practice (Pendry, 2012). Since the current structures and systems did not support good practice, both practice and structures needed to be changed.

In the original RSW model (Trowler & Goodman, 2012), the systemic units consisted of a social worker, a children's practitioner, a family therapist, a unit coordinator providing administrative support, and a consultant social worker who led the unit and had the overall case responsibility. Units held weekly meetings in which they discussed, debated, reflected and made decisions on all cases allocated to the unit. These meetings served as the main mechanism for case supervision (Trowler & Goodman, 2012). While practitioners were trained to embed the systemic approach in social work, the family therapist ensured that this approach was maintained (Pendry, 2012). In systemic practice, problems are seen to be embedded in relationships. Family members' actions are a response to others' actions. In this relational frame, the social worker is not expected to fix the family, but instead resolve the problem with members of the family's system (Pendry, 2012).

According to Pendry (2012), the key systemic theories underpinning the RSW model in Hackney were family therapy (Bateson, 1972, 1980), rooted in systems theory and cybernetics (Wiener, 1961), and the Milan school of family therapy (Palazolli et al., 1980). The idea of cybernetics was originally introduced in natural science, and it refers to the "field of control and communication, whether in the machine or in the animal" (Wiener, 1961, p.11). A cybernetic system is self-regulating, i.e., it maintains stability through information looping back into the system. Bateson applied this idea in the behavioural sciences to explain the dynamics of families. The Milan school (Palazolli et al., 1980), inspired by Bateson's work, introduced three guidelines for family therapy. These were hypothesising, circularity and neutrality. Hypothesising refers to a technique in which a practitioner is continuously looking for different descriptions and explanations for a given situation; circularity is a method by which a family therapist creates curiosity within the family and therapy systems by asking circular questions that use the language of relationships; and neutrality means actively avoiding the acceptance of any one position as more correct than another. Systemic family therapy evolves and integrates new ideas, such as social constructionist theory and narrative approaches. Common to all these ideas and to RSW was that problems are seen as embedded in relationships, i.e., they are interpersonal rather than in-

trapsychic (Pendry, 2012). These were complemented with Social Learning Theory (SLT) (Bandura, 1977; Herbert 1978, 1987). SLT developed within the domain of behavioural psychology and was chosen as one of the main approaches to problem resolution in RSW (McCafferty, 2012). The general idea in SLT is that behavioural development can be influenced by positive reinforcement of the desired behaviour, while the unwanted behaviour can be ignored.

2.3.2 Previous research on RSW

As further described in Section 3.2.3, previous research has not provided strong evidence of the effectiveness of RSW. However, several studies have addressed the content, user experiences, and implementation of the model both in Hackney and in new sites in the United Kingdom.

A mixed-method evaluation conducted in Hackney (Cross et al., 2010) concluded that social workers in RSW units were more satisfied with their work environment, social-work processes, and work-related wellbeing than colleagues in teams not applying the model. In addition, the number of children in care decreased by 30 percent between 2005 and 2009. However, as noted (but not further analysed) by Forrester et al. (2013), in large part this reduction occurred between 2005 and 2007, i.e., before the implementation of RSW.

In their realist-informed mixed-method evaluation conducted in Hackney, Forrester et al. (2013) found that practitioners in RSW units provided more intensive help for families and spent more time with them compared to conventional teams. In addition, practitioners demonstrated higher levels of direct practice skills and made higher-quality assessments. Although these findings suggest RSW reform improved the quality of practice, the researchers point out that all the practitioners in Hackney had to reapply for their positions. As described by Trowler and Goodman (2012), most applicants did not pass these rigorous tests. Hence, the positive changes identified in Hackney may partially be due to this key element of the RSW reform (Forrester et al., 2013). Based on observational data, the study outlined six features that distinguished the RSW units from conventional teams (Forrester et al., 2013, pp. 105-106): shared work, in-depth case discussion, a shared systemic approach, focused skills development, special roles in systemic units, and low caseloads. These features were interlinked. For example, sharing work necessitated more discussion of cases. It was observed that units had informal debriefings after family visits and structured in-depth discussions of every child and family on a regular basis, with a monthly discussion on each case. The study also identified key factors that supported good practice. Along with recruiting new staff, practitioners in RSW units had wider practical organizational support (e.g., adequate office space), strong administrative support, and more case supervision. The small size of teams was regarded as one of the key innovations of RSW, and caseloads were limited to ensure effective service delivery. Forrester et al. also pointed to the importance of articulating clear values, although these were also observed in conventional teams.

The RSW model has not been implemented in other UK agencies in its original form. The shortages of key elements in local implementation have resulted

in disappointing results. Qualitative studies conducted in one site (Laird et al., 2017, 2018) found only limited changes in social work practice. The fact that the site had not implemented any structural changes or provided sufficient training for managers had undermined the creation of shared values and shared teamwork (Laird et al., 2018). The social workers were already suffering from high caseloads and could not find time to build relationships and apply systemic tools and methods with their case families. A parallel analysis of case files (Laird et al., 2018) revealed the minimal use of systemic tools or a systemic approach in analysing a family's situation. In addition, practice was conducted around co-resident blood relatives of the child, predominantly with mothers, rather than, as intended, incorporating extended family members in the process. The study concluded that training alone was not enough to create change in child protection practice. The social workers at this site had received 18 days of training, which they highly valued (Laird et al., 2018). In turn, their managers had received only 8 days of training, a period which was considered insufficient for enabling them to provide systemic supervision for social workers and maintain the agreed approach. In addition, the ongoing high staff turnover rate meant that teams lost many of their trained social workers, thus making the maintenance of new ideas difficult (Laird et al., 2018).

In their mixed-methods study, Bostock et al. (2017) evaluated RSW implementation at five sites that had participated in a project aimed at embedding RSW into their practice with the support of MLA. The project involved special training for consultant social workers, a reduction in bureaucracy, the setting up of specialist RSW units targeting teenagers, and consultation with senior managers. All five sites had previously been involved with MLA and had formed systemic units. However, these had been replaced by larger systemic teams comprising three to four case-holding social workers and, in some cases, teams operated without a family therapist. The project had difficulties in recruiting CSWs, and the notions of MLA coaches on the organisational changes (e.g., smaller units and caseloads) and activities (e.g., external recruitment) needed for successful RSW implementation had not always been welcomed by local managers. Despite these challenges, practitioners primarily reported positive experiences of RSW. They appreciated the focus on reflexivity, the way unit meetings helped them to plan their sessions with the family, and the contributions of a family therapist and a unit coordinator. Bostock et al. concluded that good systemic practice required a consultant social worker trained in systemic practice, input from a family therapist, systemic case discussion, and dedicated administrative support.

In their subsequent study, Bostock et al (2019a) found that the role of family therapists was crucial in improving the quality of both supervision and direct practice. To assess the quality of systemic supervision, Bostock et al. (2019b) had identified the following key features of a systemic case discussion:

- patterns of family relationships and narratives are actively explored
- family relationships are set in the wider social context

- child and family focus are present within the conversation
- clarity about the potential risks to the child(ren)
- discussion is curious and reflective, e.g., open to different ways of thinking about the family
- different hypotheses are generated and/or established theories about the family are challenged
- discussion is collaborative and involves all group members
- hypotheses are developed into clear and actionable conversations with families

To be assessed as 'systemic', all the above elements had to be observed. In Bostock et al. (2019a), the association between the practice skills and the quality of supervision was explored by quantifying and pairing observations of supervision (n=14) with observations of direct practice (n=18). The quality of direct practice was assessed by rating recorded observations of family meetings with the coding framework developed by Whittaker et al. (2016). Based on quantitative analysis, the workers who had received non-systemic supervision scored the lowest in direct practice skills whereas those who had received systemic supervision scored the highest. The presence of a family therapist was associated with the quality of supervision. However, in their follow-up study, Bostock and Newlands (2020) identified further changes in the delivery of the model, including reductions in training, family therapist input and group supervision. In addition, to guide their practice, some sites had introduced a second approach (Signs of Safety) to accompany the RSW. Only one of the five sites in their sample had continued applying all the key RSW elements and offering training and group supervision to practitioners. In this site the number of children receiving any intervention from children's services between 2015-2019 demonstrated the largest reduction (-1.4%) compared to four other sites. The authors suggest that sustained commitment to systemic practice has beneficial outcomes for children and families.

2.3.3 Towards the Finnish Systemic Practice Model

The Finnish version of the RSW model was outlined in a workshop process organized by the Finnish Institute for Health and Welfare (THL) in 2016. This process resulted in a description of the Finnish Systemic Practice Model (Lahtinen et al., 2017). The authors argued that the organisation culture of Finnish child protection needed comprehensive reform. Problems identified included high caseloads, poor job-related well-being of practitioners, high staff turnover, increasing expenses, and problems in collaboration between child protection and other child and family services. In addition, it was noted that child protection practitioners were expected to cope on their own without sufficient support and supervision. The goal was to change the structures and resources to support practitioners and

transform practice in ways that would make it comprehensive and engage the client families.

In the Finnish model, it was suggested that a systemic team would consist of a consultant social worker, a family therapist, a unit coordinator, a case-holding social worker, other social worker(s) (such as a practitioner working with a parent), and child or family practitioners (if included in the local child protection teams). The recommended team size was 6-8 members. Teams would hold weekly meetings that would serve as a forum for case supervision, collective reflection, and learning. The core principles of the new model were as follows (Lahtinen et al., 2017, p. 16):

- adaptation of systemic culture
- family-therapeutic understanding and relationship-based practice
- child-oriented practice
- service-user participation and encounters with service users
- organisational structures that support the model
- shared values and coordinated management.

Owing to the general nature of the initial description of the model, the details of the SPM were negotiated and developed in a Training of Trainers (ToT) programme organized by THL in 2017-2018 (for more details, see section 6.2). The training applied the description of systemic family therapy proposed by Leeds Family Therapy Research Center and reported in Lorås et al. (2017). According to this definition, systemic family therapy involves 11 specific competences (Lorås et al., 2017):

- the central focus is on the system, rather than the individual
- behavioural patterns within systems are seen as circular and always evolving
- connections between circular patterns of behaviour and connections between beliefs and behaviour within systems should be considered
- paying attention to narratives and language
- applying the idea of constructivism, i.e., each individual will interpret and make sense of her world from her own frame of reference
- applying the idea of social constructionism, i.e., meaning is constructed in social interactions between people, is context-dependent and constantly changing
- considering the importance of cultural context in relation with cultural meanings and narratives
- taking a reflexive stance toward power differentials

- acknowledging that reality is constructed between the therapist and the service user
- applying self-reflexivity
- focusing on strengths and solutions.

However, it was acknowledged that child protection is not therapy (ToT training material, 2017). As applied in the child protection context, systemic practice involved collaborating and sharing expertise with families and other professionals and formulating goals and creating meanings together with families while acknowledging the social worker's authority position.

Since the ToT and subsequent local training programmes were not based on any previous training programme developed for RSW, the two models are not identical. The most important similarities between RSW and the SPM include practitioner training in systemic thinking, the inclusion of a family therapist in the team, the use of systemic tools and methods (i.e., genograms, formulating hypotheses) with families and having systemic weekly meetings. The composition of an SPM team, as proposed by Lahtinen et al. (2017), is different from that in the original RSW model: the SPM team is larger and includes a parent's social worker[1]. During ToT, new ideas were incorporated into the SPM, such as inviting the family to participate in a team meeting. New tools were also introduced such as the Collaborative Helping Map (Madsen, 2011) and Three Houses from the SoS. The use of SLT was not included in SPM training. It should be noted that the role of SLT in later RSW implementation has not been as central as it was in Hackney (Bosanquet and Goodman, personal communication, 2018) and it is not mentioned in studies on RSW dissemination (Bostock et al., 2017; Bostock & Newlands, 2020).

The concrete measures taken to disseminate the SPM were on the team level, whereas the original RSW was a whole system reform. In practice, no special training programme or on-site coaching was provided for managers during the initial stage of SPM dissemination in 2017-2018.

For the purposes of THL's national evaluation, I formulated an initial logic model for the SPM (Table 2). First, the logic model was derived from the SPM description (Lahtinen et al., 2017). Next, it was introduced to the ToT steering group. This group was led by THL and included trainers conducting the ToT. Child protection experts, some of whom had participated in formulating the SPM description were also invited. The initial logic model was refined based on discussions with the steering group and within THL. Finally, it was used as part of the dissemination material in training and workshops, and as a starting point for operationalising the data collection instruments.

[1] This idea was not implemented; instead, existing child protection teams were transformed into SPM teams by adding a clinician. However, this might explain the idea of inviting the family and its service network to weekly team meetings.

TABLE 2 The initial logic model for SPM.

Structures	Training	Practice	Outcomes
Team structure • Consultant social worker • Social workers • Family therapist • Coordinator Managers and leaders support local trainers and local teams Caseloads have been reduced to enable systemic practice Key partners have been identified and informed	Team members are trained in systemic practice and teamwork Practitioners have adopted key knowledge and skills	Case work is guided by case discussions conducted within teams Teams hold weekly meetings characterised by the following criteria • principles of systemic practice structure the discussion • consultant social worker leads the discussion • family therapist supports reflection • cases are discussed frequently and for long enough • documentation is transparent, promotes practitioner's and service users' understanding, and supports progress The whole team is familiar with all the cases that are the responsibility of the social workers in the team Team shares responsibility for these cases The majority of the work is face-to-face practice with families	Family feels they have received the help and support they needed Family feels they can affect decisions and design of their services Mutual trust and understanding between service users and practitioner Increased child well-being Increased parental ability to ensure child's safety and wellbeing Increased family capability Improved practitioner well-being and motivation, and decreased exhaustion and burden Practitioners perceive their work as more meaningful Increased safety in work Savings in costs Decrease in the number of involuntary care decisions The service system becomes more organized

2.3.4 Implementation of the SPM

The development and dissemination of the SPM was organized by the Finnish National Institute for Health and Welfare (THL) as part of the first stage of the government's key programme aimed at reforming child and family services (LAPE) during 2016-2018. The model was implemented between the autumn of 2017 and summer of 2018 in 31 municipal children's service sites located in 14 counties around Finland. The first stage of SPM implementation began in spring 2017, when THL selected participants for the national ToT programme, organised a kickoff seminar for leaders and managers responsible for local implementation, and published the SPM description. During 2017-2018, THL trained fifty-eight (n=58) social workers or family therapists to work in pairs as local trainers. ToT

consisted of seven training days and four supervision sessions. By the summer of 2018, local trainers had trained fifty-three (n=53) local teams[2] at 31 sites in use of the systemic practice model. Local training consisted of six days of training as a team. Topics included an introduction to systemic thinking and family therapy, team roles and the functioning of a systemic team, the structure of systemic weekly meetings, principles of systemic practice, documentation, and an introduction to selected methods and techniques (e.g., genogram, formulating hypotheses). Teams then implemented these components in their daily practice by holding systemic weekly meetings and applying systemic orientation and methods in casework.

[2] The number of these teams was erroneously reported as n=52 in A3.

3 PREVIOUS RESEARCH ON EFFECTIVENESS IN SOCIAL WORK AND CHILD PROTECTION

3.1 The experimental research tradition in social work

The first effort to examine the effectiveness of social work was the review by Joel Fischer published in 1973 to answer the question "Is casework effective?" In his review, Fischer examined whether professional case work (as delivered by practitioners with a master's degree from an accredited school of social work) in the United States was successful in helping clients. Effectiveness was measured as significant differences in scores between groups in achieving a goal specified in advance by the researcher. The minimum requirement for studies to be included in the review was that they had included a control group of some kind. The control group had either received no treatment at all or some form of non-professional assistance or support. Fischer surveyed recent reviews, major social work journals, dissertation abstracts, and unpublished agency reports from the 1930s onwards, and identified eleven studies that met the inclusion criteria. These dealt with children (n=6), low-income multiproblem families (n=3), aging (n=1), and young female probationers (n=1). A positive finding was that the studies had used more than one source of data, multiple outcome measures, and an acceptable study design, i.e., the allocation of research participants to study groups was based either on matching, randomisation or a combination of the two. However, the precise nature of the casework in question was inadequately defined. The main finding was that caseworkers had not been able to bring about any positive, significant, measurable change in their clients compared to controls. It is possible that the studies did not represent mainstream practice at the time, or that the context for casework, such as low-income clients, was such that caseworkers had few ways of helping their clients. In his conclusions, Fischer expressed the wish that future research would be able to validate new methodologies.

However, social work research did not extend to experimental research. Rosen et al. (1999) reviewed the studies published in thirteen US social work journals between 1993 and mid-1997. The sample comprised 1 849 articles, of which

863 (47%) were research articles. The non-research articles, such as non-systematic reviews of the literature, narrative case reports, or methodology papers without a substantive focus, were excluded from the further analysis. Of the research papers, 36 percent were descriptive in nature, 49 percent were explanatory, i.e., investigating relationships among two or more variables, and 15 percent were control-oriented, i.e., outcome studies evaluating the effects of services or interventions. The adequacy and usefulness of controlled studies was assessed by analysing whether the intervention was described in enough detail to permit practitioners or subsequent studies to replicate the intervention with minimal error. Only half of the controlled studies met this criterion, i.e., only three percent of all the published papers (Rosen et al., 1999).

Holosko (2010) analysed the study designs used in empirical social work research reported in the three most relevant, empirically oriented social work publications during the period 2005-2007. The sample comprised 329 papers, of which 23 percent (n=77) were non-research articles. Of the research articles, the majority (82%) were pre-experimental and most did not include an intervention. Five percent of the articles were quasi-experimental and only 2 percent had used a true experimental design (Holosko, 2010).

Olsson and Sundell (2016) examined 2 334 PhD theses published in Sweden in seven disciplines, i.e., public health, criminology, nursing, psychiatry, psychology, social work, and sociology, during the period 1997-2012. Of these, 13 percent investigated the effects of interventions. Of those in social work, 8 percent were outcome studies and only one-third of these were experimental or quasi-experimental designs. Olsson and Sundell concluded that future scientists are "ill prepared to produce the type of research that is necessary to inform practice of the effects of its interventions."

To summarise, the application of experimental or quasi-experimental study designs in social work research in general is scarce. One of the goals of this dissertation was to test a quasi-experimental study design and other potential research methods for studying the effectiveness of social work interventions in Finland.

3.2 Evaluating child protection on the micro, meso, and macro level

The effectiveness of child protection can be approached from various angles. The task is challenging, since child protection performs multiple functions, and no consensus exists on its outcomes (Forrester, 2017). To give an overview of empirical research on the effectiveness of child protection as an intervention or as interventions, it would be helpful to distinguish between 'child protection' and 'child protection interventions' according to the level of the analysis. On the macro level, the interest is in 'child protection' as a whole: a system of statutory child protection covering all stages, i.e., assessment, open care, out-of-home care

and after care, and all accompanying interventions. On the meso level, the focus is on one of these stages, e.g., out-of-home care or assessment. On the micro level, the focus of child protection research may be on social work practice, i.e., the interaction between a practitioner and service users, whether in the form of general meetings or distinct interventions designed for a target group or specific situation. These interventions can be divided into two types: those that child protection practitioners apply themselves and those that are provided by other professionals independently of or in collaboration with child protection. Several programmes have also been developed to reform both the system and frontline practice, i.e., all three levels. These include practice models, such as RSW.

3.2.1 Previous research on the effectiveness of child protection on the macro- and meso-levels

The majority of previous meso-level research has focused on children in out-of-home care. In their systematic review, Gypen et al. (2017) analysed the evidence of 32 primary quantitative studies and compared the outcomes in two different contexts, i.e., countries with a child protection orientation and those with a family services orientation. The study found that young people leaving care are at several disadvantages compared to peers in the general population. Former foster youth had adverse outcomes in their young adulthood in the areas of education, employment, income, housing, health, substance abuse and criminal involvement. This applied in both contexts, i.e., children coming from family service systems and children coming from child protection systems. A systematic review by Kääriälä and Hiilamo (2017) focused on out-of-home care outcomes in the Nordic countries. All 20 eligible studies were register-based cohort studies. The overall finding was that placement in care was associated with negative outcomes in young adulthood in all outcome categories, i.e., educational challenges, self-supporting problems, mental health problems, criminality, suicidal behavior, teenage parenthood, mortality, alcohol and substance use, and being on a disability pension (Kääriälä & Hiilamo, 2017).

In their subsequent studies, Kääriälä et al. (2018) examined the association between placement in care and completed secondary education across three Nordic countries. The birth cohort datasets of children born in 1987 were derived from national registers: Finland (n=58 855), Denmark (n=55 995) and Sweden (n=100 152). In all three countries, children placed in care had a significantly higher risk for early school leaving, and the risk was highest in Denmark. In Kääriälä et al. (2019), the association between individuals' care history and their early adulthood education and employment trajectories was analysed by comparing children in care with a matched peer group without care experience and the general population. The study used the longitudinal birth cohort data of children born in Finland in 1987 (n=59,476). The study found that children in care were less likely to enter trajectories characterised by education and employment and more likely to enter disadvantaged trajectories involving early parenthood, long periods of fragmented social assistance benefit receipt, and unemployment (Kääriälä et al., 2019).

A scoping review by McKenna et al. (2021) explored the associations between child protection services, open and out-of-home care, and long-term adult mental health outcomes in all geographical locations. The majority of the primary studies indicated that adults with an experience of any child protection service had higher rates of adverse adult mental health outcomes and a higher suicide risk compared to adults with no contact with child protection as children. The findings regarding the effect of placement duration, placement stability and age at entry were conflicting. The researchers argue that, based on their review, more nuanced research and detailed reporting is needed. For example, researchers should delineate outcomes by placement type, i.e., foster care with strangers, foster care with biological family members and residential care. In addition, only a few studies controlled for potential confounders, i.e., adverse childhood experiences, that are associated with poor mental health and simultaneously influence the nature of child protection services. Future research should attempt to separate the effects of the child protection services from the underlying reasons leading to the receipt of these services. Finally, McKenna et al. (2021) call for more research on the mental health outcomes associated with open care (i.e., services provided for children living at home with their parents).

An example of a study operating on the macro and meso levels is the attempt by Vinnerljung et al. (2006) to evaluate the effectiveness of the child protection system as a whole and the effectiveness of child protection assessment in Sweden. The study compared the prevalence of social problems in young adulthood between three groups: one that had received child protection interventions, one that had been referred to and investigated by child protection authorities but had never received any of the services, and general population peers. The representative sample (n=2 232) was drawn from the case files of children born in 1968-1975 and living in Stockholm from birth to age 18. Outcome data were collected from national registers, and the indicators involved teenage parenthood, criminal offences, hospitalisations for psychiatric diagnoses, and having received welfare by the age of 25. Compared with general population peers, negative outcomes were 2-3 times more common in the group that had been investigated but not received services, and 3-4 times more common in the group that had received child protection services. Vinnerljung et al. concluded that the investigation process had failed since some children had been left in adverse home conditions that increased the risk of future problems. In addition, the minor differences between those having received services and those who were only investigated could indicate that child protection services had weak long-term impacts. The researchers recommend an enhanced substantiation process and provision of primarily evidence-based interventions.

These studies provide important knowledge of the inability of systems to provide sufficient support to enable children and young people in need of statutory protection to achieve an equal position with their peers as young adults. However, to improve the operation of the system, more differentiated research on the micro level is needed.

3.2.2 Previous research on the effectiveness of child protection on the micro-level

Since social work is not a particularly method-oriented practice, few detailed methods for application in case work exist. One such is Family Group Conferencing (FGC). FGC is a participatory decision-making model that brings together the family, extended family and professionals in a family-led meeting to make a safeguarding plan (Frost et al., 2014). The original FGC model consisted of four stages: preparation, the giving of information, private family time, and making the plan and agreement. Various variations on this model have been developed since (Dijkstra et al., 2016). Underlying FGC are several, complementary theories, i.e., problem-solving, empowering, solution-focused and strengths-based approaches and increasingly also a restorative approach (Frost et al., 2014). The core idea is that families have the right to voice their views on matters concerning them and to be involved in important decisions about their children. It is expected that plans made on this basis are more likely to be carried out successfully compared to plans developed by professionals (Burford & Hudson, 2000).

The model was first developed in Aotearoa New Zealand to recognize the cultural values and traditions of the Māori community and address their ill treatment in matters of child protection (Frost et al., 2014). FGC has since been implemented or tried out in different forms in at least in 30 countries (Havnen and Christiansen, 2014), including Finland and other Nordic countries (Heino, 2009). Despite the model's apparent popularity among child protection organisations and researchers, no strong evidence exists on its effectiveness. In their review, Havnen and Christiansen (2014) found no clear evidence that FGC reduces maltreatment and abuse, out-of-home placements, or the need for further contact with child protection services. In their meta-analysis, Dijkstra et al. (2016) found no significant overall effects of FGC on child maltreatment, out-of-home placements, or on involvement in youth care. Of the large number of studies examining FGC, the majority have focused on implementation and participant satisfaction, and only 14 were controlled effectiveness studies. The quality of these studies was weak, e.g., the assignment of cases to the experimental and control group was retrospective and the data were often obtained from case files; in other words, no valid instruments had been used. No information on model fidelity, which would enable assessment of the degree to which fidelity explained the lack of positive effects, was available (Dijkstra et al., 2016).

Another topic for micro-level intervention is the elaboration of social workers' direct practice skills. In their research project, Forrester et al. (2019) identified seven key communication skills for child protection practice (Forrester et al., 2020) which they grouped into three dimensions: relationship building (including empathy, collaboration, autonomy), good authority (purposefulness, focus on child, clarity about concerns), and evocation of intrinsic motivation (i.e., support for behaviour change). By using a validated coding framework (Whittaker et al., 2016), researchers rated practitioners' skills on a five-point scale from recorded observations of service user meetings and measured child protection outcomes

by collecting data from families with validated instruments. The findings (Forrester et al., 2019) suggest that relationship-building skills predicted parent-reported engagement. However, good authority and evocation of motivation had a stronger association than relationship building with the outcome measures. In their subsequent analysis, Forrester et al. (2020) found that skilled social workers were able to combine good authority and empathic engagement, whereas workers experiencing trouble in the use of authority were also less skilled at engagement. Finally, the researchers found that all these skills can be improved by providing practitioners with training in Motivational Interviewing (Forrester et al., 2018) or by systemic supervision (Bostock et al., 2019a).

Most of the specific interventions that are also occasionally applied in the context of statutory child protection are first developed and evaluated in other contexts. For example, group-based parenting programmes are used to support parents seeking for help due to a child's challenging behaviour (Karjalainen, 2021). The content and methods of delivery of these programmes are based on well-researched theories such as cognitive-behavioural, social-learning and attachment theories. Strong evidence that these programmes have many beneficial outcomes for children and parents has been reported. Systematic reviews have found that parenting programmes can reduce risk factors of physical abuse and neglect (Barlow et al., 2006) and child maltreatment (Euser et al., 2015). Chen and Chan (2016) found that these programmes have also promoted protective factors such as positive parenting, parent-child interaction, and child-rearing attitudes. In Finland, one such programme, titled Incredible Years® (Webster-Stratton, 2010), was found effective in increasing positive parenting and reducing child behavioural problems in families currently receiving child protection services (Karjalainen et al., 2019). It was also well-received by parents (Karjalainen et al., 2020). However, the programme is delivered by trained group leaders who typically work in family counselling centres. In the research project led by Karjalainen, the programme was delivered in collaboration with child protection services: two of the leaders were from family counselling services and one (family worker) from child protection services. However, this type of delivery is, to date, an exception. Moreover, some child protection organisations were unable to identify families that they believed would benefit from the programme, despite the very likely high prevalence of these problems among their service users (Karjalainen, 2021).

3.2.3 The effectiveness of child protection practice models

Among the other practice models for child protection, RSW aims to improve micro-level practice by providing a comprehensive theoretical and practical approach to all stages of child protection. Practice models also aim at macro- and meso-level reforms, e.g., changes in child protection policies and organisations to ensure intended change on the micro level.

Barbee et al. (2011, p. 623) define a practice model as follows:

> "A practice model for casework management in child welfare should be theoretically and values based, as well as capable of being fully integrated into and supported by a child welfare system. The mode should clearly articulate and operationalize specific casework skills and practices that child welfare workers must perform through all stages and aspects of child welfare casework in order to optimize the safety, permanency and well-being of children who enter, move through and exit child welfare system."

A systematic review (Isokuortti et al., 2020), conducted parallel to this dissertation research, sought evidence on the effectiveness of all the practice models that met this definition. Local innovations that had not been disseminated to other agencies were excluded, as no evidence exists of their transferability and scalability. As a result, three eligible practice models were identified: RSW, Signs of Safety (SoS) and Solution-Based Casework (SBC).

SoS is a strengths-based, safety-focused, and participative approach that applies techniques from Solution Focused Brief therapy. The aim is to create a shared focus and understanding among families and practitioners, and to involve children and families in safety planning. SoS (Turnell & Edwards, 1999) was developed in Australia by Edwards, a child protection practitioner, and Turnell, a brief therapist, together with practitioners in a collaborative action learning process, and has continued evolving ever since. In its current form (Turnell & Murphy, 2017), the model consists of three core principles: constructive working relationships between professionals, family members, and professionals themselves; thinking critically with a questioning approach; and respecting service users' and practitioners' views on good practice. While SoS along with publicly available materials is free of charge, a licensing programme for trainers has been launched to ensure the quality of implementation of the model (https://www.elia.ngo/licensing-1). SoS involves specific registered tools, e.g., Mapping, Safety Planning, and Three Houses, and child protection organisations have also adopted these without any training in SoS. It is estimated that SoS has been implemented in some form in nearly 200 agencies in 15 countries, including Australia, Aotearoa New Zealand, Japan, Canada, United States, Cambodia, and several countries in Europe (Turnell & Murphy, 2017). In the United Kingdom, approximately one hundred local authorities use SoS: half of them use it exclusively, and half use it alongside or incorporated into some other framework (Baginsky et al., 2021).

SBC is based on three theoretical foundations: family life cycle theory, relapse prevention, and solution-focused family therapy. The aim is to attain full partnership with the family. By focusing on the patterns of everyday life of the family, practitioners help service users to identify situations that may result in behaviour associated with child maltreatment and to develop skills to avoid and interrupt such situations (Christensen & Todahl, 1999). To improve child safety, service users and practitioners co-develop action plans at both the family and individual levels (Antle et al, 2012). SBC has been developed in the US by Christensen and his colleagues together with child welfare practitioners (Christensen et al., 1999). The model is registered, and the training is provided by approved

SBC trainers. SBC has been implemented in six states in the United States and in one state in Australia (Gillingham, 2018).

RSW is a systemic approach to child protection rooted in systemic family therapy (for a more detailed description, see Section 2.3.1). RSW is not registered, but the model developers have created specific training programmes and consultation for its dissemination. The precise number of organisations using RSW is not available. Based on published studies on RSW, it has been implemented in seven local authorities in United Kingdom (Bostock et al., 2017; Forrester et al., 2013; Laird et al., 2018; Wilkinson et al., 2014).

Several studies have evaluated the implementation of these models from various perspectives. In contrast, our review (Isokuortti et al., 2020) identified only five papers, representing six studies, that have evaluated their effectiveness. To be included in the sample, studies had to be quantitative, minimally controlled, before-after studies, i.e., randomized controlled trials or a quasi-experimental studies. Four of the eligible studies addressed the effectiveness of SBC, one of SoS and one of RSW.

A quasi-experimental study by Antle et al. (2012) explored the outcomes of SBC with a sample of 4 559 child protection cases from one federal state that had been randomly selected during the period 2004-2008. This sample was subcategorised into two groups based on their fidelity score, which was measured with a tool covering 33 core elements of the SBC model. The results indicated that outcomes were better in cases where SBC fidelity was high than in cases where it was low. Outcomes were measured by using federal definitions of safety, permanency, and well-being, but no information on the validity of these measures was provided. Another quasi-experimental study by Antle et al. (2009) explored recidivism by comparing two groups: a high-fidelity SBC group and a comparison group not implementing the model (n=760). Data were derived from state data reports. Based on the findings, the SBC group had fewer recidivism referrals, i.e., another incident of substantiated or indicated abuse or neglect within a six-month period, than the comparison group. The first study reported in Antle et al. (2008) was a quasi-experimental study comparing the cases (n=48) of teams with a high and lower degree of training in SBC. The primary outcome measure was the number of case goals and objectives, collaboratively defined by the practitioner and family, achieved by the family. Based on the analysis, the mean number of goals achieved was higher in the group that had received more SBC training (Antle et al., 2008). The second study reported in Antle et al. (2008) used another sample (n=100). SBC fidelity was first assessed for each case, and the cases were then assigned to two study groups: a high-fidelity group (SBC group) and a comparison group. Goal achievement was measured as a dichotomous variable (Yes/No). Based on the findings, the families in the SBC group were more likely to achieve both family-level and individual-level objectives (Antle et al., 2008).

In the case of SoS, Reekers et al. (2018) explored its effectiveness in a quasi-experimental study comparing a group of families supported by SoS to a group of families in service as usual. The sample consisted of 40 families, and the group-

ings were based on propensity score matching. The study used standardised instruments to measure the outcomes, i.e., child maltreatment according to the social worker and to the parents, family empowerment, and service system empowerment. The study found that both the SoS and control group were equally effective in reducing the risk of child maltreatment, and in increasing family empowerment and service system empowerment (Reekers et al., 2018).

Finally, Bostock et al. (2017) explored the effectiveness of RSW in a quasi-experimental study primarily focused on the practice and experiences of RSW implementation in five local authorities. Two groups of families were compared: service users in units receiving additional training in RSW and a control group that had received less training. The study found no difference between the study groups in the number of children entering care. Comparative analysis on the other outcome measures could not be conducted due to high attrition in the follow-up data (Bostock et al., 2017).

A systematic review (Isokuortti et al., 2020) concluded that the number of controlled effectiveness studies was low, and that the quality of these studies was weak. Identified methodological problems were the risk of selection bias, small sample sizes, short-term follow-up, and reliance on single-source data. However, lack of high-quality evidence does not prove that these models are not viable. The findings highlight the importance of proceeding cautiously in local implementation and collaborating with researchers attempting to increase the evidence-base of these models (Isokuortti et al., 2020).

The searches for the review were conducted in spring 2019. Since then, two research reports on practice models have been published. In their study on the SoS model, Baginsky et al. (2020) compared the outcomes for children in 9 child protection organisations applying SoS with outcomes for children in similar organisations not applying the model. Based on their difference-in-differences analysis, no strong evidence that SoS would improve outcomes, e.g., the duration of assessments or the probability of re-referrals, was found. In their follow-up study on the RSW model, Bostock & Newlands (2020) found that in the one local organisation that had retained all the key elements of the RSW model, the reduction in overall risk (i.e., the number and percentage of children receiving any children's services) to children was greater compared to organisations only partially applying the model. The authors conclude that a sustained commitment to systemic practice may be associated with improved outcomes for children and families. However, these findings need to be interpreted with caution. The report has not been peer-reviewed, and the performance indicators used as outcome measures in the study are ambiguous, i.e., they do not necessarily reflect similar change in well-being across population.

4 THEORETICAL AND METHODOLOGICAL FRAMEWORK

4.1 Evidence-based practice and the quality of the evidence

The idea of evidence-based practice (EBP) was originally articulated in the context of health care (Sackett et al., 1996; Sackett, 2000) and only later adopted in other fields of human services, such as crime prevention, education, and social work. Although empirical research guiding practice has been conducted in all these disciplines, practitioners tend to make decisions based on intuition, expert or peer opinion, tradition, or anecdotal experience instead of the most updated research findings (Gambrill, 2006; Morago, 2006). The innovation in EBP was to guide practitioners to use scientific knowledge as part of their everyday work. According to Sackett et al. (1996),

> "Good doctors use both individual clinical expertise and the best available external evidence, and neither alone is enough. Without clinical expertise, practice risks becoming tyrannised by evidence, for even excellent external evidence may be inapplicable to or inappropriate for an individual patient. Without current best evidence, practice risks becoming rapidly out of date, to the detriment of patients."

Hence, in the context of healthcare, evidence-based medicine was later defined as "the integration of best research evidence with clinical expertise and patient values" (Sackett, 2000).

The concept of EBP expanded into the discipline of social work in the 1990s where it met with a mixed response (Morago, 2006). Social work scholars have debated the features and misrepresentations of the original ideas of EBP (Gambrill, 2006; Parrish, 2018) and whether or not this original proposition needs to be reformulated for social work purposes (Gitterman & Knight 2018; Thyer, 2018). While EBP could improve the quality of practice, and improve transparency, and accountability (Forrester, 2010) its implementation is hindered by several obstacles, such as limited material and financial resources, job pressures, and the lack of knowledge and skills of practitioners to critically appraise research findings or

deliver effective interventions (Bellamy et al., 2006; Forrester, 2010; Morago, 2006). While acknowledging the vast literature on the applicability of EBP in the social work context (e.g., Gray et al., 2013; Julien-Chinn & Lietz, 2019; Regehr et al., 2007; Tilbury, 2021), I focus here on its theoretical and methodological aspects.

The first proposition on the evidence hierarchy in studies on social programmes was formulated by Campbell and Stanley (1963). In their book, they describe a three-tier hierarchy and explained how different research designs impact the internal and external validity of research findings. In true experiments, participants are randomly assigned to the treatment group or the control group, and some outcome of interest is measured after (post-test) or before and after (pre-test post-test) the intervention. These are nowadays called randomised controlled trials (RCT). When the researcher has no control over the execution of the programme (when and who are exposed to a programme), a quasi- experimental design can be applied. Typically, this would involve two naturally assembled collectives, one of which is exposed to a programme while the other is not, and the use of pre-test post-test measurements. Even if the groups are not matched or equivalent, the use of a control group reduces the ambiguity of the research findings (Campbell & Stanley, 1963). The least preferable study designs for evaluating a social programme are what the authors call 'pre-experimental' studies. These involve designs in which only a single group is studied, or a comparison group is used, but no pre-tests or randomisation are done. Pre-experimental study designs present the most threats to validity whereas true experiments present the least (Campbell & Stanley, 1963).

Since Campbell and Stanley, various new propositions on evidence hierarchies and guidance on rating evidence have been developed (Blunt, 2015; Evans, 2003). Sackett (1981; originally published anonymously) suggested that the strength of the evidence in different study designs could be rated from strongest to weakest as follows: RCTs, cohort studies, case-control studies, case series. As systematic reviews have become more popular, these are now located at the top of the hierarchy (Elamin & Montori, 2012; Greenhalg, 2000). Some hierarchies include expert opinion, which is placed on the lowest step of the hierarchy (Blunt, 2015). However, evidence quality cannot be assessed solely by reference to the study design. Hence, more nuanced grading recommendations have been developed. For example, the Grading of Recommendation Assessment, Development and Evaluation (GRADE) (Guyatt et al., 2008) rates the quality of the evidence for each outcome by assessing various aspects across studies, including study limitations, imprecision, inconsistency of results, and indirectness of evidence. As the volume of research on the gathering and synthesising of evidence has increased, so too have hierarchies for rating the overall evidence base for an intervention based on the consistency of the effect across research studies. Such ratings are applied for example by Evidence Stores that gather and analyse evidence to provide knowledge for decision-making (see for example What Works for Children's Social Care https://whatworks-csc.org.uk/evidence-store/).

The evidence hierarchies focusing on effectiveness of interventions have been challenged in many respects. For example, Evans (2003) suggested a hierarchy that assesses the evidence not only on the effectiveness but also appropriateness of the intervention for its recipient, as well as feasibility, i.e., issues related to its implementation, cost, and change in practices. All these dimensions can impact the success of an intervention. An intervention that might in theory seem to be highly effective will not produce the desired outcomes if it cannot be properly implemented, or if it is not accepted by its users. To assess the evidence for the appropriateness and feasibility of an intervention requires research methods different from those used and valued in assessing its effectiveness. Flay et al. (2005) point out that even when it has been proven to be effective, a programme should be ready for dissemination before it can be successfully implemented in new settings. When screening and evaluating the scientific evidence, the focus should not be exclusively on the evidence of the outcomes but also on the evidence of the programme's readiness for broad dissemination, i.e., scaling up, adoption, implementation, and sustainability. These viewpoints have also guided this dissertation in which the intention is not only to evaluate the outcomes but also the features of implementing the SPM that might explain its success.

Given that health and social care practitioners work in complex situations, Greenhalg et al. (2018) suggest that producing a relevant and meaningful synthesis of the research evidence for practitioners requires both conventional systematic reviews and more interpretative and discursive narrative reviews. These different kinds of reviews should be viewed as complementary rather than as in competition (Greenhalg et al., 2018).

Although RCTs are regarded as a 'golden standard' in primary research, Deaton and Cartwright (2018) point to their limited value as a guide to decision-making. They argue that while RCTs can be used to show that an intervention works in the sample used in a specific study, the result cannot automatically be transferred to another context and population. Because a programme has worked in one place does not mean that it will work somewhere else (Cartwright & Hardie, 2012). Clarke et al. (2014) argue that the evidence on mechanisms is just as important as the correlational evidence (i.e., probabilistic dependence between arbitrary variables) provided by RCTs and other statistical trials. Evidence on mechanisms is produced in studies that use mechanism-based reasoning in trying to establish the mechanism underlying the correlation found in statistical trials. However, mechanism-based reasoning is generally located at the bottom end of an evidence hierarchy. Clarke et al. (2013) continue that without understanding the mechanisms producing the outcomes, it is not even possible to adequately design and interpret RCTs. Both types of evidence are needed to establish causal claims, transport such claims to new populations, and apply them to individual cases. Hence, the evidence on mechanisms needs to be viewed as complementary to, rather than inferior to, the evidence on correlations (Clarke et al., 2013).

To better understand what works or would work, Pawson (2006) suggested that systematic reviews should apply a realist understanding of causality. For the

purposes of conducting such reviews, Pawson formulated a new method, i.e., the realist synthesis. Pawson and Tilley (1997) had already formulated an evaluation method for conducting realist primary studies, i.e., the realist evaluation. Both methods seek to provide more in-depth understanding on whether an intervention works, and if so for whom, in what context, and why. In realist research, the aim is to explore outcome patterns instead of outcome regularities, generative mechanisms instead of successionist causality, and contextual conditions explaining whether these mechanisms can operate as intended.

I will now turn to the philosophical underpinnings of my thesis, i.e., critical realism, and introduce the main features of the realist evaluation that has informed my project.

4.2 Critical realism as the ontological and epistemological background of the study

Critical realism is a strand of scientific realism. The two defining features of scientific realism are that the world is believed to exist independently of our knowledge of it, i.e., the world and its objects are mind-independent, and that it is possible to acquire fallible theoretical knowledge about it (Kaidesoja, 2005). Critical realism makes more specific claims on the ontological structure of the world.

For Roy Bhaskar (1997 [1975]), one of the founding fathers of critical realism, the world is stratified. He makes a distinction between three domains of reality: *the empirical*, *the actual* and *the real*. The domain of the real refers to "whatever exists" (Sayer, 2000, p. 11). Objects in this domain exist irrespective of their being empirical objects for science, or whether their nature is understood. These objects have certain structures and causal power. The domain of the actual refers to 'events', i.e., what happens when these powers are activated, and the domain of the empirical is the level on which we can make experiences and observations (Sayer, 2000). However, some structures may not be observable. The existence of unobservable entities can be argued for by reference to observable effects which can only be explained as the products of such entities. In addition, realists acknowledge that powers may exist unexercised. Real objects have potential that could, in a different time, result in different outcomes than those that have happened or have been observed (Sayer, 2000).

According to Bhaskar (1997 [1975]), knowledge has two sides; 'transitive' and 'intransitive.' Transitive knowledge refers to the theories of the world. These are "the artificial objects fashioned into items of knowledge by the science of the day" (Bhaskar, 1997 [1975], p.21). By intransitive objects, Bhaskar refers to objects that do not depend upon human activity. According to Bhaskar (p. 22), these are "the real things and structures, mechanisms and processes, events and possibilities of the world." Intransitive objects of knowledge are in general "invariant to our knowledge of them" (Bhaskar, 1997 [1975], p.22). As Sayer (2000) formulates

the argument, even if our theories of the world change, the world itself does not change: "there is no reason to believe that the shift from a flat earth theory to a round earth theory was accompanied by a change in the shape of the earth itself" (p. 11). Similarly, social phenomena are unlikely to significantly change when scientists change their theory of them, although it is admitted that the social world is socially constructed. However, scientists do not 'construct' the social world, rather, they construe it (Sayer, 2000).

Critical realism is committed to epistemological relativism and judgmental rationalism (Bhaskar 1986). The former acknowledges that various interpretations of reality exist. However, while all knowledge is socially produced, it can be potentially fallible (Patomäki & Wight, 2000). Judgmental rationalism means that some theories and interpretations are more accurate than others. It is possible to prefer one theory over another based on better justified arguments (Patomäki & Wight, 2000).

For realists, the world is emergent (Sayer, 2000). Both social phenomena and structures are combinations of internally related elements and their causal powers, and they are irreducible to these individual elements. The activation of causal powers depends on other conditions. Consistent regularities are likely to occur only in special conditions, i.e., in closed systems. The social world is an 'open system', where a causal power can produce different outcomes depending on the prevailing conditions. Hence, according to critical realism, causation cannot be proved in a 'successionist' way by gathering data on regularities and repeated occurrences. Rather, causal explanation entails identifying causal mechanisms and how they work and discovering if they have been activated and under what conditions (Sayer, 2000). Mechanisms are generative, they are the "ways of acting of things" (Bhaskar, 1975, p.14) or "'powers' inherent in a system" (Pawson, 2006, p.23).

Margaret Archer in her *morphogenetic approach* (Archer, 1995) has sought to complement Bhaskar's philosophical realism to provide a more "useful and usable" social theory for social research (p. 135). In a critique of previous 'individualistic' or 'holistic' social theories, Archer suggests an approach for analysing the relationship between 'structure' and 'agency', concepts which she sees as fundamentally inseparable. According to her, social reality is unlike any other kind of reality because of its human constitution. First, the very existence of society depends in some way upon human activities; hence, society is inseparable from its human components. Second, it is characteristically transformable. What precisely society is depends upon human doings and their consequences. Third, humans as social agents are affected by the society in which they live and by their efforts to transform it. Both structure and agency possess distinctive emergent properties. Human beings are simultaneously free and constrained. Morphogenesis refers to "those processes which tend to elaborate or change a system's given form, state or structure" (Archer, 1995, p.166). It means that society takes its shape from, and is formed by agents, originating from the intended and unintended consequences of their activities. This must be acknowledged in endeavors to explain what social reality is.

As suggested by Archer (1995, p.2), "what society is held to be also affects how it is studied." This study is committed to ontological and epistemological ideas suggested within the school of critical realism, and their application in the form of realist evaluation methods.

4.3 Realist evaluation

Realist evaluation is a theory-driven form of evaluation based on a realist philosophy of science. It shares the main arguments of critical realism proposed by Bhaskar and Archer, e.g., the idea of underlying generative mechanisms and the understanding that people both create and are created by society, but it takes a critical stance (at least in the writings of Pawson) towards Bhaskar's understanding of the concepts of experiment i.e., it requires a closed system, and 'depth realism', i.e., stratified reality (Pawson, 2013). The principles of realist evaluation were first formulated by Pawson and Tilley in their seminal work "Realistic Evaluation" in 1997. In 2006, Pawson articulated the principles of realist synthesis, i.e., a realist approach to the reviewing of scientific literature and the synthesis of existing evidence. The methodological development and application of both methods has expanded, leading to more detailed guidance (Dalkin et al., 2015; Emmel et al., 2018; Manzano, 2016) and reporting standards (Wong et al., 2013, 2016). The overall aim of realist inquiry is to understand what works, for whom, under what circumstances, and how.

Realist evaluation aims at formulating and refining theories explaining why a programme does or does not work. First, the change is seen as a result of underlying mechanisms (Pawson & Tilley, 1997). As these are not always observable, they must be identified by theorising and testing theory-based hypotheses with empirical data. As opposed to data-driven research, a realist evaluation is theory-driven: the theory drives first the data gathering and then the analysis. By accumulating knowledge about an experiment makes it possible to sort out which theories of the underlying mechanisms have proven to be more robust than others (Pawson, 2013). Second, the mechanisms a programme seeks to trigger do not always fire (Pawson & Tilley, 1997). The context may or may not activate the mechanisms, and this needs to be addressed through a realist inquiry.

Realist evaluation is method neutral, i.e., it does not require any specific methods to be used. All methods and data that provide insight and help the researcher to elaborate theories about what works, for whom, under what circumstances, and why, can be used (Emmel et al., 2018). A notable exception to this principle is that the possibility of realist RCTs, as suggested by Bonell et al. (2012), has been rejected. Marchal et al. (2013) argue that the RCT design is "fundamentally built upon a positivist ontological and epistemological position" and cannot be meaningfully adapted within the realist paradigm. In their response, Bonell et al. (2013) point out that "research methods don't make assumptions, researchers do." According to them, RCTs can be conducted without commitment to ideas that contradict realist philosophy, e.g., that reality is governed by stable laws.

Tilley (2016), in turn, argues that RCTs have their place among many other methods needed to understand how an intervention works: 'anything goes' as long as it contributes to progress. I agree with these authors and believe that all pieces of evidence may be beneficial in testing and refining theories. After all, Pawson (2013) himself has argued that evidence should be sought in all its forms, "quantitative and qualitative, outcome and process, measurement and gossip" (p.11).

The key features of a realist evaluation have been stipulated by Wong et al. (2016). First, a realist evaluation approaches social programmes, or interventions, as attempts to create change by offering resources to participants to (or withdrawing them) or by changing the context in which decisions are made. Programmes enable or motivate participants to make different choices; this entails change in their reasoning and in the resources available to them. However, the context in which a programme operates affects its mechanisms by enabling them to operate or preventing them from operating, in turn affecting their outcomes. The interaction between the programme's context and its mechanisms produces the programme's outcomes. This process is formulated as follows (Pawson & Tilley, 1997):

Context + Mechanism = Outcome

The researcher's hypotheses about a programme's context, mechanisms and expected outcomes are formulated in *context-mechanism-outcome configurations* (*CMO configurations)*.

The starting point and the unit of analysis in a realist inquiry is a programme theory (Pawson, 2013). It works as a set of hypotheses that are to be tested by collecting relevant qualitative and quantitative data. A programme theory consists of "the ideas about what makes an intervention work" (p. 87). It is assumed that different stakeholders have different perspectives, information, and understanding of the working of the programme. The researcher's task is to collect the partial knowledge of the different stakeholders (i.e., their ideas, hopes, beliefs and knowledge of the programme), take into consideration the unacknowledged conditions and unintended consequences of actions, and then formulate a wider understanding of the intervention, which is then tested with empirical data (Pawson & Tilley, 1997, pp. 159-164).

Ideally, programme theories would be refined and re-tested in a continuous process. For Pawson (2013, p.11) the "hierarchy of evidence" derives from this process. Theories must be tested again and again, and it is the quality of the reasoning rather than the quality of the data that we should be interested in.

In the recent methodological literature, several contributions have been made concerning two key concepts, i.e., mechanism and context. Pawson (2013) suggests that context has four different levels:

1. The individual capacities of the stakeholders

2. The interpersonal relationships of the stakeholders that carry the programme

3. The institutional setting, such as the rules, norms, and customs local to the programme

4. The wider infrastructure, i.e., the social, economic and cultural setting of the programme.

Researchers have, however, struggled to differentiate the contextual from the mechanistic aspects of a programme. Building on the original work of Pawson and Tilley (1997), Dalkin et al. (2015) proposed a revised formula in which mechanisms are disaggregated into resources and reasoning in the following temporal order:

> Mechanism (Resources) + Context -> Mechanism (Reasoning) = Outcome.

In other words, the programme's resources are introduced into a context. The effect of doing this leads to a change in the participants' reasoning. This new reasoning then leads to change in the participants' behaviour, and hence to other outcomes.

Another debate among realist researchers has focused on the level of analysis required to identify mechanisms. Several researchers have argued that mechanisms may also be located in levels other than individual reasoning, such as within institutional structures, culture, agency, and the relational properties between them (Shaw et al. 2018; Westhorp 2018). Indeed, Dalkin et al. (2015) suggest that Pawson and Tilley have adopted a different understanding of what a mechanism is than that found in realist theory. However, Dalkin et al. do not consider this a problem. They see that a researcher may choose which concept of a mechanism she will use according to what type of change the programme is seeking. If the programme aims at a large-scale social transformation, then the researcher will look for structural mechanisms. If the programme aims at behavioral change, then the focus is on individual reasoning.

The difference between context and mechanisms becomes more problematic when it is acknowledged that mechanisms can be situated in many other levels than exclusively inside a participant's mind. Complex programmes that aim at changes on both the macro and micro level have long causal chains before the target outcomes are realised on the service user level. As illustrated by Shaw et al. (2018), improving multi-professional collaboration in social and health care can change practices with clients, and eventually the health or wellbeing of a service user. The outcomes of the programme activities targeted at the behavior of professionals will later serve as the context for client practice. Thus, Shaw et al. (2018) came to the conclusion that

> "the distinction between context and mechanism might be ultimately arbitrary, such that mechanisms occurring more distally in a causal chain that leads to a particular outcome become contexts for *subsequent* mechanisms occurring more proximally to the outcome of interest. In this way, the causes of outcomes of complex health interventions, policy programmes or any social action may be viewed either as contexts *or* mechanisms depending on where the analytic focus is located in a given causal chain." (Shaw et al. 2018, emphasis in original)

In this dissertation research, I have used realist evaluation as the methodological framework guiding the formulation of my research questions, data collection, and analysis. For me, the key components of a realist evaluation are the realist understanding that the underlying mechanisms may or may not be activated depending on the context; the attempt to analyse the nature and working of these mechanisms to understand the observed outcomes; and the idea of formulating, testing and refining programme theories in a cyclical process. To undertake a realist evaluation, various types of data and analytical tools are needed. Each additional study accumulates knowledge and increases our understanding of social programmes.

4.4 Applications of a realist evaluation in social work research

While critical realism has been advocated by several social work scholars (Kjørstad & Solem, 2017), empirical examples of realist evaluations of social work interventions are scarce (Boost et al., 2021; Caffrey & Browne, 2022). At the same time, realist understanding of the working of an intervention in a given context and of realist methodological tools occupies a well-established position in health care research, i.e., they have been included in the widely used UK Medical Research Council's guidance for evaluating complex interventions (Skivington et al., 2021).

Some of the first applications of realist evaluation in social work research were conducted by Kazi. In his book, Kazi (2003) presents examples of the realist evaluation projects he has conducted in various settings, i.e., the formative evaluation of social inclusion models with the drug-using community, single-case evaluations conducted in family centres, and the formulation of Mechanism-Context-Outcome Configurations based on quantitative analysis explaining the causal patterns of services for children and young people who have sexually harmed others. More recently, realist inquiry methods have been applied in research on child protection practice models. Forrester et al. (2013) applied a realist evaluation in their study on RSW. Later, two realist syntheses on Signs of Safety were conducted to develop (Sheehan et al., 2018) and refine (Caffrey & Browne, 2022) a programme theory for this practice model. Other examples of realist evaluation in other types of social work interventions include the study by Boost et al. (2020) aimed at understanding social work practices that focus on improving access to social services and guaranteeing social rights for everyone through interorganizational collaboration.

Blom and Morén (2010) outlined their own realist method to explain social work practice, i.e., the CAIMeR theory. The authors felt the realist evaluation proposed by Pawson and Tilley was inadequate for social work research since it lacked explicit actors and interventions; hence, CAIMeR guides the evaluator to focus on the context, actors, intervention, mechanisms, and results. While this methodological approach shares the philosophical underpinnings of realist eval-

uation, it proposes a more detailed scheme of concepts and sub-concepts for analysing social work interventions. However, there seems to be no empirical applications of this theory in the international literature.

In more general terms, critical realism has been used as a response to constructivist approaches to the analysis of social phenomena and the situation of service users in social work. For example, Houston (2001) suggested that critical realism offers a way of achieving a synthesis between objectivist theories seeking to predict and manage risk in child protection and subjectivist accounts approaching risk as socially constructed phenomenon.

In Finland, the first application of a realist evaluation in the social work context was probably that conducted by Rostila (2001) on a social work programme for the unemployed. Karjalainen (2012) applied a realist evaluation framework to a social work programme designed to empower unemployed young people. In the context of child and family social work, Svenlin (2020) applied the CAIMeR theory in her PhD thesis on the Support Family intervention by exploring informal practice theories guiding this intervention and the critical points and mechanisms explaining support relationships.

Rather than applying the method of realist evaluation, Finnish social work researchers have applied the ideas and theories introduced in the school of scientific or critical realism. Probably the first example of these was Mäntysaari (1991), who combined a critical realist understanding of social structures in his phenomenological analysis on the reasoning of organisations and practitioners regarding control and bureaucracy in social work offices. Pekkarinen (2010) applied Bhaskar's (1986) transformative model of social activity in a historical child welfare study analysing boys who had repeatedly committed crimes. Tapola-Haapala (2011) applied Archer's morphogenetic approach in a study analysing social work professionals as reflexive agents. Lunabba (2013) explored the recognition of problems and support needs of schoolboys, and social bonds between adults and pupils by applying both Bhaskar's and Archer's work on critical realism.

5 DATA AND METHODS

5.1 Study design

The overall aim of this dissertation research was to understand the effectiveness of the SPM and related factors. This has involved analysing programme outcomes (A1, A2), exploring the context affecting these outcomes, i.e., implementation fidelity and the factors affecting it (A3), and developing a coherent description of the components of SPM and a programme theory to be used in future implementation and evaluation (A4). The study applied realist evaluation as the overarching approach guiding the choices in the data collection and analysis' along with mixed-methods, including both qualitative and quantitative data - as is usual in realist evaluation (Greenhalg et al., 2017). The thesis is based on four sub-studies involving three distinct study designs and data (see Table 3).

Two sub-studies (A1, A2), comparing the family-level outcomes in two study groups with repeated measures data, were quasi-experimental. The third sub-study assessing the fidelity of implementation and the factors influencing this (A3) applied a concurrent transformative mixed-methods design. This entailed simultaneous quantitative and qualitative data collection, and mutual interpretation of the findings. The fourth sub-study (A4) was a qualitative, research-led process that aimed specifically at formulating a programme theory for the SPM by generating discussions in workshops and interviewing experts.

TABLE 3 Individual sub-studies, study designs and datasets.

	Sub-study	Research question	Study design	Data
A1	Aaltio, E. (accepted). Evaluating the effectiveness of the Systemic Practice Model of children's social care – A pilot study on child- and family-level outcomes	Did it work?	Quasi-experimental, repeated measures design	Effectiveness data Children's questionnaire Parents' questionnaire Social workers' questionnaire
A2	Aaltio, E. & Kannasoja, S. (accepted) Achieving agreement on service needs in child protection? Comparing children's, mothers', and practitioners' views over time and between approaches			
A3	Isokuortti, N. & Aaltio, E. (2020) Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland	In what context and for whom did it work or not work?	Concurrent transformative mixed-methods design	Context data Administrative data Social worker surveys Social worker interviews
A4	Aaltio, E. & Isokuortti, N. (2021) Developing a programme theory for the Systemic Practice Model in children's social care: key informants' perspectives.	Why should the SPM be effective? What are its core components and mechanisms?	Research-led programme theory formulation process	Programme theory data Workshop data Complementary data

5.2 Study conditions

A national evaluation of the dissemination of the SPM and local implementation was conducted by THL in parallel with the national dissemination project. I was the researcher responsible for conducting this evaluation from March 2017 until February 2019. Together with Nanne Isokuortti, I conducted the initial evaluation of the SPM (Aaltio & Isokuortti, 2019). Thereafter, we each continued investigating the SPM in parallel PhD projects.

I collected the data for the initial evaluation and for this dissertation in 2018 in THL. As a researcher, I did not manipulate the implementation conditions. The decisions of the sites and teams participating in the implementation were made on the local level, with municipalities and areas making an area-level plan on

how they would participate in the LAPE programme. The Ministry of Social Affairs and Health set the guidelines for participation and allocated funding. THL was responsible for the national level execution of the LAPE programme in general and for the SPM specifically.

While the local teams implementing the SPM covered all the areas of children's social care, the majority focused on child protection cases (see table 4).

TABLE 4 Areas of responsibility of the local teams participating in the implementation of SPM by summer 2018 (Aaltio & Isokuortti 2019, 69).

	Areas of responsibility			
Number of teams	Assessment	Child protection	Looked after children	Aftercare
22		x		
5		x	x	
4		x	x	x
3	x			
2	x	x		
1	x	x	x	
1			x	x
1		x		x

Note. Information on six teams is missing.

Although dissemination of the SPM was mainly based on the Training of Trainers and team-level training programmes, it was announced that the model would systemically change the whole organisation, from frontline practice to management and decision-making. In the first stage of dissemination, no specific training or coaching was provided for local leaders to bring about systemic change in their organisations. At some sites, a manager or a leader might participate in team training. To support the managers responsible for local implementation, THL organised three national workshops for managers (May 2017, September 2017 and February 2018). For the local level, these workshops provided information, guidance, and a forum for collective reflection on how to solve practical issues. For THL, the workshops provided information on local plans and the progress of the implementation.

The second stage of national dissemination began in January 2019 as part of the central government's decision to continue with the LAPE programme. The evaluation report (Aaltio & Isokuortti, 2019) on the first stage of SPM dissemination was published in the beginning of 2019, although the preliminary findings had already been presented in THL and in executive meetings on several occasions in 2018. One of the key findings was the lack of a concrete and clear description of the SPM, a factor which undermined its implementation. Together with the THL implementation team, the researchers decided to launch a workshop process aiming to formulate a coherent description and a programme theory for the SPM. This process was led by the researchers but executed in close cooperation with THL.

5.3 Effectiveness data

These data was used in two articles (A1 and A2) exploring SPM outcomes with a quasi-experimental repeated measures design (Research Question 1).

5.3.1 Procedures

To compare systemic practice with SAU, two samples of families were drawn, one supported by social workers in SPM teams and the other supported by social workers in SAU teams. Data were collected from nine SPM teams and nine SAU teams at three municipal children's services sites.

A convenience sample was recruited. First, the sites were selected from all the sites that had signed up to participate in the national implementation of the SPM by autumn 2017. The selection criteria were: 1) the organisation had several children's services teams, some of which had been assigned for SPM training (SPM teams), leaving the others to continue providing services as usual (SAU teams); and 2) the organisation would be able to implement the SPM in March 2018 at the latest, enabling baseline and follow-up data to be collected by the end of 2018. Three suitable sites were identified, one in the metropolitan area and two in medium-sized regional centers in Central and Eastern Finland. All three sites were willing to participate in the study. The SAU teams did not receive any particular training and did not implement any specific model to guide their practice during the study.

Second, cases were allocated to the study groups in real-world conditions, i.e., following the common procedures in each participating organisation. To reduce selection bias, the practitioners in teams participating the study were first asked to fill in research forms, preferably on all cases they were actively working with during the first data collection period. Due to high caseloads and difficulties in the implementation process, this plan had to be abandoned. Moreover, teams reported that they had only selected some families for systemic practice (see A3 for more details). Finally, a target number of cases to be included in the study was set in consultation with the practitioners in each team based on their workload and time resources. This process yielded a total of 312 cases. Both new and older cases were included in the study as the numbers of new cases per team per month were small (as low as one or two).

The first three-month data collection period (T1) took place at each site in Spring 2018 after the SPM team had been trained in systemic practice and the weekly team meetings with a family therapist had started. The SAU teams were instructed to start their data collection as soon as the SPM teams at the same site had begun practising. The researcher gave the practitioners both oral and written instructions. The second stage of the data collection (T2) started at each site approximately six months after the end of first stage.

5.3.2 Participant characteristics

The unit of analysis was a case involving a child and his or her parents. In the total sample, half of these children were girls and half were boys. The children ranged in age from one to 17, with a mean age of 13. The majority (93%) of children were Finnish speaking. In the majority of cases (64%), the parents were separated. One in four children were living with both of their birth parents. At baseline, a small proportion of cases (7%) were in out-of-home care. The only significant difference between the study groups was that the children in the SAU group (M = 13.5) were on average older than children in SPM group (M = 11.0) (U = 2023, p = .004). The mean age of the participating mothers was 41, and the majority (91%) of the mothers were Finnish speaking.

5.3.3 Instruments

The first draft of the questionnaires were designed based on the initial SPM logic model (see section 2.3.3). Next, to specify the intended outcomes, the researcher arranged a reflective discussion with the national evaluation steering group. Since no detailed consensus exists on the desired outcomes of child protection, deciding on the outcome measures was not straightforward. According to the Finnish Child Welfare Act (417/2007), child protection services should promote the child's development and wellbeing, provide support for parents in the child's upbringing and care, prevent child and family problems, and intervene sufficiently early if problems are observed. In assessment and decision-making, the best interests of the child must be taken into account, e.g., the child's development, wellbeing, close relationships, and safety. Based on these premises, it was concluded that key outcomes would include improving the child's well-being and family dynamics, and decreasing any forms of child abuse or neglect. In general, the overall goal would be to decrease the need for child protection from the practitioner's perspective, or outsider help from the service user's perspective. To identify suitable outcome measures, previous research on RSW was also explored. However, many of the measures had not been translated into Finnish and could not be used within the time scale and resources of the study. Finally, in consultation with the steering group, two standardised outcome measures already translated into Finnish were selected for this study. Additional outcome measures were also formulated and included

Data were collected from three groups of informants (practitioners, children, and parents) with questionnaires on wellbeing and the need for child protection and on the number of meetings and service-user feedback. Two versions of the children's questionnaire were in use, one for 7- to 12-year-olds and one for 13- to 17-year-olds.

5.3.4 Measures

Child subjective well-being was measured by KINDL-R, a generic instrument for assessing children's and adolescents' subjective experience of their physical, psychological, and psychosocial quality of life (Ravens-Sieberer & Bullinger, 1998). Supporting and ensuring the child's general well-being was seen as one of the ultimate goals of child protection and systemic practice. KINDL-R was chosen to measure this outcome as it had previously shown good validity (Bullinger et al., 2008; Erhart et al., 2009; Ravens-Sieberer & Bullinger, 1998) and reliability (Bullinger et al., 2008; Ravens-Sieberer & Bullinger, 1998) and had already been translated into Finnish. Based on their review, Korpilahti et al. (2015) concluded that KINDL-R is valid, reliable, and feasible, although more research with a Finnish population is needed. In this study, KINDL-R showed good internal consistency reliability (T1: $\alpha = .84$, T2: $\alpha = .84$).

Family dynamics according to the child and to the parent were measured using SCORE-15, an instrument developed to monitor progress and outcome in systemic family therapy (Jewell et al., 2013; Stratton 2010). Systemic practice was intended to improve family functioning and communication. SCORE-15 was chosen as it had previously been used in two studies assessing systemic practice (Bostock et al., 2017; Wilkinson et al., 2014) and the forms for adults and children had been translated into Finnish. In their review of previous research on SCORE instruments, Carr and Stratton (2017) concluded that SCORE-15 has good reliability, validity, and sensitivity to change. In this study, SCORE-15 showed good internal consistency reliability (Mothers at T1: $\alpha = .90$, T2: $\alpha = .89$, children at T1: $\alpha = .84$, T2 $\alpha = .86$).

Several measures designed expressly for this study were used. These included *need for outsider help according to the child, need for outsider help according to the parent*, *need for child protection according to the social worker*, *signs of abuse or neglect, number of meetings during the past six months*, and *child and parent service user feedback* (for more details, see A1 and A2).

5.3.5 Sample size, power and attrition

Given the absence of published studies on the association between the selected outcome measures and systemic practice model or any similar programme aiming to improve the effectiveness of child protection services, a power analysis was not possible. Instead, a target sample size was negotiated in each team participating in the data collection based on the team considered realistic given the implementation and the study timescale, and the practitioners' time resources (i.e., caseloads).

The sample size was 112 cases at T1 and 65 cases at T2. These numbers are for the cases in which the practitioner's form was completed. The number of completed children's and parent's self-assessment forms was lower. See table 5 for the number of completed research forms.

TABLE 5 Completed research forms in the SPM and SAU groups.

	T1		T2	
	SPM group	SAU group	SPM group	SAU group
Practitioner forms	56	56	28	37
Children's forms	28	43	19	29
Mothers' forms	41	40	23	25
Fathers' forms	12	13	6	8

Missing data analyses on the social workers' and children's forms showed that missing data were not related to study group or child demographics. Missing data on the outcome measures were related to just one of the outcome variables at baseline, i.e., neglect and abuse assessed by social worker. The amount of neglect and abuse was higher at T1 among those participating at both measurement points compared to those participating only at T1. This finding indicates that the cases participating at both measurement points were more severe than those participating only at the baseline. It is possible that a number of less severe cases had been successfully closed by the time of the T2 data collection, but unfortunately this could not be confirmed as no information on the reason for dropping out was available in 37 cases.

Among mothers, non-Finnish speaking mothers were less likely to participate at both measurement points. Among fathers, missing data were related to living arrangements, marital status, and age. Due to the small number of fathers who participated at both measurement points (n=9), fathers' subjective outcome and feedback measures were omitted from the analysis.

5.3.6 Statistical analyses

In article A1, the research question on changes in outcomes was addressed using repeated measures analysis of variance. In cases where the assumption of normality was not met, additional non-parametric tests were conducted. Differences in the frequency of meetings and service user feedback between the study groups at T2 was examined using the Mann-Whitney U-test. The association between the number of meetings and feedback variables was analysed using Pearson's Product-Moment correlation.

In article A2, first, informant discrepancy (mean-level differences) on the need for service intervention and the family dynamics was assessed using paired-sample t-tests. Change in this discrepancy between T1 and T2 across the whole sample and between the study groups was assessed by calculating the difference between the group means, and using these new variables in repeated measures analysis of variance. The extent of agreement between the informants was analysed using Pearson's Product-Moment correlation. Fisher's Z-test was used to examine whether the extent of agreement between the informants had increased more in the SPM than SAU group. The effect of agreement on outcomes was assessed by calculating change scores (T2–T1) and conducting multiple regression analysis. SPSS Statistics 26 was used in A1 and SPSS Statistics 27 in A2.

5.4 Context data

Using a mixed-method study design, data on context were used in A3 to explore the fidelity and the factors influencing the implementation of the SPM (Research Question 2). This article also used social worker interview data collected by Isokuortti (for a description of this dataset, see A3).

5.4.1 Administrative data

These data comprise basic information on the implementation sites and teams and were collected via emailed research surveys from local contact persons in May 2017 and February 2018. The first round of data gathering revealed differences between the informants in their views on the model and its definition. It was difficult to obtain precise information on team size, the availability of certain team members (especially the clinician, i.e., family therapist, and coordinator) or on the moment the SPM was implemented in each site or individual team. Hence, the researcher participated in organising workshops for managers together with the implementation team in THL and used this forum to increase the validity of the data. The second form was sent to informants prior to the second workshop and completed in a researcher-led session in the workshop. If information was still missing after this, the researcher reminded each site twice via email to complete the forms. The workshops were intended for high-level managers or leaders responsible for children's social care who would thus be in a position to promote systemic change within the organisation and organise sufficient resources for SPM implementation. In practice, however, the profile of workshop participants varied widely from temporary project workers to high-level leaders. Permission to use the information collected via the forms and in the workshops was requested from respondents and workshop participants.

The first research form included items seeking background information on the current situation in each site, such as the number of social workers in children's social care, the number and profile of teams, and open-ended questions on local implementation plans. This information was used to identify potential sites for participation in the outcome study.

The second research form included closed-ended questions on the SPM teams and implementation schedules at each site. This information was used to assess the fidelity of the SPM team structure. Three measures were used: *adoption of the team structure*, *the extent of the clinician's work in a team*, and *the extent of the coordinator's work in a team.* The dissemination material did not provide any concrete criteria for or indicators of successful implementation. Hence, the fidelity thresholds for each measure were defined by the researchers based on stakeholder discussions and the researchers' interpretation and understanding of the content and aims of the model.

A data subset on the 23 sites participating in this study was derived from the administrative data collected from all the sites participating in the implementation of the SPM. This subset was gathered from 39 teams. Data from one team

were missing. It is likely that at the time of the data collection, the team's implementation plans were incomplete, and hence managers were unable to provide the required information.

The data were organised, recoded, and analysed using SPSS Statistics 24.

5.4.2 Social worker survey

A full sample was used to collect the survey data, i.e., all the sites (n=31) that had informed THL that they were planning to start implementing the SPM at the latest in March 2018 were asked to participate in the study. Research permission was received from 27 sites. The email addresses of the social workers were collected from managers after permission to conduct the research at each site had been given. Three sites failed to deliver the social workers' email addresses and one site decided to withdraw from the study due to implementation-related difficulties. Thus, 23 sites, accounting for 74 percent of all the SPM sites, participated in the study. The SPM was implemented at these sites by 39 teams, accounting for 75 percent of all known SPM teams.

The survey was conducted in September-October 2018. The invitation and the link to the survey, and two follow-up reminders, were sent via email. All the participants gave their informed consent prior to their inclusion in the study.

The majority of the respondents had received SPM training and additional post-training coaching. While 45 percent of the SPM teams had an ideal team structure, only 27 percent of the respondents worked in such a team, i.e., in a team composed of a consultant social worker, a maximum of three social workers, a clinician, a coordinator and a maximum of eight team members in total. 12 percent of the respondents were unable to say if the model was still in use in their team. This reflected the incoherence of the model and its local implementation. On average, the respondents were responsible for 40 cases.

The design of the first draft of the survey was based on the THL's description of the SPM and the initial logic model. Items on the adoption of skills, knowledge, systemic tools and methods, and systemic practice were formulated based on the ToT syllabus and discussions within THL. As in the course of dissemination the content of the SPM evolved and became more concrete, the researcher participated in and observed discussions in the ToT sessions and manager workshops and used these observations to specify what implementation of the SPM was intended to achieve. Finally, the survey was refined based on reflective discussions with the steering group of the national evaluation and the first author of the fidelity study (A3).

The survey included items on the team structure, caseloads, practitioner skills and knowledge, availability of support, features of face-to-face practice, relationship with service users, work-related well-being, and experiences and ratings of SPM training and implementation in one's team.

Fidelity was assessed using the following measures: *the number of cases discussed in the team meetings, the use of key techniques, intensive case work* (i.e., the number of cases post-training in which respondent had been able to work more intensively than previously) and *SPM clarity* (8 items). The factors influencing fidelity were

explored with the following measures: *satisfaction with the training* (7 items), *experiences of the weekly team meetings* (2 items), *experiences of support received from a clinician* (2 items), *willingness to continue using the SPM*, *willingness to recommend the SPM to colleagues*, *the burden caused by the implementation*, and *caseload*.

The survey was sent to all the social workers (n=126) reported to have been members of the SPM teams participating in the study and whose email addresses had been delivered to the researcher. The number of social workers whose contact information had not been delivered to the researcher was not known due to missing information on some of the SPM teams and their composition. In total, 56 social workers completed the survey, yielding a response rate of 44 percent. The attrition was partly related to the fact that some of the social workers had already changed their workplaces and were not reached at the time of the survey. Their number was not available, as their contact information with local managers was not always updated. Discussions in the field with the social workers who also participated in the outcome data collection revealed that many suffered from research fatigue and general distress.

The quantitative data were organised and recoded and percentage frequencies calculated. The Spearman correlation coefficient was calculated to examine the correlation between the variables, and a one-way analysis of variance (ANOVA) was used to examine between-group differences in mean fidelity scores. The quantitative analyses were performed using SPSS Statistics 24.

5.5 Programme theory data

The programme theory data were used in the article A4, the aim of which was was to create a detailed description of programme theory for the SPM. The data were collected from key informants in a researcher-led process, in which reflective discussions were organised for the purpose of clarifying and specifying the components and intended outcomes of the SPM and explicating the assumptions regarding the causal chains between them in given contexts. As we were already familiar with the programme, we were able to identify the main issues that had caused confusion in the field or undermined implementation of the SPM, pinpoint contradictions, and formulate more precise research questions. The data collection and analysis were guided by a specific protocol (Gugiu & Rodriguez, 2007) designed for constructing logic models.

The research idea was presented to the ToT national steering group, all of whose members were invited to participate in the workshop process. Thus, the workshop participants (n=12) were key individuals in the development, training programme, and national dissemination of the model. Of these twelve, eight were practitioners in children's social care and six worked in an SPM team. Supplementary data comprised information drawn from the research literature, notes from the workshops and training sessions, written comments received from workshop participants, and personal communications with four informants:

three workshop participants in charge of SPM dissemination and one key informant outside the workshop process.

Data were collected during five workshops between January and August 2019. In addition, the researchers and most of the workshop participants participated in a 5-day systemic training course in between workshops run by two trainers from Collective Space (see Section 2.3.1). We also conducted two interviews, the first with these trainers and the second with one Finnish workshop participant. Supplementary data were collected by making notes in meetings with informants.

The semi-structured interview protocol (SSIP) proposed by Gugiu and Rodriguez (2007) for constructing logic models guided our data collection and analysis. We adapted this protocol for the purposes of a realist evaluation by first outlining a researcher's theory (Pawson & Tilley 1997, pp. 159-161) that was then refined in the workshops. We also elaborated the core components and logic model elements from a realist perspective by formulating hypotheses on the mechanisms of the SPM in different contexts. Finally, we formulated CMO configurations.

The transcripts were analysed with ATLAS.ti. As the first author, I initially coded the data by using key SPM components and other key SPM-related concepts as coding labels. Next, I and the second author produced a detailed description of each core component. We also identified contradictions and alternative ways of understanding the components and their functions. The supplementary data were used to add details and references to theories and methods. The first draft of the programme theory was then introduced to the research participants for collective discussion and refinement. Finally, the coded data were further categorised into context, mechanisms, and outcomes. We then constructed hypotheses on the causal chains generated by the SPM and created narrative and tabular versions of the two CMO configurations.

5.6 Research ethics

The study followed the Finnish Advisory Board on Research Integrity 2012 guidelines on the responsible conduct of research. Necessary research permits were acquired from all sites and stakeholders participating in the data collection. The research was not expected to cause any risks, damage, or harm to research subjects. Ethical approval was granted by the Finnish Institute of Health and Welfare Research Ethics Committee (2017–09).

My position as a researcher was a combination of evaluator and colleague in a team responsible for disseminating the SPM in THL. Although my role was not to make any decisions on the content of the SPM, I participated in formal and informal reflective discussions on the training and dissemination processes, shared previous research findings with the team and other stakeholders, and presented preliminary observations and findings based on my own data during the dissemination process. Lacking any previous concrete examples of formative

evaluation research projects in Finnish social work development projects, I often found it difficult to balance between critically appraising the dissemination process and being a team member supporting this process. On the other hand, being a member of the dissemination team allowed me to refine the hypotheses guiding my data collection and analysis along the way. To resolve situations that could cause tension or concerns, I consulted my supervisors and co-authors as well as the steering group surveilling the national evaluation.

The study processed personal data both on children's safety, service needs, health, and well-being and parents' health, well-being, and service needs. In the European Union's General Data Protection Regulation (GDPR), 'health' is the only information category that is listed in the 'special category of personal data'. However, all personal information associated with child protection that is collected from a child and a family is highly sensitive and has been treated as such in this study. This entails risk assessment and a specific technical and organisational protocol to ensure appropriate data security.

Personal data were collected as part of routine practice. To be able to use the children's forms in the study, informed consent from guardians and from children over 12 years old was required. The original plan was to collect written consents from all guardians. Since most of the children participating in the study were living with one guardian only and had no or minimal contact with their other guardian, this plan proved unfeasible. After consulting the Ethics Committee, it was decided to include in the study all cases in which consent had been received from at least one guardian. In the case of the parents' forms, parents were asked for their informed consent.

The data collected from the practitioners were used with the permission of the municipalities participating in the study. In Finland, municipalities control child protection records. The research forms only requested the kind of information relevant to child protection casework that is available to practitioners as part of their everyday practice. The reason for not asking service users themselves for permission to use these data was also to obtain data from service users with negative attitudes towards child protection or a low motivation or capacity to take part in research. Thus, the aim was to reduce bias and increase the diversity of the data. The decision to ask permission solely from the municipalities could be regarded as not fully respecting service users' autonomy. However, previous experience had taught me that service users in child protection may not be very active in the matter of taking part in research. Thus, to obtain as large and diverse a sample as possible, it was assumed that this strategy, along with the practitioners' contribution, would ensure a sample of this kind. Assessing the effectiveness of the SPM was seen an ethical task in itself, since providing knowledge on the model's intended (and possible unintended) outcomes would be needed in decision-making and ethical considerations in the future.

Once the outcome data had been collected, it was clear that, for reasons further elaborated in Limitations, the sample was considerably smaller than had been hoped. New plans to collect additional data could not be implemented due to circumstances related to COVID-19. Although it was understood that the small

sample size would limit the statistical analysis, it felt unethical to waste the contribution practitioners and service users had made in collecting the outcome data. Hence, the available data was analysed as thoroughly as possible.

Finally, a notably small number of fathers (n=9) participated in the whole study (both T1 and T2). Since the level of observation was a single case, the information collected from mothers and fathers could not be combined, e.g., *need for help according to mother* was a separate variable from *need for help according to father*. Different options for using the information collected from fathers were repeatedly discussed in supervision sessions. Unfortunately, there was no statistical tool that would have enabled comparison of the viewpoints of this group of informants with those of the other three groups, i.e., children, mothers, and practitioners. The low participation of fathers in this study might be related to their lower participation in child protection meetings compared to mothers. In future, both researchers and practitioners need to develop strategies to increase the participation of fathers in both processes.

6 EMPIRICAL FINDINGS

The aim of this dissertation research was to study the effectiveness of the *Systemic Practice Model for children's social care* (SPM) by applying a realist evaluation framework. More specifically, I wanted to find out whether the SPM worked, or could work, and, if so, for whom, under what circumstances, and how. In this section, I answer to these research questions by bringing together the findings of the four sub-studies.

6.1 Did it work? Comparing the effectiveness of the systemic approach and service-as-usual

The short-term family-level outcomes of the SPM were explored in two studies, in the pilot study (A1) and the agreement study (A2). The hypotheses were that the SPM should improve family dynamics and decrease child abuse and neglect and thereby decrease the overall need for child protection and improve the subjective well-being of the child. It was expected that these goals would be achieved by improving the quantity and quality of child protection practice, e.g., service user participation, relationship-based practice, power sharing, and more intensive casework. The pilot study focused on detecting changes in the child's well-being, safety, and family dynamics, and identifying the quality and quantity of services during the follow-up. The agreement study elaborated the analysis further by comparing children's, mothers', and practitioners' reports, and exploring the relationship between agreement at baseline and child protection outcomes at follow-up. Agreement was approached by analysing *informant discrepancy*, i.e., the mean-level differences between participant's scores on their families' service needs and family dynamics, and the *extent of agreement*, i.e., inter-rater correspondence. Comparisons were made between families supported by practitioners in teams applying the SPM and families supported by practitioners in service-as-usual teams.

Along with improvements in practice, the SPM's intention was to increase support for practitioners and improve their work-related well-being. This was measured in the social worker survey and the findings were reported in the article on implementation fidelity and influencing factors (A3) and in the THL report (Aaltio & Isokuortti, 2019).

6.1.1 Child's subjective well-being

In the pilot study, no significant change over time was found across the sample. In addition, study group had no significant effect on the changes in group means. In other words, systemic practice and service-as-usual had an equal effect on improving the child's well-being.

Based on the descriptive data on all children participating in the study, the raw estimates of the group means in the whole sample were somewhat lower at T1 but about on the same level at T2 as among the general population. In a study conducted among Finnish schoolchildren aged 7-17, the KINDL-R mean was 72.1 (SD 11.0) (Paakkonen & Paakkonen, 2018). In this sample, the mean was 70.5 at T1 and 72.8 at T2 among the children for whom the score could be calculated for both measurement points.

The agreement study explored the associations between the outcome measures. The results showed that the child's subjective well-being was associated with the child's assessment of the family's dynamics at both measurement points, and with mother's assessments of the family's dynamics at T1. This finding indicates a logical relationship between these measures: poor family dynamics was associated with poor well-being. Interestingly, children's subjective well-being did not correlate with practitioners' assessments of neglect or abuse or the need for child protection.

These finding will be further considered in the Discussion.

6.1.2 Family dynamics

The mothers in the pilot study reported that their family's dynamics improved. This could indicate that child protection practice in general had supported positive communication and family strengths and assuaged some of the families' difficulties. However, the children did not report a similar improvement. Again, no differences were found between the SPM approach and service-as-usual in their relative ability to generate progress in the family's functioning.

The agreement study explored discrepancies and the extent of agreement between the participants. Across the sample, no significant mean level differences between the children's and mothers' groups were found at either T1 or T2. In other words, the children and their mothers, on average, gave the same assessment of the quality of their family's dynamics. The differences between the group means showed no significant effect of time or effect of group-by-time. This finding indicates that both approaches were equally effective in improving the level of agreement between children and mothers on family's functioning.

Although no discrepancy was found between the children and their mothers on the group level, a separate analysis was needed to explore the extent of agreement between child and mother on the case level. For this purpose, inter-rater correspondence was investigated. The results for the whole sample showed a significant association, indicating a high level of agreement between children's and mothers', in their assessments of their family's dynamics at both T1 and T2. However, no difference in inter-rater correspondence was found between the study groups.

6.1.3 Need for service intervention

All three informants - children, mothers, and practitioners - were asked to assess the family's need for service intervention. Across the sample, the need for help according to the child or mother did not show any significant change at T2. In addition, no significant difference was found between the study groups. When the family's situation was observed from the practitioner's perspective, the need for child protection showed a main effect of time. Across the sample, the mean decreased from 6.5 to 4.6, indicating an improvement in the child's situation. Again, no differences were found between the approaches. Thus, the findings of the first sub-study (A1) suggest that a positive change in the family's overall situation occurred during the follow-up according to the practitioner but not according to the family.

The second sub-study (A2) examined the agreement between children, mothers, and practitioners more closely. In general, the family's situation was appraised as more serious by the mothers than by their children. Across the sample, the results indicated a significant mother-child discrepancy at both T1 and T2. No significant differences between mothers' and practitioners' assessments were found, but a significant child-practitioner discrepancy was found at T1.

The difference in means between the practitioners' and children's groups in the whole sample decreased from 3.1 to 1.3 over the follow-up period, meaning that their views converged. However, practitioners' and mothers' views did not show a similar level of convergence.

On case-level agreement, the children's scores showed a medium-level correlation with their mothers' scores across the sample at T1 but not at T2. In contrast, practitioners' assessments of the need for service intervention did not correlate with those of children or mothers. These findings indicate that the extent of agreement between practitioners and family members was poor.

No significant differences were found, between the study groups in either analysis, i.e., both approaches were equally effective in bringing practitioners' and children's views closer together.

6.1.4 Abuse and neglect

The pilot study found no significant differences in the amount of abuse and neglect assessed by practitioners either over time or between the study groups.

When the associations between outcome measures were explored in the agreement study, a significant association was found between practitioner-assessed signs of abuse and neglect and practitioner-assessed need for service intervention: the correlation was strong at T1 and moderate at T2. This indicates that practitioner's assessments of the child's situation were coherent. However, as previously noted, practitioners' assessments of neglect or abuse did not correlate with children's subjective well-being (for further discussion, see 7.1). A moderate correlation was found between neglect or abuse and child-assessed family dynamics at T2. No significant between-group differences in the correlations were found.

The second article (A2) found that, across the sample, the change in abuse or neglect was predicted by children's and mothers' agreement on the need for service intervention at T1. Child-mother disagreement predicted a negative and child-mother agreement a positive outcome. Specifically, the greater the disagreement between child and mother at the beginning of the process, the greater the predicted increase in abuse and neglect. Correspondingly, the greater the agreement between child and mother, the greater the predicted decrease in abuse and neglect. Interestingly, this predicted decrease was greater when a child and mother agreed that they did not need help (Figure 1).

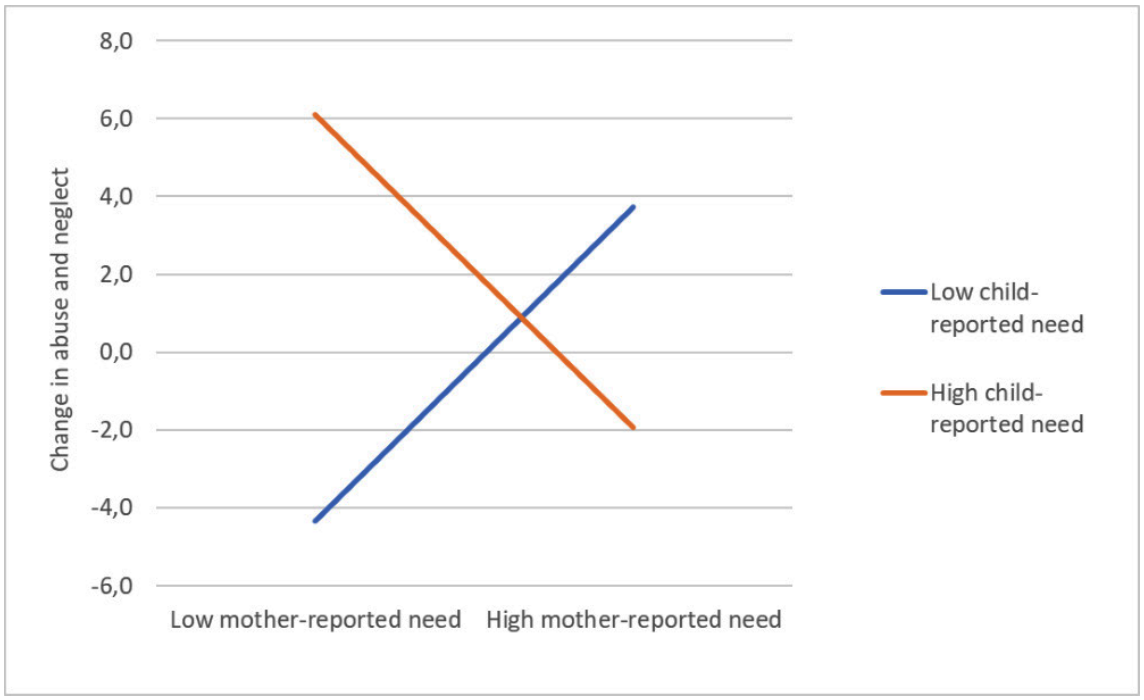


FIGURE 1 Interaction between child-reported and mother-reported need at T1 predicted change in practitioner-reported abuse or neglect.

A more unexpected finding was that when both child and mother reported not being in need of outsider help, the amount of abuse and neglect decreased indicating that child protection practitioners had carried out a successful intervention. This finding is considered further in the Discussion. An expected result was that

when both child and mother reported needing outsider help, this agreement resulted in a positive outcome in terms of practitioner-rated abuse and neglect.

6.1.5 Intensity and quality of the practice

The findings of the pilot study suggest that systemic practice did not differ from service as usual in any outcome measures. Nevertheless, it was found to be more intensive in terms of the volume of meetings between practitioners and family members.

First, while meeting children without the presence of other family members was infrequent in both groups, the social workers in the SPM group met the child twice whereas the social workers in the SAU group met the child only once during the six-month follow-up. Second, in the systemic teams, the total of all meetings between a case-holding social worker and a family was more than double (M = 11.1) that in the service-as-usual teams (M = 5.0). Third, the mean total number of all meetings between any child protection practitioner (social worker or family practitioner) and a family in the SPM group (M = 22.6) was nearly threefold that in the SAU group (M = 8.5). Half of the families in the SAU group had met their child protection practitioner less than once a month, compared to only one in ten in the SPM group. These results reveal that systemic practice was considerably more intensive than service as usual.

Interestingly, based on service-user feedback, more intense practice did not result in better quality of practice. The results of the pilot study (A1) showed no differences between the study groups in children's or mothers' sense of participation, sense of trust, or overall satisfaction with child protection.

Across the sample, mothers were more satisfied with the process than children. This result could be linked to the fact that children seldom met their social worker. Interestingly, the number of private meetings between a child and a case-holding social worker correlated negatively with the child's trust in their social worker. Even more surprisingly, this variable correlated negatively with mothers' satisfaction with services. This finding is further discussed in section 7.2.

6.1.6 Practitioner-level outcomes

The intention of the SPM model is to support practitioners and increase their work-related well-being. The findings were inconsistent. The majority, 76 percent, of the survey respondents reported that collective reflection in team meetings had improved their practice, and 78 percent perceived they had received the support they needed from these meetings (Aaltio & Isokuortti, 2019). Moreover, 58 percent of the respondents reported that the SPM pilot study had increased their work motivation. As reported in the fidelity study (A3), 79 percent wished to continue with the model and 76 percent were willing to recommend it to colleagues. However, 44 percent of the practitioners experienced the implementation of the model as a burden. Half of these respondents wanted to continue us-

ing the model, but half of them did not. Thus, one fifth of the practitioners participating in the study felt that implementing the SPM had negatively affected their work-related well-being.

6.2 In what context does the SPM work or not work, and why?

The key finding from the two outcome studies was that the SPM did not outperform service as usual in either the selected outcome measures or service user feedback. Practitioners, in turn, were divided in their assessments on SPM. This finding needs to be interpreted in relation to SPM implementation. The fidelity study (A3) analysed fidelity in implementing the SPM, i.e., the extent to which the implementers adhered to the programme as it was designed by the developers, and the potential factors influencing adherence. Based on the findings, we formulated hypotheses on the possible relationships between the implementation context and fidelity. Here, I focus on the relationship between SPM fidelity and the family-level and practitioner-level outcomes.

The SPM was introduced in a local context unaccompanied by any specific organisation-level changes and was primarily implemented by the pre-existing child protection teams. Originally, the SPM was intended to initiate systemic change in the organisational culture and structures, including decreasing caseloads. However, during the first stage of dissemination, no concrete resources were provided for such changes. Instead, ascertaining what organisational reform would actually involve in their local contexts was left to sites themselves. Moreover, covering the costs of the additional resources needed to reduce caseloads was also left to the local organisation. At some sites, the caseloads of the SPM teams were reduced by recruiting additional social workers, but this was rare. Rather, the budgets and expenditure on child protection for the year of implementation had already been approved before the managers responsible for SPM implementation realised that the LAPE programme did not provide any funds earmarked for this purpose (Aaltio & Isokuortti, 2019). Hence, with a few exceptions, the programme did not change the organisational context in which the SPM teams were working.

The intended key resources of the programme were the specific team structure, weekly team meetings as a forum for case supervision, the contribution of a clinician in case supervision, the input of a coordinator in administrative tasks, and the training of practitioners in systemic skills and thinking. These resources were expected to change practitioners' reasoning in at least in two ways. First, systemic case discussions and training was expected to result in systemic thinking, which would then result in systemic practice. Systemic practice, in turn, was expected to result in improved family-level outcomes. Second, the team structure, teem meetings and coordinator's input were expected to give practitioners a sense of collective support and shared responsibility, leading to improved work-related wellbeing.

The fidelity analysis showed that 45 percent of the teams had the intended team structure. Three teams out of four had the intended clinician resource available, i.e., a clinician was present at the weekly team meetings. In contrast, the contribution of the coordinator was much lower than expected. The minimum contribution of a coordinator to serve the original idea of giving practitioners extra support had been projected to be 50 percent of the individual's weekly working hours. Only one team in five had a full-day or half-day coordinator. Moreover, a quarter of the teams did not have a coordinator at all. One third of the teams had more than three social workers, i.e., these teams were larger than intended.

When asked about the training programme, only a quarter of the respondents felt it had been concrete enough. Overall, 43 percent of the respondents reported dissatisfaction with the training they had received. This was evident in the interviews, where some of the practitioners expressed dissatisfaction with the training and felt that it had not furnished them with concrete suggestions and guidance on systemic practice and techniques. However, 40 percent of the survey respondents were satisfied with the training.

Based on the survey, the number of cases discussed in a team meeting during the intervention period was rather low. On average, social workers had brought only one case per month to the case discussion. Only five percent of the respondents had had a weekly opportunity to present one of their own cases to the team. This was expected given the large team size and the social workers' high caseloads (M = 40.3). As a result, the total number of cases that within one team could amount to as much as two or three hundred. Not surprisingly, the vast majority of these cases were not exposed to collective reflection.

According to the interviews, some teams had been able to form a safe learning environment resulting in shared learning motivation. In some teams, learning was also supported by hands-on coaching provided by a clinician or other skillful team member. These individual capacities of team members and the interpersonal relationships that developed between them supported the take up of the SPM. Positive experiences of practice strengthened their feelings of competence and commitment to systemic practice. At the same time, the low fidelity of the model lowered participants' commitment to the SPM. For example, in many teams, high caseloads forced practitioners to select cases for applying the new approach. When practitioners felt unable to implement the model as intended, their motivation dropped.

As described above, the implementation sites encountered considerable challenges in providing the requisite SPM resources. As a result, the practitioners' reasoning did not change as comprehensively as had been hoped. In some teams, many positive changes were reported by practitioners in their interviews, e.g., considering a family's situation from multiple perspectives with the help of a family therapist and avoiding making hasty decisions. Nevertheless, the adoption of systemic thinking and techniques varied between teams. In sum, some practitioners had managed to embrace a systemic orientation, while others expressed frustration and confusion on how to use systemic ideas in real-world settings.

The service user interviews indicated that families were satisfied with the help received from child protection in general but could not identify any significant impact of systemic practice on these services. Similarly, in the outcome data, parents and children both gave positive feedback on the child protection services they had received, and no differences were found between those supported by the systemic teams and those receiving service as usual.

Thus, with the notable exception of a higher volume of service user meetings, no comprehensive change in practitioners' reasoning or practice was detected. At first glance, it seems surprising that while practice was significantly more intense in the SPM group, this did not lead to better outcomes or more positive service-user feedback. However, this finding could serve as an important reminder that in human services, the quality of practice is as important as its quantity – if not even more so. The fidelity analysis on 23 sites indicated that the majority of practitioners had not in fact used systemic techniques with service users; only a few had adopted a systemic orientation in general, and only a very low proportion of cases had been exposed to systemic supervision and collective reflection. If the key resource of the SPM, i.e., systemic understanding on what the case could be about and how to proceed with the family, was absent, increasing the number of meetings did not result in better outcomes. This could happen even when SPM methods have been used. In the pilot study, additional analysis was conducted to evaluate the effect of fidelity on outcomes. Three fidelity groups were compared: i) cases in the SPM group where none of the key SPM methods had been applied, ii) cases in the SPM group where at least one of the key SPM methods had been applied, and iii) cases of service as usual. The results showed that the use of SPM methods did not result in more effective practice at the three sites participating in the pilot study.

Another apparent paradox was the finding that the number of private meetings between a child and a case holding social worker correlated negatively with the child's trust in her or his social worker, and with mothers' satisfaction with services. The general context of child protection practice could explain this finding. Quite often, the practitioner responsible for building a direct relationship with the child and meeting the child individually is not a social worker but a family practitioner working as a partner in a given case. It is possible, that in this context, service users have reservations about a social worker in this role. This division of tasks offers another possible explanation. Social workers who have no previous experience of working with children directly may lack the skills needed to create a trustful relationship with them. Either way, merely increasing the number of meetings is not enough to improve the effectiveness of practice.

Finally, the practitioner-level findings were also mixed. One fifth of the practitioners participating in the study felt the implementation of the SPM had negatively affected their work-related well-being. In the interviews, practitioners reported lack of time for team meetings, family meetings, implementing new techniques. In some teams, the clinician and coordinator resources were terminated. Practitioners also mentioned low commitment on the part of their agency

leaders, disorganised implementation, and stress due to simultaneous organisational reforms unrelated to the SPM. The findings indicate that in some cases, simultaneous and interacting contextual factors inhibited the firing of the mechanism intended to promote a sense of collective support and shared responsibility.

Nevertheless, the majority of the survey respondents appreciated the opportunity for collective reflection in the team meetings and assessed it as having improved their practice. For nearly two thirds, the SPM pilot had increased their motivation to work, four out of five wished to continue with the model, and an equal number were willing to recommend it to colleagues. Interestingly, caseloads were not associated with willingness to continue with the model or experiencing it as a burden. Instead, respondents with higher caseloads were less willing to recommend the model to their colleagues. While the implementation of the SPM was challenging and relatively far from ideal, for many social workers the model seemed to possess great potential. During the first stage of implementation, the mere idea of such a model together with some of the implementary baby-steps seemed to provide a resource that in the right conditions, i.e., a positive team atmosphere, could lead to learning and the adoption of a systemic orientation, and hence to experiencing a positive motivation and inspiration in relation to one's work.

6.3 Why should the SPM work?

Finally, to clarify the key components, intended outcomes, expected mechanisms, and understanding of the ideal context of implementation, the programme theory sub-study (A4) was conducted together with a group of key informants. The aims were to enable high quality implementation, and meaningful outcome and process evaluations of the SPM. It was also hoped that the findings would aid organisations and practitioners in assessing their readiness for implementing the SPM and preventing adverse effects.

The key question guiding this sub-study was "Why should the SPM work?" What are its key components, and how do these differ from service as usual? What are the intended outcomes, and why should the SPM have more potential in achieving these outcomes than service as usual? What kinds of resources do the key components provide, and what kind of reasoning are they expected to initiate in practitioners and service users? Finally, what kind of context is needed for these mechanisms to operate as intended?

The process was motivated by observations of the challenges encountered during the first stage of implementing the SPM. Together with the various stakeholders we hoped to deepen our understanding of the potential causal powers of the SPM, produce hypotheses on how and in what context the mechanisms could work, address some of the contradictions in the design of the SPM, and aid programme designers to prioritise its elements and clarify its target outcomes. The

analysis yielded three core components (a systemic team, systemic weekly meetings, and systemic practice) and two CMO configurations representing the causal chains of the model.

6.3.1 Systemic team

An ideal systemic team would comprise the following: a consultant social worker, a systemic family therapist (a clinician), a coordinator and two to three social workers. A team could include family practitioners. While this description did not differ from the description used in the initial logic model (see Section 2.3.3), but the process clarified that the systemic team should not include "other social worker(s)", such as a practitioner working with a parent, as mentioned in the first SPM description. These other practitioners working with and around the family could be invited to weekly or other meetings, but they were not to form part of the permanent systemic team. Instead, they were part of multiagency practice. The number of case-holding social workers was considered to be no fewer than two, so that they could fill in and provide each other with peer support, but no more than three to ensure time and space for systemic team discussion. During the first stage of implementation, the role of a consultant social worker had been unclear. Now it was agreed that the consultant social worker should function as a mentor leading the team and supervising cases alongside the clinician. The clinician, in turn, should present new insights and generate reflexive curiosity while supporting the practitioners in adopting systemic thinking, methods, and tools.

6.3.2 Systemic weekly meetings

Weekly meetings were defined as a forum for systemic case discussions, in which all team members would actively participate. Collective reflection from multiple perspectives should aid the case-holding social worker to generate hypotheses and plan the next steps. The nature and theoretical underpinnings of a systemic case discussion were specified: the discussion should be curious and respectful. By applying the three Domains of Action theory (Lang et al., 1990), practitioners should consider the ethical aspects of their work (domain of aesthetics), explore various ideas and perspectives with curiosity (domain of explanation), and decide how to act based upon the case discussion (domain of production). They should also reflect on their role as part of the system and, possibly also, in maintaining the family's problems.

6.3.3 Systemic practice

A definition for systemic practice was formulated in the workshop process. According to the participants, systemic practice was:

> "purposeful, relation-ship based, and systemic-oriented work with families, in which the focus is on the relationships between child and family, and the physical and immaterial living-environment of both, including the child protection system and practitioners."

The key methods and techniques of systemic practice would include the use of genograms, the formulation of hypotheses (Cecchin, 1987), and the use of circular and other types of questions (Tomm, 1988). The latter technique was not included in the SPM training in the first stage of SPM dissemination. Similarly, some of the techniques were now excluded or, at least, not given a central role. The key methods were expected to provide insight and new perspectives on families' dynamics, history, patterns, and narratives for both practitioners and service users. They should increase practitioners' understanding. For family members, the use of these methods should generate compassion, self-understanding, empowerment, interaction within the family, and also help them to create their own solutions. Systemic practice required a trusting and affirmative relationship between the practitioner and the child and parents. To create such a relationship, practitioners should 'seek permission' (Aggett et al., 2015), acknowledge the limits of their knowledge (Anderson, 1997), and be respectful and curious (Cecchin, 1987). The aim would be power-sharing between the social worker and the family while keeping the child safe.

6.3.4 Competing ideas

The process aroused differences of opinion on three main themes. First, the workshop participants had different ideas of the composition of a systemic team. For some, the team could mean a "network" that included professionals from other organisations while others favored the original idea of a permanent team as in the RSW. Although multiagency practice was regarded as highly important, it was finally decided that a systemic team would comprise permanent members only. Second, views on the number of cases that should be discussed in a team meeting varied. Some argued that the meeting should focus on one or two cases, thereby allowing in-depth discussion, while others saw that most of the ongoing cases should flow through the weekly meetings. Finally, it was decided that each case-holding social worker should have an opportunity to bring one or two cases to be discussed in the meeting each week. Third, three possible functions for the weekly team meetings were identified: forum for collegial reflection and learning, forum for practice, and forum for multiagency practice. The first option would provide an exclusive space for practitioners, enabling them to reflect on possible prejudices and emotions which might unconsciously affect their decision-making. The second option entailed inviting families to the weekly meetings, and the third option suggested inviting both families and professionals working with a family to the meeting. Arguments for the different options were both practical and ideological. Some participants saw practitioners as needing a forum for learning and reflection without the presence of the family, while others argued that all the discussions should be open to the families. The decision to restrict weekly meetings to the permanent members of the systemic team was practical: having families and other professionals present would take up more time.

6.3.5 CMO configurations

Based on our synthesis of the findings presented above, we formulated two CMO configurations representing the ideal operation of the SPM, first on the team level and then on the family-level.

To operate properly, a systemic team should be embedded in a systemic organisational culture and management that supported the team's frontline practitioners and daily routines (Table 6).

TABLE 6 The intended causal chains at the practitioner level (as presented in A4).

Practitioner level			
Resources	**Context**	**Reasoning**	**Intermediate outcomes**
Systemic team + Systemic training and coaching + Systemic weekly meetings providing systemic case supervision	Organisation and leaders who implement daily practices and resources to support the proper functioning of the systemic teams, e.g., ensuring reasonable caseloads and recruiting the requisite team members + A team that forms a safe space for learning	Practitioners expand their understanding of family dynamics and begin to approach problems systemically from multiple perspectives + Mutual learning and reflection + Systemic team shares responsibility which provides emotional and practical support for practitioners	Intensive systemic practice, including respectful, power-sharing and curious relationships with families + Improved work-related wellbeing + Decreased staff turnover

For practitioners, an ideal team provides a safe learning environment. With systemic training and coaching, the SPM will equip practitioners with new knowledge and understanding, new communication skills, and strategies that can help in addressing sensitive questions with family members. Shared responsibility acquired through team discussions will help social workers make better decisions for families. Practitioners' reasoning will be changed by systemic training and case supervision, thereby improving their skills (i.e., permission-seeking and being respectful and curious) and changing service-as-usual into systemic practice. Practitioners' work-related wellbeing will improve when they have time and permission to use acquired skills and expertise, they will share responsibility and receive collegial support from the team.

The intermediate outcomes on the practitioner level now form the context and mechanisms for the family level (Table 7).

TABLE 7 The intended causal chains at the family level (as presented in A4).

Family level			
Resource	**Context**	**Reasoning**	**Intermediate and long-term outcomes**
Skillful and permanent practitioners who experience wellbeing + Systemic social work practice	The family meeting context is respectful, power-sharing, and curious + Regular and frequent meetings with the same practitioner/s	Identifying multiple perspectives + Identifying problematic interaction and communication, e.g., family patterns or scripts + Family's own motivation for change aroused	New insights and change in beliefs + Improved family dynamics + Increased safety + Improved parent/ child wellbeing

Systemic practice and practitioners thus serve as a resource for family members seeking to change their behaviour or situation. A systemic social worker seeks to improve the family's communication and dynamics by aiding family members to identify problematic interaction and communication. Instead of telling the family what to do, a skillful social worker tries to evoke the motivation for change in family members themselves. If family members are provided with new insights, this may increase the likelihood of longer-lasting behaviour change. The final outcomes of systemic practice include the increased safety and wellbeing of parent/s and child/ren. Since the change is built on the family members own motivation and resources, it can be assumed that systemic practice has the potential to achieve outcomes that last longer than those induced by service as usual.

6.4 Limitations

6.4.1 Limitations of the study design and data used to evaluate the effectiveness of the SPM

First, the small sample size limits the representativeness of the effectiveness data and the attrition between the two measurement points hampers between-group comparisons. Based on observations in the field, the high attrition seemed to be connected with practitioners' frustration with high caseloads, lack of supervision, insufficient training for the new model, simultaneous organisational reforms, and staff turnover. In some instances, attitudes towards the research were highly

negative. Although the researcher offered to provide hands-on guidance to support practitioners in completing the forms, her offer met with no response from either teams or managers. The data collection was successful at one site, where the manager responsible for both the SPM implementation and local data collection actively supported the frontline practitioners and maintained contact with the researcher. At two other sites the practitioners' motivation was random, and in general diminished after a positive beginning. In the future, more collaboration is needed between the researcher and local sites to support key persons in motivating individual practitioners. In addition, it would be worth considering whether one researcher can reasonably be expected to succeed in collecting representative data. In the field of social work research, it might have been a tradition to collect small samples of qualitative data with no intention to generalise the findings. However, if the future intention is to produce more evidence to support social work practice, studies need large samples and carefully designed data collection strategies to ensure the representativeness of the data.

Second, the study used a quasi-experimental design, as the researcher could not determine in advance the allocation of the participants to the study groups. In theory, the teams participating in the SPM implementation should have been randomly selected. This, however, would have entailed high-level decision on the national level, whereas the implementation strategy was to leave these decisions to the local level. Similarly, changing the rules of case allocation to enable randomisation of the families participating in the research would have entailed considerable changes in established procedures and hence high-level decisions. As a doctoral researcher solely responsible for planning and conducting the collection of the outcome data, I had no authority to intervene in the case allocation. However, this study provides important information for conducting future experimental studies in the social work context.

Prior to the data collection, the researcher visited each site or team once or twice to present the research project and instruments, and to negotiate on the practicalities of the data collection. Due to high caseloads, the original attempt to collect a whole sample from each time had to be abandoned. To decrease selection bias, the social workers were instructed not to exclude any cases based, for example, on their native language (non-Finnish speakers) or willingness to cooperate. While some practitioners reported filling in the forms through an interpreter, some reported they had not had time for this. In the future, research forms should be translated into several languages to ensure the equal participation of different language groups.

The interval between the baseline and follow-up measurement was approximately six months. This interval is often regarded as too short to detect change in outcomes. On the other hand, it could be argued that if a child is in immediate need of protection, such an intervention should bring about at least some immediate or short-term relief in the child's situation. Another question is how to decide what these changes should centre on. No consensus over child protection outcomes exists (Forrester, 2017), and few validated measures are available for this population. While the two validated outcome measures used in this study,

i.e., KINDL-R and SCORE-15, have been validated in other contexts, it is not yet clear that these measures are feasible, valid, and sensitive to change in the context of child protection social work. Neither of these outcome measures distinguishes between a child's living arrangements, i.e., the possibility that the child has two homes instead of one and a sense of belonging to two separate families instead of one. Based on the feedback from some of the children participating in the study, some of the questions in these measures were difficult to answer when a child has two homes.

Finally, the systemic approach had not been fully implemented during the study period. Hence, the intended study design, i.e., comparing systemic practice to service-as-usual, was watered down. On the other hand, combined, these findings address the need to increase the leverage of SPM as a programme. The volume of resources provided in the first stage of SPM implementation was not enough to change practice in a way that would produce superior outcomes.

6.4.2 Limitations of the study design and data used to evaluate fidelity in relation to the SPM outcomes

First, the fidelity data were based on self-reports. Observations of team meetings or direct practice would have provided richer data for analysing the implementation of the SPM and especially for formulating hypotheses on the mechanisms underpinning the practitioner-level outcomes and thus also the family-level outcomes.

Second, no validated or any pre-existing tool for measuring the fidelity of the SPM was available. Hence, we had to formulate fidelity indicators and thresholds based on our personal understanding of the model's design.

Third, it was not possible to relate the fidelity data to the outcome-data on the team level. The survey data on the respondents' worksite had been collected on the municipality level and not the team level. A more in-depth fidelity analysis on practitioners' reasoning was performed on the team level at three sites participating in the outcome study on the data from interviews conducted by another researcher. Unfortunately, these ratings could not be related to the outcome data since at one site the information on the team to which the case had been allocated had been inaccurately coded. The teams at this site had been divided into smaller systemic teams, and this new information was not available for the outcome data. Hence, no statistical analyses on the effect of systemic reasoning at the team level on case-level outcomes could be conducted. It was, however, possible, using practitioners' reports, to analyse the effect of the use of key SPM methods on outcomes at the case level. Unfortunately, this data did not allow me to assess the quality of use of these methods.

Fourth, the fidelity data focused on the implementation of the core components of the SPM but did not explicitly address the underlying mechanisms. The hypotheses on the mechanisms that would lead to intended outcomes had not yet been formulated by the time of the data collection. In fact, this situation motivated the third research task, i.e., developing a refined programme theory for the SPM.

6.4.3 Limitations of the study design and data used to develop the programme theory

To achieve a national consensus on the SPM in collaboration with the agency responsible for its development and dissemination meant that our inclusion criteria for research participants had to be narrow. The data did not involve service users and only a very restricted sample of practitioners was recruited, i.e., those previously involved with the national development and dissemination of the SPM. Our prior understanding might have constrained the choices we made in sampling. Our findings at the time indicated that it had remained unclear to many practitioners precisely how the SPM differed from service as usual, and their use of systemic tools and methods with service users was rare. Since it seemed that the SPM was not properly in use at the service-user level, we excluded both practitioners and service-users who did not seem to have experience of the key components and mechanisms of the SPM. Without this pre-knowledge, we might have searched for sites that had been more successful in SPM implementation and from these recruited frontline practitioners and service-users to participate in the study. It is possible that such sites could have been found outside those we were familiar with.

When we began the process, we were completing our national evaluation of the implementation of the SPM and had conducted preliminary analyses on the data used in the fidelity study. Thus, we had an insight into the key issues that needed clarification and critical appraisal. In addition, we had observed that the meanings of some of the components and terms differed among both the stakeholders in the field and those at the national level. Our position as formative evaluators enabled us to observe the dissemination process both from a close distance and, partly, as an outsider. On the other hand, our observations on the challenges presented during implementation and our position as evaluators might have had a negative effect on the researcher-participant relationship. In any case, our impression was that recruiting and motivating research participants was straightforward and initiating discussions in workshops was effortless.

The focus of the programme theory is on the team and family levels. For the SPM to operate as intended, many organisation-level changes are required. The operationalisation of these was outside our remit, although the need for such a process was acknowledged.

In conclusion, the programme theory represents an ideal description of the SPM. The next step would be to test these hypotheses in real-world settings.

7 DISCUSSION AND CONCLUSIONS

7.1 Discussion

The purpose of this dissertation was to evaluate the effectiveness of the SPM, analyse how its mechanisms affected or could affect its outcomes, and analyse the components and mechanisms that should make the SPM different and more beneficial than service as usual.

The key finding was that during its initial implementation the SPM did not outperform service as usual as far as the selected outcome measures were concerned. The observed outcomes were analysed in the light of realist understanding of causation: the activation or de-activation of the mechanisms that would produce the outcomes depend on the context. The analysis of the fidelity of implementation, i.e., the extent to which the implementers adhered to the programme as it was designed by the developers, found several barriers that inhibited practitioners from adopting the SPM and applying systemic practice as intended: the core components of the SPM were unclear, training was insufficient, and organisational factors (i.e., high caseloads, high staff turnover, inadequate support from leaders) hindered its implementation. Finally, to aid future implementation and research, a programme theory, including a detailed description of the SPM's core components, was formulated together with key informants involved in the national development and dissemination of the SPM.

While the challenges in the data collection and implementation fidelity limit the conclusions that can be drawn from this thesis, the findings regarding changes in outcomes across the sample merit consideration. First, no significant change over time was found in the child's subjective well-being. This finding could be unsettling for child protection practice in general. Although no consensus exists on the desired outcomes of child protection, the Finnish system has been described as a child-oriented system that aims to secure the 'best interest of the child'. One of the components of the child's best interest is the child's well-being. Hence, it could be argued that child protection should aim at short-term positive changes in the child's well-being even when it is acknowledged that this

is not always possible for reasons outside sphere of influence of the child protection system. In a child-oriented system, the child's subjective assessment of his or her well-being should be one of the key indicators of the system's success.

However, due to lack of tradition in measuring children's well-being within child protection, we have little knowledge of the validity and reliability of these measures for this population. It is possible that the standardised instrument used in this study, KINDL-R, is not best suited to the target population. It might be too generic to detect the effects of neglect and abuse on a child's well-being. The raw estimates of the group means across the present sample were somewhat lower at T1 and about on the same level as among the general population at T2 (Paakkonen & Paakkonen, 2018). If the change in the KINDL-R scores had been significant, this could have indicated that the child protection services had succeeded in bringing service user children's well-being closer to that of their peers. However, the present raw estimates do not permit such a conclusion. This comparison serves to illustrate the possible benefits of the child protection system using standardised measurements to assess its short-term contribution on the group level.

The analysis of agreement between child, mother, and practitioner showed that children's subjective well-being was not associated with practitioners' assessments of neglect or abuse or need for child protection. Assuming that a practitioner's concern about a child is justified, this finding could be explained in several ways. First, it is possible that service users are unwilling to expose their true situations to practitioners, especially as practitioners have the power to separate families. The disparity between a practitioners' and a child's reports on the child's situation might also be due to the child's attempt to protect the family's private sphere. Previous research has found that loyalty to parents may prevent children from discussing their family life with social workers (Pölkki et al., 2012; Wilson et al., 2020). Second, lack of information on child protection services and lack of the support needed to participate in the process may impair children's ability or willingness to be involved (Diaz, 2020). Moreover, the psychological and emotional stress caused by the process of investigating abuse, and mistrust of social workers may hinder children from disclosing abuse and neglect (Wilson et al., 2020). Third, in some cases, violence experienced at home (Paterson-Young, 2021) or by peers in the school environment (Allnock & Atkinson, 2019) may become 'normalised'. For some children, emotional disengagement may serve as a coping mechanism (Paterson-Young, 2021). Thus, social worker may assess as problematic an environment or experiences that seem normal to the child. Fourth, it is possible that either of the subjective measures used in the study, i.e., KINDL-R or SCORE-15, are sufficiently sensitive to the problems of families in child protection, despite the fact that SCORE-15 was especially developed to describe aspects of family strengths, difficulties, and communication. SCORE-15 has shown to be valid, but only in the therapeutic context. It is possible that this measure is not the best possible in the non-therapeutic context, i.e., child protection.

The extent of agreement between practitioners and family members on the need for service intervention was poor. This finding indicates that a family's situation looks different from different perspectives. Factors that a practitioner considers as presenting a potential risk, and hence a need for child protection, might be regarded by the family members as manageable difficulties or normal behavior. Previous research has shown that some parents who have a history of childhood trauma themselves may have difficulties in distinguishing between abusive behavior and acceptable boundary setting when administering discipline (Siverns & Morgan, 2019). Children's perspectives and understanding of their situation also differ from that of their mothers (Miller & Bentovim, 2003). Another possible explanation is that the expressed need for help does not reflect the true experience of a child or mother. Distrust of child protection agents or researchers could affect the responses given. Even if the reasons for disagreement are not known, awareness of its occurrence is important for practice and research.

The agreement study found that greater disagreement between child and mother at the beginning of the process predicted a greater increase in abuse and neglect, and that greater agreement between a child and a mother predicted a greater decrease in abuse and neglect. Interestingly, this decrease was greater when the child and mother agreed that they did not need help than when they agreed that they needed help. The first interpretation of these finding is that disagreement between child and parent is valuable information that should be further explored by the practitioner and used in planning measures for the child's protection. A previous study on a psychotherapy intervention for children with psychosocial deficits, Goolsby et al. (2018) reported a similar finding, i.e., parent-child discrepancy was associated with poorer treatment outcomes. The authors suggest that possible factors causing the discrepancy could be family functioning, awareness of problems, motivation to be involved in treatment, and conflicting perspectives. These factors could also explain child-mother discrepancy in the context of child protection social work. But what would explain the finding that child-mother agreement on no need for outsider help resulted in a decrease in abuse and neglect? A first possible explanation for this is that the need for child protection was relatively high in such cases even if it was rejected on the part of the family. Positive changes in parents' behaviour generated by a child protection intervention could be manifested in several areas of child's treatment simultaneously, resulting in greater change in the outcome measure of interest. Another, possibly related, explanation is that when a child is deemed to need child protection while the family rejects the need for this, practitioners focus more closely on assessing the child's situation, identify more signs of abuse and neglect, and make more effort to support the family. Unfortunately, the data did not allow us to explore these hypotheses further.

Fidelity during the first stage of SPM implementation varied and in general was modest. The use of key SPM techniques was moderate only, and the number of cases discussed in a team meeting was very low. Based on the survey responses from 23 sites, the social workers had, on average, brought only one case per month for case discussion. In comparison, in Hackney, each case of each social

worker was discussed once a month in a team meeting (Forrester et al., 2013). In England, in the sample examined by Bostock et al. (2019), 9 cases on average had been discussed in one session in sites implementing RSW, whereas Finnish sites reported that only one or two cases can fit into a weekly meeting. Moreover, in several sites these meeting have been held less often than once a week: according to a survey conducted by THL in 2021, of the 220 systemic teams around Finland, 41% held systemic meetings less often than once a week (Yliruka & Tasala, 2022). This survey did not ask for team size or caseloads. When the preliminary findings of this dissertation, have been presented for the practitioners and managers working in SPM teams, they have reported that their team sizes were still too large and that social workers rarely have a caseload of 35 cases, as has been the national recommendation since January 2022 (amendment to the Child Welfare Act 417/2007). Hence, even in theory, only a small portion of these cases can be discussed in systemic meetings. If this remains the case, it would seem unlikely that the SPM can have as much leverage as the RSW in the quality of child protection on the team level as previously reported (Bostock et al., 2019a; Forrester et al., 2013). Nevertheless, it is possible that the SPM has a beneficial effect on the few cases that have been discussed in systemic weekly meetings or where the practitioner has applied other systemic tools and methods. However, according to the pilot study, controlling for the use of key SPM methods on the case level did not change the results, i.e., no differences were found between cases where systemic methods had been applied compared to cases in the SPM group where these had not been applied or cases of service as usual. This finding could suggest that the use of SPM methods did not improve outcomes. However, given the difficulties encountered in implementing the SPM in general, it is possible that SPM methods were not used as skillfully during the first stage of implementation as had been hoped. To assess the quality of direct practice, observational data would be needed.

The Finnish SPM had an additional feature that was not part of the original RSW model, i.e., that of inviting families to be present in systemic team meetings. To date, no further research on service users' experiences of these meetings and their outcomes has been conducted, although positive experiences have been reported in THL working papers (Petrelius & Uutela, 2020; Ryynänen & Heinola, 2021). Moreover, the programme theory that was formulated as part of this dissertation should be updated. At the time of implementation, inviting families to team meetings was left out of the national description, and for this reason no detailed description of its delivery, intended outcomes or expected mechanisms was formulated.

7.2 Conclusions

The overall conclusion on the effectiveness of the SPM is that, during its first stage of implementation, it did not outperform conventional social work practice in any of the selected outcome measures. It did not result in improved family-

level outcomes or higher service user satisfaction with services. Moreover, agreement between practitioners and service users was not higher in cases supported by SPM teams than in cases supported by conventional teams. The lack of differences between teams in service outcomes needs to be interpreted in the light of the fidelity analysis: the content, dose and coverage of SPM was in general low. Some of the resources provided by the SPM were unfinished (training), or cut short (clinician and coordinator), or poorly implemented (large team size), and these were introduced in a context that did not support implementation (high caseloads, lack of managerial support, organisational reforms). As a result, on average, practitioners' reasoning (systemic thinking) did not substantially change, and their practice did not turn into systemic practice. Further, while the implementation of the SPM resulted in more intense casework, the increased volume of meetings did not change service user outcomes. In sum, when it comes to practice, quantity does not make up for quality. To improve implementation quality, stakeholders need a clear vision of the key components, intended outcomes and expected mechanisms of the programme, and an understanding of the context required for the mechanisms to operate successfully, i.e., a programme theory. Moreover, without such a theory, it is difficult to evaluate whether a programme has achieved its outcomes, and why. Evaluation should be a continuous process. After the programme theory has been formulated, it should be tested with new data, and refined based on the findings. As Bhaskar (1978) suggests, "Theory without experiment is empty. Experiment without theory is blind" (p., 191).

Despite the many challenges of using research evidence in daily practice, the social work profession has an ethical responsibility to critically investigate the intended and unintended consequences of social work interventions. All human services should do their best to ensure that practitioners do more good than harm. The history of health care and social care has witnessed interventions that despite their seemingly good intentions have caused serious harm (Gambrill, 2010; Morago, 2006). In Finland, for example, Hytönen et al. (2016) interviewed nearly 300 people who had been placed in out-of-home care as children during the years 1937-1983 and discovered that experiences of neglect and violence during placements were frequent. These findings led to a public apology by the Ministry of Social Affairs and Health and to reforms in social work practice and surveillance. Despite good intentions, children continue to report receiving violence and maltreatment in residential care (Koivula et al., 2021; Mononen et al., 2021). Policy reforms, legislation or more specific interventions do not necessarily have enough leverage or the right design to produce the desired outcomes. To understand if a programme is effective (or not), why, and how to improve it, we need evaluation research.

The findings of my dissertation illustrate how limited our current knowledge is on the effectiveness of child protection interventions. Previous research has primarily focused on the meso-level, such as out-of-home care (e.g., Kääriälä et al., 2018; Kääriälä & Hiilamo, 2017; McKenna et al., 2021). Attempts to understand micro-level mechanisms and test their operation and short-term

outcomes remain scarce, as illustrated by the findings of systematic reviews on child protection interventions (Dijkstra et al., 2016; Isokuortti et al., 2020; Sheehan et al., 2018). Small-scale attempts, such as this, cannot fill such an enormous gap in knowledge. It is, nonetheless, a small step forward in introducing and applying such methods in the field of social work research.

7.3 Implications for social work evaluation research

The lack of an experimental research tradition in social work has led to a situation in which researchers are not involved in the implementation of new interventions in a way that would lead to robust study designs. Currently, the initiative for developing services and models comes from practitioners and perhaps even more often from 'practice developers', i.e., experts in service development. The inner logic of 'practice development' is not compatible with experimental or quasi-experimental study designs. First, developers often find the idea of implementing highly specified interventions uncomfortable. It may be that social work is seen as a practice that needs autonomy and flexibility – it may even be considered as 'art' (Gitterman & Knight, 2018). On this thinking, the idea of controlling the application of a novel intervention is impossible. Second, the implementation of the new service models is based on the willingness of practitioners and service users. While this makes perfect sense for developers, for the researcher, this makes it difficult to randomise the research participants. As a result, the content of the intervention is random, while the sample is biased. In the future, more negotiation and collaboration are needed between researchers, developers and local organisations implementing novel interventions. If we want stronger evidence on the effectiveness of social work interventions, their implementation cannot be as free and creative as in the past (see also Baginsky et al., 2021).

Second, the challenge of deciding the key outcomes for social work interventions and child protection (Forrester, 2017) needs to be addressed. As for the most part social work research has been qualitative, researchers have not needed to operationalise outcome measures. In children's social care, documentation is primarily qualitative and narrative. Hence, information recorded in the client information systems does not provide quantitative data on a child's situation and well-being but only on activities and administrative decisions. However, this might change in the future due to national- (Kanta services) and area-level projects (e.g., Apotti) aimed at developing the client information systems for social care. Developing feasible measures for use in the child protection context and validating these would benefit future research. Ideally, a selection of valid and feasible measures would be in routine use in the field and scores would be recorded in a client information system. In this way, practitioners and service users would not have the bother of additional data collection, and the evaluation of new interventions could be performed with fewer resources.

Third, experience and expertise in the research methods needed in experimental or quasi-experimental research and process evaluation are scarce in social

work research. As suggested by Olsson and Sundell (2016), researchers should produce scientific evidence that is of interest in and applicable to practice. To do this, discipline-specific education is required that is not only grounded in professional practice but also prepares future researchers in how to apply evaluation methods that meet high quality research standards. To give an example, the UK Medical Research Guidance (Skivington et al., 2021) for developing and evaluating complex interventions could be beneficial for the purposes of evaluating social work interventions. While social work professionals often claim that healthcare interventions are more straightforward compared to those in social work, not all healthcare interventions or programmes are simple. Both of these service sectors operate around complex human needs and human behaviour, and have developed complex interventions for these purposes. A complex intervention refers to an intervention that has all or some of the following features: several interacting components, intention to affect participants' behavior in several ways, intention to affect several groups or levels in the organization, has several intended outcomes, and allows for flexibility and tailoring of the intervention (Craig et al., 2008). To produce useful information on such interventions, various study designs and methods can and should be applied. According to Skivington et al. (2021), complex intervention research can be viewed from four perspectives: efficacy, effectiveness, theory based and systems. The first two perspectives seek to explore the extent to which the intervention produces the intended outcomes in ideal (efficacy) or real-world (effectiveness) settings. The theory-based perspective aims to understand – like realist evaluation - what works, in what circumstances, and why. The systems perspective seeks to understand how a system and intervention adapt to each other. These perspectives are overlapping and can be used in combination. The core elements when evaluating complex interventions are stakeholder engagement, and developing, refining and retesting the programme theory (Skivington et al., 2021).

7.4 Implications for social work programme designs

A detailed description of an intervention is a prerequisite for meaningful evaluation and quality implementation. Both researchers and participants play an active role in producing such a description, e.g., a programme theory. As suggested by Chen (1990), an ideal description, or a normative theory helps organisations and practitioners to identify the changes needed in their service provision. In addition, Urk et al. (2016) suggest that theory-based interventions may be more effective than non-theory-based interventions, as they may result in a better understanding of human behaviour. The underlying theory can reveal weak links and potential conflicts in the hypothesised changes, as illustrated in this project. It is also crucial to take into account the possible unintended and even harmful outcomes a programme might have. These are often insufficiently theorised or empirically explored (Evans et al., 2015). In the case of the SPM, one fifth of the practitioners had experienced the SMP implementation process as a negative burden.

As Bonell et al. (2015) suggest, a programme theory addressing the resources and context needed to conduct a programme successfully could prevent such adverse effects.

Based on previous research, quality implementation requires training and support for practitioners and those designing the local implementation of an intervention prior its execution, and ongoing coaching during its implementation (Meyers et al., 2012). In the case of the SPM, the unclear design of the intervention, incoherent training material, and lack of a manual undermined its implementation and hence outcomes. In the light of these findings, the idea of co-producing new interventions in an on-going process should be re-evaluated. Although this may be justified in some cases, in others where the purpose is to disseminate a model which has specific mechanisms that should operate to achieve predefined outcomes, another strategy is needed. Moreover, if a model is being developed in an on-going process of co-production, it might be almost impossible for the researchers to keep up with the inputs and intended outputs let alone succeed in the data collection (Baginsky et al., 2021).

7.5 Implications for social work practice

For organisations planning SPM implementation, it is important to acknowledge that no evidence yet exists that is it more effective than service as usual. This is not to say that it could not be more effective if fully implemented, but at the present juncture this is only a hypothesis. On the other hand, several factors that may have a negative effect on the quality of SPM implementation have been identified that should be considered before deciding whether to go ahead with it. In addition, no research thus far has analysed the innovation included in the SPM, i.e., that of inviting families to weekly meetings. Before the large-scale dissemination of this new core component, it would be important to examine under what conditions such a meeting is beneficial, for which families, and what are the intended outcomes and mechanisms of this component.

In general, if an organisation wishes to use a social work intervention that is not yet evidence-based, its implementation should be done with care and on a small scale. The organisation's managers or other professionals responsible for organizing practice need to monitor the intervention to see whether brings more benefits than harm. Intended and un-intended outcomes should be evaluated from the perspective of both service users and practitioners. To do this, information should be collected from multiple sources from various levels - not just from those responsible for training or dissemination, or from most enthusiastic supporters of the intervention.

The use of valid and feasible measures could serve everyday practice in general as a way of monitoring change on the case level, and especially when implementing new interventions without external research resources. In Finland, the client information systems currently used in social work do not primarily involve structured assessment tools or categorised data on the factors related to the

need for child protection and child wellbeing. While such information could be derived from case files, taking the challenges of the existing client information systems into account (Salovaara & Ylönen, 2021) and quantifying narrative information in a comparable form would require significant time and effort.

Sundell et al. (2010) characterised Swedish social work as an opinion-based rather than evidence-based practice. This argument can also be applied in the Finnish context. In their review, Gray et al. (2013) identified the barriers to implementing evidence-based practice. These included lack of time and access to the research literature, lack of knowledge and skills to search and apply research evidence, poor applicability of the findings, negative attitudes, and inadequate support. However, none of this matters if relevant primary research on child protection interventions and approaches is scarce or nonexistent.

From the service user perspective, services are not currently based on knowledge gained from research – at least not in any clear or explicit manner. While this would be regarded as a major ethical problem in health care, it seems not to be the case in social work. Some scholars see social work practice as an artistic approach to engaging and helping clients. Social workers should have the ability to "go with the flow", be informal and spontaneous (Gitterman & Knight, 2018). In this view, evidence-based practice is seen as restricting practitioners. However, as argued by Gambrill (2010), good intentions are not enough to ensure good practice. Moreover, if new innovations are developed and disseminated to provide social work practitioners with a new, commonly shared set of skills and knowledge, they are clearly intended to restrict practitioners in certain ways. If not, they would not be 'a model', 'a practice model' or 'a shared approach'. If an agency wishes to disseminate something that aims to change the way organisations and practitioners operate, this is an intervention in the current state-of-the-art. It must, of course, be acknowledged that such interventions do not always work as intended; they may not always be transferable to a new context, or they might have unintended consequences (Pawson & Tilley, 1997). This has been found over and over again in previous research and is one of the arguments for promoting evidence-based practice (Morago, 2006). Outcome and process evaluation research has established methods and frameworks to critically address these issues to benefit the future development and application of social work interventions.

YHTEENVETO (SUMMARY IN FINNISH)

Tämän väitöstutkimuksen tavoitteena on tutkia *systeemisen lastensuojelun toimintamallin* (myöh. systeeminen malli) vaikuttavuutta realistisen arvioinnin viitekehyksessä. Systeemisessä mallissa yhdistetään systeemisen perheterapian orientaatioita, menetelmiä ja välineitä lasten ja perheiden sosiaalityöhön. Realistisessa arvioinnissa pyrkimyksenä on ymmärtää mikä toimii, kenelle, missä olosuhteissa, ja miksi. Arviointia ohjaa ohjelmateoria, joka on kuvaus siitä, miten tutkittavan intervention odotetaan toimivan. Tutkimukseni kolme päätutkimuskysymystä noudattavat realistisen arvioinnin tutkimustehtävää. Tutkimuksen käynnistymisvaiheessa systeemiselle mallille muotoiltiin logiikkamalli, joka toimi aineistonkeruuta ja analyyseja ohjaavana alustavana ohjelmateoriana. Tutkimuksen ensimmäinen tutkimuskysymys koski systeemisen mallin vaikuttavuutta, jota olen arvioinut vertailemalla systeemisen työskentelyn asiakastason tuloksia ja työskentelyn laatua tavanomaisen lastensuojelun sosiaalityön tuloksiin ja laatuun. Tutkimuksen toinen tutkimuskysymys koski systeemisen mallin käyttöönottoa ja sitä, miten käyttökonteksti on vaikuttanut mallin toimivuuteen ja havaittuihin tuloksiin. Lopuksi systeemiselle mallille muotoiltiin tarkennettu ohjelmateoria, jossa on kuvattu mallin ydinelementit, toivotut tulokset, ja oletukset siitä, millä mekanismeilla ja missä olosuhteissa nämä tulokset on mahdollista saavuttaa.

Reclaiming Social Work -malli ja aiempi tutkimus

Systeeminen lastensuojelun toimintamalli on suomalainen sovellus *Reclaiming Social Work* -mallista (RSW). Molempien mallien alkuperäisenä ideana on ollut sosiaalityön johtamisen ja organisaation muokkaaminen siten, että sosiaalityöntekijät ja muut asiakastyöhön osallistuvat työntekijät voivat tehdä systeemistä asiakastyötä. Systeemisessä asiakastyössä hyödynnetään systeemisen perheterapian tapaa ymmärtää perheenjäsenten välisiä suhteita ja tarkastellaan perhesysteemin suhdetta muihin systeemeihin työntekijä ja palvelujärjestelmä mukaan lukien. Ongelmien ei nähdä paikantuvan yksilöihin, vaan yksilöiden ja systeemien välisiin suhteisiin, jotka täten asetetaan työskentelyn kohteeksi. Systeemisen työskentelyn tueksi sosiaalityöntekijöistä muodostetaan systeemisiä tiimejä, joihin kuuluu tiimin asiakastyötä ohjaava konsultoiva sosiaalityöntekijä, systeemistä näkökulmaa ja menetelmäosaamista vahvistava systeeminen perheterapeutti eli kliinikko, käytännön asioita hoitava koordinaattori, ja perhetyöntekijä. Systeeminen tiimi pitää systeemisiä viikkokokouksia, joissa tarkastellaan asiakastapauksia yhteisesti reflektoiden systeemisiä menetelmiä ja välineitä, kuten sukupuuta ja hypoteesityöskentelyä, hyödyntäen.

RSW-malli kehitettiin alun perin Hackneyn alueen sosiaalitoimistossa Lontoossa, missä se otettiin käyttöön 2007. Hackenyä koskevissa arviointitutkimuksissa RSW-mallin on todettu parantaneen asiakastyön laatua ja työntekijöiden työhyvinvointia. Tarkkaa tietoa mallin levinneisyydestä ei ole, mutta julkaistujen arviointitutkimusten perusteella mallia on Briteissä sovellettu ainakin neljällä

uudella paikkakunnalla ja kolmessa Lontoon kaupunginosassa. Tutkimusten perusteella mallin käyttöönotto ei ole uusissa organisaatioissa useimmiten onnistunut tarkoitetulla tavalla. Organisaatioissa ei ole välttämättä toteutettu tarvittavia muutoksia systeemisen työskentelyn mahdollistamiseksi. Liian suuret asiakasmäärät ja tiimit sekä organisaation toimintaa ohjaavan yhteisen arvoperustan puute ovat estäneet RSW-mallin ideoiden toteutumista. Kaikkiin tiimeihin ei ole rekrytoitu perheterapeuttia, mutta juuri perheterapeutin on havaittu olevan yhteydessä laadukkaaseen systeemiseen case-keskusteluun, minkä puolestaan on osoitettu olevan yhteydessä parempaan asiakastyön laatuun. Näistä myönteisistä tutkimustuloksista huolimatta RSW-mallin vaikuttavuudesta ei kuitenkaan ole sellaista tutkimusnäyttöä, jonka perusteella voitaisiin sanoa, että mallia soveltavissa tiimeissä tai organisaatioissa olisi saatu aikaan parempia asiakastyön tuloksia ei-systeemiseen työskentelyyn verrattuna.

Suomalainen systeeminen lastensuojelun toimintamalli

Väitöstutkimukseni kohteena oleva suomalainen systeeminen lastensuojelun toimintamalli lähti liikkeelle RSW-mallin perusideoista, muttei pyrkinyt noudattamaan mallia sellaisenaan. Systeemisen mallin valtakunnallinen kehittäminen käynnistyi 2016 Terveyden ja hyvinvoinnin laitoksen (THL) vetämässä työpajaprosessissa osana Lasten ja perheiden muutosohjelman (LAPE) ensimmäistä vaihetta vuosina 2016–2018. Vuonna 2017 mallin levittämistä varten THL julkaisi mallin kuvauksen ja käynnisti kouluttajakoulutuksen, johon osallistui yhteensä 58 perheterapeuttia tai sosiaalityöntekijää. Heistä muodostetut kouluttajaparit vastasivat paikallisista tiimeittäin toteutetuista koulutuksista, joissa oli koulutettu kesään 2018 mennessä yhteensä 53 tiimiä kaikkiaan 31 paikkakunnalla. Nykyisin THL osallistuu systeemisen mallin levittämiseen esimerkiksi järjestämällä kouluttajakoulutusta, ylläpitämällä kouluttajapankkia ja koulutusmateriaalia, järjestämällä verkostotapaamisia ja koordinoimalla mallin kansallisesta kehittämisestä vastaavaa ohjausryhmää. Tämän lisäksi mallin sisällöllistä kehittämistä ja käyttöönoton vakiinnuttamista on tehty vuosina 2020–2022 sosiaali- ja terveysministeriön rahoittamissa alueellisissa lastensuojelun kehittämishankkeissa.

Keskeisimpiä eroja RSW-mallin ja suomalaisen mallin välillä on ensinnäkin se, että systeeminen tiimi on alkuperäistä RSW-mallin systeemistä yksikköä isompi. Suomalaisessa mallissa tiimin kokoonpanoon voi kuulua lastensuojeluorganisaation pysyvien työntekijöiden lisäksi asiakastapauksittain vaihtuvia jäseniä, kuten lapsen ja perheen verkostoon kuuluvia työntekijöitä. Lisäksi suomalaisessa mallissa työntekijöitä on kehotettu kutsumaan asiakasperheitä mukaan systeemisen tiimin viikkokokouksiin, jolloin tapaamisesta tulee asiakas- tai verkostotapaaminen siinä missä RSW-mallissa viikkokokoukset ovat toimineet työntekijöiden keskinäisinä asiakastyön ohjauksen foorumeina. Systeemisen mallin koulutukseen valitut menetelmät ovat myös eronneet RSW-mallissa sovelletuista menetelmistä. Lisäksi mallin levittämisstrategiat ovat eronneet toisistaan työntekijöille tarjotun koulutuksen ja organisaation johdolle tarjotun muu-

tostuen määrän ja toteutustavan osalta. Nämä erot on otettava huomioon mietittäessä RSW-mallia koskevien tutkimustulosten siirrettävyyttä suomalaiseen kontekstiin.

Aineistonkeruuprosessi

Olen kerännyt systeemisen mallin vaikuttavuutta ja käyttöönottoa koskevat aineistoni THL:n tutkimushankkeessa, jossa vastasin systeemisen mallin kansallisesta arvioinnista maaliskuusta 2017 helmikuuhun 2019. THL:n tutkimushanke eteni rinnakkain systeemisen mallin kehittämisestä ja levittämisestä vastanneen hankkeen kanssa. Tutkimus- ja kehittämishankkeet olivat tiiviissä vuoropuhelussa siten, että tutkijana minun oli mahdollista osallistua mallin kouluttajakoulutukseen ja kehittämistilaisuuksiin saadakseni paremman käsityksen mallin sisällöistä ja tavoitteista, ja toisaalta välittää aiempaa tutkimustietoa ja arviointitutkimuksessa muodostunutta tietoa ja tilannekuvaa mallin kehittämishankkeen hyödynnettäväksi. Lisäksi tutkimus- ja kehittämishankkeet järjestivät yhteisiä työpajoja mallin paikallisesta toimeenpanosta vastanneille esihenkilöille. Kansallisen arvioinnin myötä havaittiin, että mallin kouluttajat ja työntekijät toivoivat mallin keskeisten ideoiden ja käsitteiden tarkempaa avaamista ja konkretisoimista. Tämän vuoksi vuonna 2019 kerättiin vielä uusi työpaja-aineisto, jonka avulla laadittiin tarkennettu kuvaus mallin ydinelementeistä ja ohjelmateoriasta. Tutkijoiden vetämä työpajaprosessi toteutettiin yhteistyössä THL:n kanssa.

Tutkimustehtävät ja tutkimusaineisto

Väitöstutkimuksen kahdessa ensimmäisessä osatutkimuksessa tarkastellaan systeemisen mallin vaikuttavuutta. Nämä osatutkimukset ovat kvasikokeellisia, eli niissä on käytetty vertailuasetelmaa, mutta vertailuryhmiä muodostettaessa tutkittavia ei ole satunnaistettu. Aineisto kerättiin yhdeksästä systeemistä mallia pilotoivasta tiimistä (koeryhmä n=56 asiakastapausta) ja yhdeksästä tavanomaisesta tiimistä (kontrolliryhmä n=56 asiakastapausta) vuoden 2018 aikana kolmelta paikkakunnalta, jotka osallistuivat mallin pilotointiin. Vaikuttavuusaineisto muodostui lapsilta, vanhemmilta ja työntekijöiltä kerätystä lomakeaineistosta, jolla kartoitettiin lapsen ja perheen hyvinvointia ja lastensuojelutarvetta kahdessa mittauspisteessä noin puolen vuoden seurantajaksolla. Jälkimmäisessä aineistonkeruussa kartoitettiin myös seurantajakson aikaista työskentelyä ja lapsen ja vanhemman siitä antamaa palautetta. Lomakkeisiin oli sisällytetty kaksi validoitua mittaria (lapsen subjektiivista hyvinvointia kartoittava KINDL-R ja perhedynamiikkaa kartoittava SCORE-15) sekä tätä tutkimusta varten laadittuja kysymyksiä koskien lastensuojelun tarvetta, lapsen kohtelussa havaittuja puutteita, lapsen ja vanhempien hyvinvointiin vaikuttavia tekijöitä sekä seurantajakson aikaista työskentelyä. Ensimmäisessä artikkelissa vertailtiin koe- ja kontrolliryhmässä tapahtuneita muutoksia lapsen hyvinvoinnissa, perhedynamiikassa, lapsen kohtelussa ja lastensuojelutarpeessa sekä ryhmien välisiä eroja työskentelyn laadussa (lapsen ja vanhempien palaute) ja määrässä (asiakastapaamisten

määrä). Toisessa artikkelissa tarkasteltiin lapsen, vanhempien ja työntekijän välisten näkemysten yhtenevyyttä koe- ja kontrolliryhmien välillä sekä lähtötilanteessa havaitun yksi- tai erimielisyyden vaikutusta lapsen ja perheen tilanteessa seurantajakson aikana tapahtuneisiin muutoksiin.

Väitöstutkimuksen kolmannessa osatutkimuksessa tarkastellaan systeemisen mallin fideliteettiä ja siihen vaikuttaneita tekijöitä. Fideliteettiä tarkasteltaessa arvioitiin sitä, kuinka pitkälle mallin paikallinen toimeenpano vastasi mallin kansallisesta kuvauksesta ja koulutussisällöistä johdettuja sisällöllisiä ja määrällisiä tavoitteita. Osatutkimus oli monimenetelmällinen, ja siinä kerättiin määrällistä ja laadullista aineistoa rinnakkain. Aineistot analysoitiin erikseen, minkä jälkeen analyysin tuloksia tulkittiin yhdessä. Osatutkimusta varten keräsin vuoden 2018 syksyllä kyselyaineiston mallin ensivaiheen pilotointiin osallistuneilta 23 paikkakunnalta. Kyselyllä kartoitettiin sosiaalityöntekijöiden kokemuksia mallin koulutuksesta, käyttöönotosta ja työhyvinvoinnista sekä asiakasmääriä. Kysely lähetettiin kaikille tutkimuspaikkakunnilla pilotointiin osallistuneille sosiaalityöntekijöille, ja siihen vastasi 56 työntekijää (vastausprosentti 44 %). Kyselyaineiston lisäksi keräsin keväällä 2018 mallin paikallisesta toimeenpanosta vastanneilta esihenkilöiltä tietoa pilottitiimien kokoonpanosta ja toimenkuvasta.

Väitöstutkimuksen neljännessä osatutkimuksessa on muotoiltu tarkennettu kuvaus systeemisen mallin ydinelementeistä ja ohjelmateoriasta. Aineisto kerättiin tutkijoiden vetämässä työpajaprosessissa, johon kutsuttiin systeemisen mallin kansalliseen kehittämiseen, kouluttamiseen ja levittämiseen vuosina 2017–2019 osallistuneita avainhenkilöitä (n=12). Tutkimusta varten kokoonnuttiin viiteen työpajaan, joissa osallistujia pyydettiin määrittelemään ja perustelemaan mallin ydinelementit, keskeiset käsitteet, tavoitteet ja mallin tarjoamat resurssit, sekä muotoilemaan hypoteeseja mallin toimintamekanismeista, vaikutusketjuista ja ihanteellisesta toimintaympäristöstä. Lisäksi tutkijat osallistuivat työpajan osallistujien rinnalla THL:n järjestämään viisipäiväiseen RSW-koulutukseen, toteuttivat kaksi haastattelua ja hyödynsivät muuta täydentävää kirjallista aineistoa.

Tulokset

Vaikuttavuutta koskevien osatutkimuksen valossa systeeminen malli ei tuottanut tavanomaista työskentelyä parempia tuloksia. Ensimmäisessä osatutkimuksessa testattiin hypoteesia, jonka mukaan systeeminen malli parantaa perheen sisäisiä vuorovaikutussuhteita ja vähentää lapsen kaltoinkohtelua ja laiminlyöntiä, minkä ansiosta lastensuojelun tarve vähenee ja lapsen koettu hyvinvointi paranee. Tässä tutkimuksessa ei havaittu eroja koe- ja kontrolliryhmän välillä näissä muuttujissa tapahtuneissa muutoksissa. Toisin sanoen, systeeminen työskentely ei ollut tällä aineistolla tarkasteltuna tavanomaista työskentelyä vaikuttavampaa. Toisessa osatutkimuksessa testattiin hypoteesia, jonka mukaan lasten, vanhempien ja työntekijöiden näkemys perheen tilanteesta olisi systeemisissä tiimeissä tavanomaisia tiimejä lähempänä toisiaan, koska systeemisen tiimin työntekijät saivat koulutusta ja tukea reflektoidakseen tilannetta perheiden kanssa ja

lisätäkseen heidän osallisuuttaan työskentelyssä. Myöskään tämä hypoteesi ei saanut tällä aineistolla vahvistusta. Tarkasteltaessa asiakastyön intensiteettiä kuitenkin havaittiin, että systeemisissä tiimeissä asiakastapaamisten määrä oli kontrolliryhmään verrattuna selvästi suurempi. Systeemisissä tiimeissä sosiaalityöntekijä oli puolen vuoden seurantajakson aikana tavannut lasta kahden kesken keskimäärin kahdesti, kun kontrollitiimeissä näin oli tapahtunut keskimäärin vain kerran. Tutkimukseen osallistuneissa perheissä seurantajakson aikainen asiakastapaamisten kokonaismäärä oli systeemisissä tiimeissä keskimäärin kolminkertainen kontrollitiimeihin verrattuna kun huomioidaan kaikki sosiaalityöntekijän, sosiaaliohjaajan, lapsen ja/tai vanhemman tapaamiset millä tahansa kokoonpanolla. Asiakastyön korkeampi intensiteetti ei kuitenkaan näkynyt lapsen tai vanhemman suurempana tyytyväisyytenä lastensuojelun työskentelyyn. Lapsen tai vanhemman kokemus osallisuudesta työskentelyssä tai luottamus ja yleinen tyytyväisyys lastensuojelun työskentelyyn olivat seurantahetkellä yhtä hyvällä tasolla systeemisissä ja tavanomaisissa lastensuojelutiimeissä.

Mallin implementointia tarkastelleessa tutkimuksessa paikannettiin useita tekijöitä, jotka estivät työntekijöitä soveltamasta toimintamallia ja systeemistä asiakastyötä tarkoitetulla tavalla. Näitä olivat systeemisen työotteen epäselvyys, koulutuksen riittämättömyys ja organisaatioon liittyvät tekijät, kuten korkeat asiakasmäärät, työntekijöiden suuri vaihtuvuus ja johdon tarjoaman tuen puute. Vaikka asiakastapaamisia oli systeemisissä tiimeissä merkittävästi enemmän, työskentelyn sisältö ei muuttunut tarkoitetulla tavalla. Systeemisiä menetelmiä, kuten sukupuuta ja hypoteesien muotoilua, oli sovellettu vähäisesti, ja vain hyvin pieni osa pilottitiimien asiakkaista oli tuotu systeemisen tiimin yhteiseen käsittelyyn systeemiseen viikkokokoukseen. Fideliteettiä tarkastelleen analyysin perusteella työntekijöiden ajattelussa oli pilotoinnin ensivaiheessa tapahtunut vaihtelevasti toivottuja muutoksia. Fideliteettiä kontrolloitiin asiakastason tuloksia koskevassa vaikuttavuusanalyysissa tarkastelemalla sitä, oliko työntekijä hyödyntänyt systeemisiä menetelmiä kussakin yksittäisessä asiakastapauksessa. Tämän huomioiminen ei kuitenkaan muuttanut tuloksia.

Työntekijöiden kokemukset systeemisestä mallista jakautuivat. Suurin osa työntekijöistä näki mallissa potentiaalia ja oli halukkaita jatkamaan mallin mukaista työskentelyä. Moni kuitenkin koki mallin keskeneräiseksi, ja arvioi paikallisen koulutuksen ja käyttöä tukevan materiaalin puutteelliseksi. Osa työntekijöistä turhautui paikallisen toimeenpanon heikkoon fideliteettiin, joka ei vastannut heidän odotuksiaan niistä muutoksista, joita asiakastyössä ja organisaatiossa oli odotettu tapahtuvaksi. Lähes puolet kyselyvastaajista oli kokenut mallin pilotoinnin kuormittavana. Osa heistä oli tästä huolimatta halukkaita jatkamaan mallin mukaista työskentelyä, kun taas viidennes vastaajista koki systeemisen mallin kuormittavana ja halusi luopua siitä. Työntekijöiden asiakasmäärät eivät olleet yhteydessä työntekijöiden kokemaan kuormitukseen tai halukkuuteen jatkaa.

Mallin jatkolevitystä ja -tutkimusta varten mallille muodostettiin lopuksi ohjelmateoria ja kuvaus mallin ydinelementeistä yhdessä mallin kansalliseen ke-

hittämiseen ja levittämiseen osallistuneiden avainhenkilöiden kanssa. Osatutkimuksessa mallin ydinelementeiksi määriteltiin systeeminen tiimi, systeeminen viikkokokous ja systeeminen asiakastyö. Työpajatyöskentelyn aikana tarkennettiin systeemiseen tiimiin kuuluvien sosiaalityöntekijöiden määräksi vähintään kaksi, jotta he voivat tarjota toisilleen vertaistukea mutta korkeintaan kolme, jotta tiimissä on riittävästi aikaa ja tilaa tiimin yhteisille keskusteluille. Systeeminen viikkokokous määriteltiin työntekijöiden keskinäisen reflektiivisen keskustelun foorumiksi. Systeemisen asiakastyön tavoitteita, periaatteita ja menetelmiä konkretisoitiin ja täsmennettiin. Työpajaprosessissa nousi myös esiin kilpailevia näkemyksiä mallin ydinelementtien sisällöstä ja funktiosta, joista käytiin keskusteluja ja neuvotteluja. Erityisesti systeemisen viikkokokouksen funktio nähtiin eri tavoin: osa kannatti sen rajaamista työntekijöiden keskinäiseksi oppimisen ja asiakastyön ohjauksen foorumiksi, kun osa kannatti asiakkaiden ja verkoston jäsenten kutsumista näihin kokouksiin. Lopulta viikkokokous päädyttiin rajaamaan työntekijöille. Aineiston pohjalta tutkijat muotoilivat kuvauksen siitä, miten systeemisen mallin tarjoamat resurssit määrätynlaisessa kontekstissa muokkaavat työntekijöiden ja asiakkaiden järkeilyä ja toimintaa, ja saavat aikaan määriteltyjä tuloksia.

Systeemisen ja tavanomaisen työskentelyn erojen tarkastelun lisäksi väitöstutkimuksen kaksi ensimmäistä osatutkimusta tuottivat koko aineistoa koskevia tuloksia. Tutkimuksessa ensinnäkin havaittiin, että työntekijöiden arvioima lastensuojelun tarve väheni ja äitien arvio perhedynamiikasta parani seurantajakson aikana. Lasten kokemus perhedynamiikasta ei sen sijaan muuttunut, eivätkä havaitut muutokset heijastuneet vastaavaan muutokseen lapsen koetussa hyvinvoinnissa tai lasten ja vanhempien arvioissa ulkopuolisen avun tarpeesta. Näkemysten yhteneväisyyttä tarkasteltaessa havaittiin, ettei työntekijän näkemys lastensuojelutarpeesta korreloinut lapsen ja äidin näkemyksen kanssa. Lapsen ja äidin yhtenevä näkemys ulkopuolisen avun tarpeesta lähtötilanteessa ennusti työntekijän arvioiman lapsen kaltoinkohtelun ja laiminlyönnin vähenemistä.

Johtopäätökset

Väitöstutkimukseni on tuonut esiin, ettei systeemisen mallin vaikuttavuudesta ole vielä tutkimusnäyttöä mallin suosiosta huolimatta. Sama tutkimusnäytön vähäisyys koskee myös muita lastensuojelun toimintamalleja ja valtaosaa lasten ja perheiden sosiaalityössä sovellettuja menetelmiä ja välineitä. Systeemisen mallin käyttöönottoa harkitsevien organisaatioiden on hyvä huomioida, että mallin vaikuttavuutta koskevan puuttuvan näytön lisäksi mallin käyttöönottoon liittyy useampia haasteita, joita tässä tutkimuksessa on eritelty tarkemmin.

On mahdollista, että systeeminen asiakastyö on tavanomaista työskentelyä vaikuttavampaa, mikäli systeemisen mallin toimeenpanossa onnistutaan nyt tutkimuksen kohteena ollutta ensivaihetta paremmin. Tämän tutkiminen edellyttää uuden aineiston keruuta organisaatioissa, joissa malli on mahdollisimman täysimääräisesti käytössä ja näissä tehdyn asiakastyön laadun ja tulosten vertailua tavanomaiseen työskentelyyn. Jotta systeemisen työskentelyn sisällöstä, määrästä

ja kattavuudesta saadaan selkeä kuva, myös jatkotutkimuksissa olisi tärkeä arvioida mallin käytön fideliteettiä. Kysely- ja haastatteluaineistojen lisäksi fideliteettiä olisi hyvä tarkastella tiimikokouksista ja asiakastapaamisista kerätyn havainnointiaineiston valossa. Erityisen tärkeää olisi päivittää mallin ohjelmateoria ainakin siltä osin, että siinä kuvattaisiin miten ja miksi asiakkaan on tarkoitus olla läsnä systeemisessä tiimikokouksessa, ja millaisia vaikutuksia sillä uskotaan olevan työskentelyyn. Tämä elementti ei sisälly mallista laadittuun ohjelmateoriaan, eikä siitä ole tutkimusnäyttöä Iso-Britanniasta, koska se ei sisälly RSW-malliin.

Kuten tässä väitöstutkimuksessa on tuotu esiin, vaikuttavuuden ohella on tärkeää tutkia interventioiden käyttöönottoa, jotta voidaan arvioida käyttöönoton vaikutuksia tuloksiin ja toisaalta tulosten siirrettävyyttä uuteen kontekstiin. Sekä arviointitutkimusta että käyttöönottoa varten on hyvä muotoilla intervention sisältöä, tavoitteita, toivottuja toimintamekanismeja ja käyttöympäristöä kuvaava ohjelmateoria. Tällöin kaikille osapuolille ja käyttöä harkitseville on selvää, mitä interventiolla tavoitellaan ja millaisissa olosuhteissa sen voidaan odottaa tuottavan toivottuja tuloksia. Ohjelmateorian muotoilu auttaa myös intervention kehittäjiä havaitsemaan mahdolliset ristiriidat tai heikot kohdat omissa suunnitelmissaan. Sekä käyttöönotossa että arvioinnissa on toivottujen tulosten tarkastelun ohella otettava huomioon, että toisinaan interventioilla tai niiden puutteellisella käyttöönotolla voi olla myös ei-toivottuja tuloksia. Väitöstutkimukseni tuo esiin, millaisia haasteita kehitysvaiheessa olevan intervention käyttöönottoon ja arviointiin liittyy. Mallin hajautetulla kehittämisellä voi olla ei-toivottuja seurauksia, kuten epäselvyyttä mallin sisällöstä ja ristiriitaisia tulkintoja mallin ydinelementeistä. Tulosten valossa on syytä kriittisesti pohtia, mitä lastensuojelun kehittämistyöllä tavoitellaan, ja miten alueellista tai valtakunnallista kehittämistyötä on näihin tavoitteisiin nähden tarkoituksenmukaista jatkossa tehdä.

Näyttöön perustuvan sosiaalityön perusedellytys on, että sosiaalityön interventioista on tuotettu tutkimusnäyttöä. Toistaiseksi sellaista on niukasti. Tutkijan näkökulmasta näytön tuottamisen keskeinen haaste on tällaisen tutkimusperinteen puute. Sosiaalityön kentällä on tehty vähäisesti kokeellista tutkimusta, eikä uusia interventioita pääsääntöisesti pyritä kehittämään sellaisilla prosesseilla, joiden puitteissa niiden vaikuttavuutta olisi mahdollista arvioida. Vaikuttavuuden arvioinnissa onnistuminen edellyttää myös, että sosiaalityön tavoitteet on selkeästi määritelty, ja niiden vertailukelpoista arvioimista varten on kehitetty soveltuvia mittareita. Jatkossa kehittäjien, tutkijoiden ja paikallisten organisaatioiden olisi syytä tehdä parempaa yhteistyötä, jotta uusien menetelmien ja mallien hyödyistä voidaan tuottaa luotettavaa tutkimusnäyttöä. Tässä voidaan hyödyntää vaikuttavuuden arvioinnin ja prosessinarvioinnin piirissä kehitettyjä vakiintuneita tutkimusmenetelmiä ja kompleksisten interventioiden kehittämiseen ja arviointiin laadittuja erityisiä suosituksia.

REFERENCES

Aaltio, E. (2015). *Lastensuojelutarpeen muutoksen arviointi perheen toimintakykyä mittaamalla. Lastensuojelun vaikuttavuusmittari -projektin loppuraportti. [To assess the change in need for child protection by assessing the family's functioning. Outcome measure for child protection. Final report].* Kuntaliitto.

Aaltio, E., & Isokuortti, N. (2019). *Systeemisen lastensuojelun toimintamallin pilotointi. Valtakunnallinen arviointi [The initial implementation of the Systemic Practice Model for child protection. National evaluation]*. Report 3/2019. Finnish Institute for Health and Welfare.

Aaltio, E. & Isokuortti, N. (2021). Developing a programme theory for Systemic Practice Model in children's social care. Key informants' perspectives. *Child & Family Social Work*, 27(3), 444-453. https://doi.org/10.1111/cfs.12896

Aaltio, E. & Kannasoja, S. (accepted). Achieving agreement on service needs in child protection. Comparing children's, mothers', and practitioners' views over time and between approaches. *Journal of Children's Services*.

Allnock, D., & Atkinson, R. (2019) 'Snitches get stitches': School-specific barriers to victim disclosure and peer reporting of sexual harm committed by young people in school contexts. *Child Abuse & Neglect*, *89*, 7-17. https://doi.org/10.1016/j.chiabu.2018.12.025

Antle, B. F., Barbee, A. P., Christensen, D. N., & Martin, M. H. (2008). Solution-Based Case-work in Child Welfare: Preliminary Evaluation Research. *Journal of Public Child Welfare*, *2*(2), 197-227. https://doi.org/10.1080/15548730802312891.

Antle, B. F., Barbee, A. P., Christensen, D. N., & Sullivan, D. J. (2009). The prevention of child maltreatment recidivism through the Solution-Based Casework model of child welfare practice. *Children and Youth Services Review, 31*(12), 1346-1351. https://doi.org/10.1016/j.childyouth.2009.06.008.

Antle, B. F., Christensen, D. N., van Zyl, M. A., & Barbee, A. P. (2012). The impact of the Solu-tion Based Casework (SBC) practice model on federal outcomes in public child welfare. *Child Abuse & Neglect, 36*(4), 342-353. https://doi.org/10.1016/j.chiabu.2011.10.009.

Aggett, P., Swainson, M., & Tapsell, D. (2015). 'Seeking permission': An interviewing stance for finding connection with hard to reach families. *Journal of Family Therapy, 37*(2), 190–209. https://doi.org/10.1111/j.1467-6427.2011.00558.x

Anderson, H. (1997). *Conversation, language, and possibilities*. Basic Books.

Anttonen. A., & Sipilä, J. (2000). *Suomalaista sosiaalipolitiikkaa*. Vastapaino.

Archer, M. (1995). *Realist Social Theory: the Morphogenetic Approach.* Cambridge University Press.

Bandura, A. (1977). *Social Learning Theory*. Prentice Hall.

Baginsky, M., Hickman, B., Harris, J., Manthorpe, J., Sanders, M., O'Higgins, A., Schoenwald, E., & Clayton, V. (2020). *Evaluation of MTM's Signs of Safety Pilots. Evaluation report.* Department for Education.

Baginsky, M., Ixer, G., & Manthorpe, J. (2021). Practice Frameworks in Children's Services in England: An Attempt to Steer Social Work Back on Course? *Practice, 33*(1), 3-19. https://doi.org/10.1080/09503153.2019.1709634

Barbee, A. P., Christensen, D., Antle, B., Wandersman, A., & Cahn, K. (2011). Successful adoption and implementation of a comprehensive casework practice model in a public child welfare agency: Application of the Getting to Outcomes (GTO) model. *Children and Youth Services Review, 33*(5), 622–633. https://doi.org/10.1016/j.childyouth.2010.11.008.

Barlow, J., Johnston, I., Kendrick, D., Polnay, L., & Stewart-Brown, S. (2006). Individual and group-based parenting programmes for the treatment of physical child abuse and neglect. *Cochrane Database of Systematic Reviews, 19*(3). https://doi.org/10.1002/14651858.CD005463.pub2

Bateson, G. (1972). *Steps to an Ecology of Mind: Mind and Nature*. Ballantine Books.

Bateson, G. (1980). *Mind and Nature: a Necessary Unity*. Fonata/Collins.

Bellamy, J., Bledsoe, S., & Traube, D. (2006). The Current State of Evidence-Based Practice in Social Work A Review of the Literature and Qualitative Analysis of Expert Interviews. *Journal of Evidence-Based Social Work, 3*(1), 23-48. https://doi.org/10.1300/J394v03n01_02

Bergh, A. (2004). The Universal Welfare State: Theory and the Case of Sweden. *Political Studies*, *52*, 745-766. https://doi.org/10.1111/j.1467-9248.2004.00506.x

Berrick J.D., Dickens J., Pösö T., & Skivenes M. (2015). Children's involvement in care order decision-making: A cross-country analysis. *Child Abuse & Neglect, 49*, 128–141. https://doi.org/10.1016/j.chiabu.2015.07.001

Berrick, J., Dickens, J., Pösö, T., & Skivenes, M. (2016). Time, Institutional Support, and Quality of Decision Making in Child Protection: A Cross-Country Analysis. *Human Service Organizations: Management, Leadership & Governance, 40*(5), 451–468. https://doi.org/10.1080/23303131.2016.1159637.

Beveridge, W. (1942). *Social Insurance and Allied Services*.

Bhaskar, R. (1997 [1975]). *A Realist Theory of Science*. Verso Books.

Bhaskar, R. (1986). *Scientific Realism & Human Emancipation*. Verso Books.

Blom, B., & Morén, S. (2010). Explaining Social Work Practice – The CAIMeR Theory. *Journal of Social Work, 10*(1), 98-119. https://doi.org/10.1177/1468017309350661

Blunt, C. (2015). *Hierarchies of evidence in evidence-based medicine*. A Thesis. London School of Evonomics and Political Science. http://etheses.lse.ac.uk/3284/1/Blunt_heirachies_of_evidence.pdf

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: a new approach to evaluating complex

public health interventions. *Soc Sci Med. 75*(12), 2299-2306. https://doi.org/10.1016/j.socscimed.2012.08.032.
Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2013). Methods don't make assumptions, researchers do: a response to Marchal et al. *Soc Sci Med. 94*, 81–82. https://doi.org/10.1016/j.socscimed.2013.06.026
Bonell, C., Jamal, F., Melendez-Torres, G.J., & Cummins, S. (2015). 'Dark logic': Theorizing the harmful consequences of public health interventions. *Epidemiology & Community Health, 69*(1), 95–98. https://doi.org/10.1136/jech-2014-204671
Boost, D., Raeymaeckers, P., Hermans, K., & Elloukmani, S. (2020). Overcoming non-take-up of rights: A realist evaluation of Integrated Rights-Practices. *Journal of Social Work, 21*(4), 831-852. https://doi.org/10.1177/1468017320948332
Bostock, L., Forrester, D., Patrizo, L., Godfrey, T., Zonouzi, M., Antonopoulou, V., Bird, H., & Tinarwo, M. (2017). *Scaling and deepening the Reclaiming Social Work model. Evaluation Report.* Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625227/Scaling_and_deepening_the_Reclaiming_Social_Work_model.pdf
Bostock, L, Patrizo, L., Godfrey, T., & Forrester, D. (2019a). What is the impact of supervision on direct practice with families? *Children and Youth Services Review, 105*, Article 104428. https://doi.org/10.1016/j.childyouth.2019.104428
Bostock, L., Patrizo, L., Godfrey, T., Munro, E., & Forrester, D. (2019b). How do we assess the quality of group supervision? Developing a coding framework. *Children and Youth Services Review*, 100, 515-524. https://doi.org/10.1016/j.childyouth.2019.03.027
Bostock, L., & Newlands, F. (2020). *Scaling and deepening the Reclaiming Social Work model: longitudinal follow up.* Evaluation report. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/933120/Reclaiming_Social_Work_-_Bedfordshire.pdf
Bronsard, G., Alessandrini, M., Fond, G., Loundou, A., Auquier, P., Tordjman, S., & Boyer, L. (2016). The Prevalence of Mental Disorders among Children and Adolescents in the Child Welfare System: A Systematic Review and Meta-Analysis. *Medicine, 95*(7), e2622. https://doi.org/10.1097/MD.0000000000002622
Bullinger, M., Brütt, A.L., Erhart, M., Ravens-Sieberer, U., & BELLA Study Group. (2008). Psychometric properties of the KINDL-R Questionnaire: Results of the Bella-study. *Eur Child Adolesc Psychiatry*, *17*, 125–132. https://doi.org/10.1007/s00787-008-1014-z
Burford, G., & Hudson, J. (2000). General Instoduction: Family Group Conference Programming. In G. Burford & J. Hudson (Eds.), *Family Group*

Conferencing. New Directions in Community-Centered Child & Family Practice (pp. xix-xxvii). Routledge.
Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-experimental Designs for Research on Teaching*. Rand McNally & Company.
Caffrey, L., & Browne, F. (2022). Understanding the social worker–family relationship through self-determination theory: A realist synthesis of Signs of Safety. *Child & Family Social Work*, January 2022. https://doi.org/10.1111/cfs.12903
Carr, A., & Stratton, P. (2017). The Score Family Assessment Questionnaire: A Decade of Progress. *Family Process, 56*(2), 285-301. https://doi.org/10.1111/famp.12280
Cartwright, N., & Hardie, J. (2012). *Evidence-Based Policy. A Practical Guide to Doing It Better*. Oxford University Press.
Cecchin, G. (1987). Hypothesizing, circularity, and neutrality revisited: An invitation to curiosity. *Family Process, 26*(4), 405–413. https://doi.org/10.1111/j.1545-5300.1987.00405.x
Chen, H.-T. (1990). *Theory-driven evaluations*. Sage Publications.
Chen, M., & Chan, K. L. (2016). Effects of parenting programs on child maltreatment prevention: A meta-analysis. *Trauma, Violence, & Abuse, 17*(1), 88-104. https://doi.org/10.1177/1524838014566718
Christensen, D. N., Todahl, J., & Barrett, W. G. (1999). *Solution-based casework: An introduction to clinical and case management skills in casework practice*. Aldine DeGruyter.
Clarke, B., Gillies, D., Illari, P., Russo, F., Williamson, J. (2013). The evidence that evidence-based medicine omits. *Preventive Medicine, 57*(6), 745-747. https://doi.org/10.1016/j.ypmed.2012.10.020.
Clarke, B., Gillies, D., Illari, P., Russo, F., & Williamson, J. (2014). Mechanisms and the Evidence Hierarchy. *Topoi 33*, 339-360. https://doi.org/10.1007/s11245-013-9220-9
Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ,* 337, Article 1655. https://doi.org/10.1136/bmj.a1655
Cross, S., Hubbard, A., & Munro, E. (2010). *Reclaiming Social Work. London Borough of Hackney Children and Young People's Services. Part 1: Independent evaluation.* The London School of Economics and Political Science.
Dalkin S.M., Greenhalg, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science 10*(49), 1–7. https://doi.org/10.1186/s13012-015-0237-x
Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized con-trolled trials. *Social Science & Medicine, 210*, 2-21. https://doi.org/10.1016/j.socscimed.2017.12.005.
Diaz, C. (2020). *Decision making in child and family social work: perspectives on participation*. Policy Press.

Dijkstra, S., Creemers, H., Asscher, J., Deković, M., & Stams, G. (2016). The effectiveness of family group conferencing in youth care: A meta-analysis. *Child Abuse & Neglect, 62*, 100-110, https://doi.org/10.1016/j.chiabu.2016.10.017

Elamin, M. B., & Montori, V. M. (2012). The hierarchy of evidence: From unsystematic clinical observations to systematic reviews. In *Neurology: An Evidence-Based Approach* (pp. 11-24). Springer New York. https://doi.org/10.1007/978-0-387-88555-1_2

Emmel, N., Greenhalg, J., Manzano, A., Monaghan, M., & Dalkin, S. (2018). *Doing Realist Research*. Sage.

Erhart, M., Ellert, U., Kurth B.M., & Ravens-Sieberer U. (2009). Measuring adolescents' HRQoL via self reports and parent proxy reports: an evaluation of the psychometric properties of both versions of the KINDL-R instrument. *Health and Quality of Life Outcomes, 7*, 77–88. https://doi.org/10.1186/1477-7525-7-77

Esping-Andersen, G. (1990). *The three worlds of welfare capitalism*. Polity.

Euser, S., Alink, L. R., Stoltenborgh, M., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2015). A gloomy picture: a meta-analysis of randomized controlled trials reveals disappointing effectiveness of programs aiming at preventing child maltreatment. *BMC Public Health, 15*(1), 1068. https://doi.org/10.1186/s12889-015-2387-9

Evans, D. (2003). Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing 12*(1), 77-84. https://doi.org/10.1046/j.1365-2702.2003.00662.x

Evans, R., Scourfield, J., & Murphy, S. (2015) Pragmatic, formative process evaluations of complex interventions and why we need more of them. *Journal of Epidemiology and Community Health*, 69, 925-926. http://dx.doi.org/10.1136/jech-2014-204806

Falconer, R., & Shardlow, S. (2018). Comparing child protection decision-making in England and Finland: supervised or supported judgement? *Journal of Social Work Practice, 32*(2), 111-124. https://doi.org/10.1080/02650533.2018.1438996

Fischer, J. (1973). Is casework effective? a review. *Social Work, 18*(1), 5-20.

Flay, B., Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., Moscicki, E., Schinke, S., Valentine, J., & Ji, P. (2005). Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science, 6*(3), 151-175. https://doi.org/10.1007/s11121-005-5553-y

Forrester, D., Westlake, D., McCann, M., Thurnham, A., Shefer, G., Glynn, G., & Killian, M. (2013). *Reclaiming social work? An evaluation of systemic units as an approach to delivering children's services.* University of Bedfordshire.

Forrester, D. (2017). Outcomes in Children's Social Care. *Journal of Children's Services, 12*(2/3), 144-157. https://doi.org/10.1108/JCS-08-2017-0036

Forrester, D., Westlake, D., Killian, M., Antonopoulou, V., McCann, M., Thurnham, A., Thomas, R., Waits, C., Whittaker, C., & Hutchison, D. (2018). A randomized controlled trial of training in Motivational

Interviewing for child protection. *Children and Youth Services Review*, 88, 180–190. https://doi.org/10.1016/j.childyouth.2018.02.014
Forrester, D., Westlake, D., Killian, M., Antonopolou, V., McCann, M., Thurnham, A., Thomas, R., Waits, C., Whittaker, C., & Hutchison, D. (2019). What Is the Relationship between Worker Skills and Outcomes for Families in Child and Family Social Work? *The British Journal of Social Work, 49*(8), December 2019, 2148–2167. https://doi.org/10.1093/bjsw/bcy126
Forrester, D., Killian, M., Westlake, D., & Sheehan, L. (2020). Patterns of practice: an exploratory factor analysis of child and family social worker skills. *Child and Family Social Work 25*(1), 108-117. https://doi.org/10.1111/cfs.12664
Frost, N, Abram, F., & Burgess, H. (2014). Family group conferences: context, process and ways forward. *Child & Family Social Work, 19*(4), 480-490. https://doi.org/10.1111/cfs.12047
Gambrill, E. (2001). Social Work: An Authority-Based Profession. *Research on Social Work Practice, 11*(2), 166-175. https://doi.org/10.1177/104973150101100203
Gambrill, E. (2006). Evidence-Based Practice and Policy: Choices Ahead. *Research on Social Work Practice, 16*(3), 338-357. https://doi.org/10.1177/1049731505284205
Gambrill, E., (2010). Evidence-Informed Practice: Antidote to Propaganda in the Helping Profession? *Research on Social Work Practice, 20*(3), 302-320. https://doi.org/10.1177/1049731509347879
Gilbert, N. (Ed.) (2018). *Combatting Child Abuse: International Perspectives and Trends*. Oxford University Press.
Gilbert, N., Parton, N., & Skivenes, M. (Eds.) (2011). *Child Protection Systems. International Trends and Orientations*. Oxford Scholarship Online. https://doi.org/10.1093/acprof:oso/9780199793358.003.0006
Gillingham, P. (2018). Evaluation of Practice Frameworks for Social Work with Children and Families: Exploring the Challenges. *Journal of Public Child Welfare, 12*(2), 190-203. https://doi.org/10.1080/15548732.2017.1392391
Gitterman, A., & Knight, C. (2018). Evidence-guided practice: Integrating the science and art of social work. *Families in Society: The Journal of Contemporary Social Services, 94*(2), 70–78. https://doi.org/10.1606/1044-3894.4282
Goodman, S., & Trowler, I. (2012). *Social work reclaimed: Innovative frameworks for child and family social work practice*. Jessica Kingsley Publishers.
Goolsby, J., Rich, B., Hinnant, B., Habayeb, S., Berghorst, L., De Los Reyes, A., & Alvord, M. (2018). Parent-Child Informant Discrepancy is Associated with Poorer Treatment Outcome. *Journal of Child and Family Studies, 27*(4), 1228-1241. https://doi.org/10.1007/s10826-017-0946-7
Gray, M., Joy, E., Plath, D., & Webb, S. (2013). Implementing Evidence-Based Practice: A Review of the Empirical Research Literature. *Research on Social Work Practice, 23*(2), 157–166. https://doi.org/10.1177/1049731512467072

Greenhalg, T. (2000). *How to read a paper. The basics of evidence based medicine*. 2nd ed. BMJ.

Greenhalgh, T., Pawson, R., Wong, G., Westhorp, G., Greenhalgh, J., Manzano, A., & Jagosh, J. (2017). *Realist evaluation, realist synthesis, realist research – what's in a name?* The RAMESES II Project.

Greenhalg, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *Eur J of Clin Invest., 48*(6), e12931. https://doi.org/10.1111/eci.12931

Gugiu, P. C., & Rodriguez-Campos, L. (2007). Semi-structured interview protocol for constructing logic models. *Evaluation and Program Planning, 30*(4), 339–350. https://doi.org/10.1016/j.evalprogplan.2007.08.004

Guyatt, G., Oxman, A., Vist, G., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ 2008, 336*, 924. https://doi.org/10.1136/bmj.39489.470347.AD

Gypen, L., Vanderfaeillie, J., De Maeyer, S., Belenger, L., & Van Holen, F. (2017). Outcomes of children who grew up in foster care: Systematic review. *Children and Youth Services Review, 76*, 74-83. https://doi.org/10.1016/j.childyouth.2017.02.035

Havnen, K., & Christiansen, Ø. (2014). *Knowledge Review on Family Group Conferencing: Experiences and Outcomes*. Regional Centre for Child and Youth Mental Health and Child Welfare. Uni Research Health.

Heino, T. (2007). *Keitä ovat lastensuojelun uudet asiakkaat? Tutkimus lapsista ja perheistä tilastolukujen takana [Who are the new service users in child protection? Research on children and their families behind statistics]*. Working Paper 30/2007. Finnish Research and Development Centre for Welfare and Health.

Heino, T. (2009). *Family group conference from a child perspective. Nordic research report.* Report 9/2009. Finnish Institute for Health and Welfare.

Hellsten, K. (1993). *Vaivaishoidosta hyvinvointivaltion kriisiin. Hyvinvointivaltion kehitys ja sosiaaliturvajärjestelmän muotoutuminen Suomessa [From poor relief to welfare state crisis. The development of the welfare state and social security system in Finland]*. [Doctoral Dissertation, University of Helsinki]. University of Helsinki.

Herbert, M. (1978). *Conduct Disorders of Childhood and Adolescence*. Wiley.

Herbert, M. (1987). *Behavioural Treatment of Children with Problems* (2nd ed.). Academic Press.

Holosko, M. (2010). What Types of Designs are We Using in Social Work Research and Evaluation? *Research on Social Work Practice, 20*(6), 665-673. https://doi.org/10.1177/1049731509339586

Hoppania, H., Karsio, O., Näre, L., Vaittinen, T., & Zechner, M. (2022). Financialization of Eldercare in a Nordic Welfare State. *Journal of Social Policy*, 1-19. https://doi.org/10.1017/S0047279422000137

Houston, S. (2001). Transcending the fissure in risk theory: critical realism and child welfare. *Child & Family Social Work, 6*(3), 219-228. https://doi.org/10.1046/j.1365-2206.2001.00205.x

Hytönen, K-M., Malinen, A., Salenius, P., Haikari, J., Markkola, P., Kuronen, M., & Koivisto, J. (2016). *Lastensuojelun sijaishuollon epäkohdat ja lasten kaltoinkohtelu 1937-1983 [The defects and child maltreatment in out-of-home care 1937-1983]*. Reports and memorandums 2016:22. Ministry of Social Affairs and Health.

Isokuortti, N., Aaltio, E., Laajasalo, T., & Barlow, J. (2020). Effectiveness of child protection practice models: a systematic review. *Child Abuse & Neglect, 108*, Article 104632. https://doi.org/10.1016/j.chiabu.2020.104632.

Isokuortti, N., & Aaltio, E. (2020). Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland. *Children and Youth Services Review, 119*, Article 105647. https://doi.org/10.1016/j.childyouth.2020.105647

Jackson, D. (1957). The question of family homeostasis. *Psychiatry Quarterly Supplement, 31*, 79-99. https://doi.org/10.1007/BF00936266

Jewell, T., Carr, A., Stratton, P., Lask, J., & Eisler, I. (2013). Development of a children's version of the SCORE index of family function and change. *Family Process, 52*(4), 673–684. https://doi.org/10.1111/famp.12044

Julien-Chinn, F., & Lietz, C. (2019). Building learning cultures in the child welfare workforce. *Children and Youth Services Review, 99*, 360–365. https://doi.org/10.1016/j.childyouth.2019.01.023

Julkunen, R. (2001). *Suunnanmuutos. 1990-luvun sosiaalipoliittinen reformi Suomessa [Changing the course. Social policy reform of the 1990s in Finland]*. Vastapaino.

Junnonen, S.-R., Hämäläinen, J., Töttö, P., Väisänen, R., Rantonen, O., & Salo, P. (2019). Sosiaalityöntekijöiden työn kuormittavuus ja siinä tapahtuneet muutokset vuosina 2000-2012 [Social workers' workload and its changes in 2000-2012]. *Janus, 27*(2), 144–163. https://doi.org/10.30668/janus.65729

Jutila, M. (2011). Narrowing of Public Responsibility in Finland, 1990-2010. *Social Policy & Administration, 45*(2), 194-205. https://doi.org/10.1111/j.1467-9515.2010.00764.x

Kaidesoja, T. (2005). The Trouble with Trancendental Arguments, *Journal of Critical Realism, 4*(1), 28-61. https://doi.org/10.1558/jocr.v4i1.28

Kananoja, A., Lavikainen, M., & Oranen, M. (2013). *Toimiva lastensuojelu. Selvitysryhmän loppuraportti [Functioning Child Protection. Final report of the clearing group]*. Publications 2013:19. Ministry of Social Affairs and Health.

Kantola, A. & Kananen, J. (2013). Seize the Moment: Financial Crisis and the Making of the Finnish Competition State. *New Political Economy*, 18(6), 811-826. https://doi.org/10.1080/13563467.2012.753044

Karjalainen, P. (2012). Realistinen arviointi kuntouttavan sosiaalityön vaikuttavuudesta [Realist evaluation on the effectiveness of empowering social work]. In A. Pohjola, T. Kemppainen, & S. Väyrynen (Eds.),

Sosiaalityön vaikuttavuus [Social work effectiveness] (pp. 162-198). Lapland University Press.

Karjalainen, P., Kiviruusu, O., Aronen, E.T., & Santalahti, P. (2019). Group-based parenting program to improve parenting and children's behavioral problems in families using special services: A randomized controlled trial in a real-life setting. *Children and Youth Services Review, 96*, 420–429. https://doi.org/10.1016/j.childyouth.2018.12.004

Karjalainen, P., Kiviruusu, O., Santalahti, P., & Aronen, E.T. (2020). Parents' perceptions of a group-based parenting programme in families with child protection and other family support services in a real-life setting. *Child & Family Social Work*, 1–12. https://doi.org/10.1111/cfs.12787

Karjalainen, P. (2021). *Parenting Intervention to help children with behaviour problems in child protection and other family support services.* [Academic Dissertation, University of Helsinki].

Kazi, M. (2003). *Realist Evaluation in Practice : Health and Social Work*. Sage Publications.

Kestilä, L. Paananen, R., Väisänen, A., Muuri, A., Merikukka, M., Heino, T., & Gissler, M. (2012). Kodin ulkopuolelle sijoittamisen riskitekijät. Rekisteripohjainen seurantatutkimus Suomessa vuonna 1987 syntyneistä. [The risk factors for out-of-home care. Registry-based follow-up study on the 1987 birth cohort in Finland]. *Yhteiskuntapolitiikka, 77*(1), 34-52.

Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money*. A Harvest Book, Harcourt Inc.

Kjørstad, M., & Solem, M.-B. (2017). *Critical Realism for Welfare Professions*. Routledge.

Koivula, M., Ellonen, N., Ikonen, R., & Paavilainen, E. (2021). Nuorten kaltoinkohtelukokemukset ammatillisissa perhekodeissa ja lastensuojelulaitoksissa [Youg people's experiences of maltreatment in residential care], *Yhteiskuntapolitiikka 86*(3), 263-274.

Korpi, W. (1980). Social Policy and Distributional Conflict in the Capitalist Democracies: A Preliminary Comparative Framework. *West European Politics, 3*, 296-316. https://doi.org/10.1080/01402388008424288

Korpi, W., & Palme, J. (1998). The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries. *American Sociological Review, 63*(5), 661–87. https://doi.org/10.2307/2657333

Korpilahti, U., Lassander, M., Valkeinen, H., & Hakulinen-Viitanen, T. (2015). KINDL-R: Lasten ja nuorten terveyteen liittyvää elämänlaatua arvioiva geneerinen mittari [KINDL-R: The generic instrument for assessing the health-related quality of life of children and youth]. *Sosiaalilääketieteellinen aikakauslehti, 52,* 170–181.

Koskiaho, B. (2008). *Hyvinvointipalvelujen tavaratalossa [In the department store of welfare services]*. Vastapaino.

Kotiranta, T., & Mäntysaari, M. (2017). Vaikuttavuus tulee ja menee – ja palaa taas takaisin [Effectiveness comes and goes – and returns]. In A. Tuulio-

Henriksson, L. Kalliomaa-Puha, & P.-L. Rauhala (Eds.), *Harkittu, tutkittu, avoin: Marketta Rajavaaran juhlakirja* (pp. 35-50). Kansaneläkelaitos. http://hdl.handle.net/10138/178926

Kuusi, P. (1961). *60-luvun sosiaalipolitiikka [The Social Policy for the Sixties]*. WSOY.

Kuusi, P. (1964). *The Social Policy for the Sixties.* Finnish Social Policy Association.

Kääriälä, A., & Hiilamo, H. (2017). Children in out-of-home care as young adults: A systematic review of outcomes in the Nordic countries. *Children and Youth Services Review, 79*, 107-114. https://doi.org/10.1016/j.childyouth.2017.05.030

Kääriälä, A., Berlin, M., Lausten, M. Hiilamo, H., & Ristikari, T. (2018). Early school leaving by children in out-of-home care: A comparative study of three Nordic countries. *Children and Youth Services Review, 93*, 186-195. https://doi.org/10.1016/j.childyouth.2018.06.007

Kääriälä, A., Haapakorva, P., Pekkarinen, E., & Sund, R. (2019). From care to education and work? Education and employment trajectories in early adulthood by children in out-of-home care. *Child Abuse & Neglect, 98*, Article 104144. https://doi.org/10.1016/j.chiabu.2019.104144

Laird, S. E., Morris, K., Archard, P., & Clawson, R. (2017). Working with the whole family: What case files tell us about social work practices. *Child & Family Social Work, 22*(3), 1322-1329. https://doi.org/10.1111/cfs.12349

Laird, S. E., Morris, K., Archard, P., & Clawson, R. (2018). Changing practice: The possibilities and limits for reshaping social work practice. *Qualitative Social Work, 17*(4), 577–593. https://doi.org/10.1177/1473325016688371

Lahtinen, P., Männistö, L., & Raivio, M. 2017. *Kohti suomalaista systeemistä lastensuojelun toimintamallia: Keskeisiä periaatteita ja reunaehtoja [Toward the Finnish Systemic Practice Model for child protection. Core principles and preconditions]*. Discussion paper 7/2017. Finnish Institute for Health and Welfare.

Lang, P., Little, M., & Cronen, V. (1990). The systemic professional. Domains of action and the question of neutrality. *Human Systems, 1*, 39–55.

Lastensuojelu 2020 [Child welfare 2020] Statistical report 19/2021. Official Statistics of Finland. National Institute for Health and Welfare.

Lorenz, W. (1994). *Social Work in a Changing Europe*. Routledge.

Lorås, L., Bertrando, P., & Ness, O. (2017). Researching Systemic Therapy History: In Search of a Definition. *Journal of Family Psychotherapy, 28*(2), 134–149. https://doi.org/10.1080/08975353.2017.1285656

Lunabba, H. (2013). *När vuxna möter pojkar i skolan- insyn, inflytande och sociala relationer [When adults encounter boys in school - insight, influence and social relations].* [Academic Dissertation, University of Helsinki]. Mathilda Wrede Institute Research servies 1/2013.

Maasalo, K., Fontell, T., Wessman, J., & Aronen, E. (2016). Sleep and behavioural problems associate with low mood in Finnish children aged 4–12 years: an epidemiological study. *Child Adolesc Psychiatry Ment Health, 10*(37). https://doi.org/10.1186/s13034-016-0125-4

Madsen, W. (2011). Collaborative Helping Maps: A Tool to Guide Thinking and Action in Family-Centered Services. *Family Process 50*(4), 529-543. https://doi.org/10.1111/J.1545-5300.2011.01369.X

Manzano, A. (2016). The craft of interviewing in realist evaluation. *Evaluation, 22*(3), 342-360. https://doi.org/10.1177/1356389016638615

Marchal, B., Westhorp, G., Wong, G., Van Belle, S., Greenhalgh, T., Kegels, G., & Pawson R. (2013). Realist RCTs of complex interventions - an oxymoron. *Soc Sci Med, 94*, 124-8. https://doi.org/10.1016/j.socscimed.2013.06.025.

McCafferty, S. (2012). Behavioural-based Interventions: Social Learning Theory. In S. Goodman, & I. Trowler (Eds.), *Social Work Reclaimed. Innovative Frameworks for Child and Family Social Work Practice* (pp.34–56). Jessica Kingley Publishers.

McKenna, S., Donnelly, M., Onyeka, I.N., O'Reilly, D., & Maguire, A. (2021). Experience of child welfare services and long-term adult mental health outcomes: a scoping review. *Social Psychiatry and Psychiatric Epidemiology, 56*(7), 1115-1145. https://doi.org/10.1007/s00127-021-02069-x

Metteri, A. (2012). *Hyvinvointivaltion lupaukset, kohtuuttomat tapaukset ja sosiaalityö [Promises of a welfare state, non-fulfilment of these promises and social work].* [Academic Dissertation, University of Tampere.]

Meyers, D., Durlak, J., & Wandersman, A. (2012). The Quality Implementation Framework: A Synthesis of Critical Steps in the Implementation Process. *American Journal of Community Psychology, 50*(3-4), 462–480. https://doi.org/10.1007/s10464-012-9522-x.

Miller, L.B. & Bentovim, A. (2003). Assessing Families: The Family Assessment of Family Competence, Strengths and Difficulties. In M. Bell and K. Wilson (Eds), *The Practitioner's Guide to Working with Families* (pp.57-84). Palgrave.

Mishra, R. (1977). *Society and Social Policy*. Macmillan.

Mononen, E., Hakovirta, M., Rasinkangas, J., & Ikonen R. (2021). Lasten vertaissuhdeväkivalta lastensuojelun sijaishuollossa [Children's peer violence in out-of-home care]. *Yhteiskuntapolitiikka, 86*(3), 275-286.

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350, h1258. https://doi.org/10.1136/bmj.h1258

Morago, P. (2006). Evidence-based practice: from medicine to social work. *European Journal of Social Work, 9*(4), 461-477. https://doi.org/10.1080/13691450600958510

Murad, M.H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *BMJ Evidence-Based Medicine, 21*, 125-127. http://dx.doi.org/10.1136/ebmed-2016-110401

Myrdal, G. (1957). *Economic Theory and Under-developed Regions*. Duckworth.

Mänttäri-van der Kuip, M. (2015). *Work-related well-being among Finnish frontline social workers in an age of austerity*. [Academic Dissertation, University of Jyväskylä].

Mänttäri-van der Kuip, M. (2016). Moral distress among social workers: The role of insufficient resources. *International Journal of Social Welfare, 25*(1): 86–97. https://doi.org/10.1111/ijsw.12163

Mäntysaari, M. (1991). *Sosiaalibyrokratia asiakkaiden valvojana : byrokratiatyö, sosiaalinen kontrolli ja tarpeitten sääntely sosiaalitoimistoissa [Social byreaucracy as an overseer of service users. Bureaucracy, social control and regulation of needs in social work offices]*. Vastapaino.

Mäntysaari, M. (2003). Realism as a foundation for social work knowledge. *Qualitative Social Work, 4*(1), 87-98. https://doi.org/10.1177/1473325005050202

Mäntysaari, M. (2013). Hyvinvointivaltion ideologi – Richard M. Titmussin näkemys sosiaalihuollosta [The ideologist of a welfare state – Richard M. Titmuss's view on social care]. In J. Saari, S. Taipale, & S. Kainulainen (Eds.), *Hyvinvointivaltion moderneja klassikoita [The modern classics of welfare state]* (pp. 327-344). Diakonia-ammattikorkeakoulu.

National Institute for Health and Care Excellence (2014). *Developing NICE guidelines: the manual [PMG20]*. RCP. http://www.nice.org.uk/article/PMG20, accessed 08/07/15.

Olsson, T., & Sundell, K. (2016). Research that Guides Practice: Outcome Research in Swedish PhD Theses Across Seven Disciplines 1997–2012. *Prevention Science, 17*, 525-532. https://doi.org/10.1007/s11121-016-0640-9

Paakkonen, T., & Paakkonen, H. (2018). Finnish schoolchildren's perceived health-related quality of life deteriorates remarkably with age. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology 6*(4), 152-158. https://doi.org/10.21307/sjcapp-2018-014

Palazolli, M.S., Boscolo, L., Cecchin, G., & Prata, G. (1980). Hypothesizing-circulatity-neutrality: three guidelines for the conductor of the session. *Family Process, 19*(1), 3-12. https://doi.org/10.1111/j.1545-5300.1980.00003.x

Parrish, D. (2018). Evidence-Based Practice: A Common Definition Matters, *Journal of Social Work Education, 54*(3), 407-411. https://doi.org/10.1080/10437797.2018.1498691

Parton, N., & Berridge, D. (2011). Child Protection in England. In N. Gilbert, N. Parton & M. Skivenes (Eds.), *Child Protection Systems. International Trends and Orientations*. Oxford Scholarship Online. https://doi.org/10.1093/acprof:oso/9780199793358.003.0006

Paterson-Young, C. (2021) Exploring how children subjected to violence in the home cope with experiences in Secure Training Centres. *Child Abuse & Neglect, 117*, Article 105076. https://doi.org/10.1016/j.chiabu.2021.105076

Patomäki, H. (2007). *Uusliberalismi Suomessa. Lyhyt historia ja tulevaisuuden vaihtoehdot [Neoliberalism in Finland. Short history and options for the future]*. WSOY.

Patomäki, H., & Wight, C. (2000). After Postpositivism? The Promises of Critical Realism. *International Studies Quarterly, 44*(2), 213-237. https://doi.org/10.1111/0020-8833.00156

Pawson, R. (2006). *Evidence-based Policy. A Realist Perspective*. Sage.
Pawson, R. (2013). *The science of evaluation. A Realist Manifesto*. Sage.
Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. SAGE.
Pekkarinen, E. (2010). *Stadilaispojat, rikokset ja lastensuojelu : Viisi tapaustutkimusta kuudelta vuosikymmeneltä [Stadi-boys, Crime and Child Protection - Five Case Studies from Six Decades]*. [Academic Dissertation, University of Helsinki]. Nuorisotutkimusseura.
Pekkarinen, E., & Tapola-Haapala, M. (2009). Kriittinen realismi sosiaalityössä – tiedontuotannosta emansipaatioon [Critical realism in social work – from knowledge production to emancipation]. In M. Mäntysaari, A. Pohjola, & T. Pösö (Eds.), *Sosiaalityön teoria [Theory in social work]* (pp. 183–205). PS-kustannus.
Pendry, N. (2012). Systemic Practice in a Risk Management Context. In S. Goodman & I. Trowler (Eds.), *Social work reclaimed: Innovative frameworks for child and family social work practice* (pp. 26-33). Jessica Kingsley Publishers.
Petrelius, P., & Uutela, A. (2020). *Systeemisen lastensuojelun toimintamallin käyttöönotto ja juurruttaminen. Konsultoivien sosiaalityöntekijöiden kokemuksia [The implementation and deepening of Systemic practice model. Experiences of consultative social workers].* Working paper 24/2020. Finnish Institute for Health and Welfare.
Pohjola, A., Kemppainen, T., & Väyrynen, S. (2012). *Sosiaalityön vaikuttavuus. [Social work effectiveness]*. Lapin yliopistokustannus.
Pohjola, A. (2012). Moniulotteinen vaikuttavuus [Multidimensional effectiveness]. In A. Pohjola, T. Kemppainen & S. Väyrynen (Eds.), *Sosiaalityön vaikuttavuus [Social work effectiveness]* (pp.9-17). Lapin yliopistokustannus.
Pölkki, P., Vornanen, R., Pursiainen, M., & Riikonen, M. (2012). Children's Participation in Child-protection Processes as Experienced by Foster Children and Social Workers. *Child Care in Practice, 18*(2), 107-125. https://doi.org/10.1080/13575279.2011.646954
Pösö, T. (1997). Finland: Child Abuse as a Family Problem. In N. Gilbert (Ed.) *Combatting Child Abuse: International Perspectives and Trends* (pp. 143-163). Oxford University Press.
Pösö, T. (2011). Combatting Child Abuse in Finland. From Family to Child-centered Orientation. In N. Gilbert, N. Parton & M. Skivenes (Eds.), *Child Protection Systems. International Trends and Orientations*. Oxford Scholarship Online. https://doi.org/10.1093/acprof:oso/9780199793358.003.0006
Pösö, T., Pekkarinen, E., Helavirta, S., & Laakso, R. (2018). 'Voluntary' and 'involuntary' child welfare: Challenging the distinction. *Journal of Social Work, 18*(3), 253–272. https://doi.org/10.1177/1468017316653269
Rauhala, P.-L. (1996). *Miten sosiaalipalvelut ovat tulleet osaksi suomalaista sosiaaliturvaa? [How did the social services became a part of Finnish social security?]* [Academic Dissertation, Tampere University]
Ravens-Sieberer, U., & Bullinger, M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric

and content analytical results. *Qual Life Res 1998, 7*(5), 399-407. https://doi.org/10.1023/a:1008853819715

Reekers, S. E., Dijkstra, S., Stams, G. J. J., Asscher, J. J., & Creemers, H. E. (2018). Signs of effectiveness of signs of safety? – A pilot study. *Children and Youth Services Review, 91*, 177-184. https://doi.org/10.1016/j.childyouth.2018.06.012.

Regehr, C., Stern, S., & Shlonsky, A. (2007). Operationalizing Evidence-Based Practice: The Development of an Institute for Evidence-Based Social Work. *Research on Social Work Practice, 17*, 408–416. https://doi.org/10.1177/1049731506293561

Riihinen, O. (Ed.) (1992). *Sosiaalipolitiikka 2017. Näkökulmia suomalaiseen yhteiskunnan kehitykseen ja tulevaisuuteen [Social policy 2017. Perspectives in the development and future of the Finnish society]*. WSOY.

Rosen, A., Proctor, E. K., & Staudt, M. M. (1999). Social work research and the quest for effective practice. *Social Work Research, 23*(1), 4–14. https://doi.org/10.1093/swr/23.1.4

Rostila, I. (2001). *Sosiaalisen kuntoutuksen mekanismit : Monet-projektin realistinen arviointi [The mechanisms of empowering social work. Realist evaluation of Monet project]*. Finnish Research and Development Centre for Welfare and Health.

Rousu, S. (2007) *Lastensuojelun tuloksellisuuden arviointia organisaatiossa. Näkymätön tuloksellisuus näkyväksi. [Assessment of Child Welfare Effectiveness in Organisations. Making the Invisible Visible]*. [Academic Dissertation, University of Tampere]. The Association of Finnish Local and Regional Authorities.

Ryynänen, J. & Heinola, V. (2021). Kokemuksia systeemisen otteen kehittämisestä ja asiakkaan osallistumisesta tiimeihin Espoon lastensuojelussa [Experiences of developing the systemic approach and of the service users' participation in teams in child protection in Espoo]. In. P. Petrelius, L. Yliruka, & N. Miettunen (Eds.), *Systeemisiä kokeiluja – kohti jatkuvaa yhdessä oppimista [Systemic experiments – towards ongoing collective learning]* (pp. 77-87). Working paper 7/2021. Finnish Institute for Health and Welfare.

Sackett, D. (1981). How to read clinical journals: IV. To determine etiology or causation. *CMAJ, 124*, 985-990.

Sackett, D., Rosenberg, W., Muir-Gray, J., Haynes, R., & Richardson, W. (1996). Evidence based medicine: what it is and what it isn't. *BMJ, 312*(13), 71-72. https://doi.org/10.1136/bmj.312.7023.71

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-Based Medicine: How to Practice and Teach EBM* (2nd Ed.). Churchill Livingstone.

Salovaara, S., & Ylönen, K. (2021). Client information systems' support for case-based social work: experiences of Finnish social workers. *Nordic Social Work Research*, https://doi.org/10.1080/2156857X.2021.1999847

Satka, M., Julkunen, I., Kääriäinen, A., Poikela, R., Yliruka, L. & Muurinen, H. (2016) *Johdanto – Käytäntötutkimus tietona ja taitona [Introduction - Practice*

research as knowledge and skill]. In M. Satka, I. Julkunen, A. Kääriäinen, R. Poikela, L. Yliruka & H. Muurinen (Eds.), *Käytäntötutkimuksen taito [The skill of practice research]* (pp.8-30). Heikki Waris -instituutti and Mathilda Wrede -institutet.

Sayer, A. (2000). *Realism and Social Science*. Sage.

Schiltroth, K. (2012). Preconditions: Structure, Continuity and Momentum. In S. Goodman & I. Trowler (Eds.), *Social work reclaimed: Innovative frameworks for child and family social work practice* (pp. 69-79). Jessica Kingsley Publishers.

Scruggs, L., & Allan, J. P. (2008). Social Stratification and Welfare Regimes for the Twenty-First Century: Revisiting The Three Worlds of Welfare Capitalism. *World Politics, 60*(4), 642–664. https://doi.org/10.1353/wp.0.0020

Shanks, Backe-Hansen, E., Eriksson, P., Lausten, M., Lundström, T., Ranta, H., & Sallnäs, M. (2021). Privatisation of residential care for children and youth in Denmark, Finland, Norway, and Sweden. *Nordisk Välfärdsforskning, 6*(3), 128–141. https://doi.org/10.18261/issn.2464-4161-2021-03-02

Shaw, J., Gray, S.C., Baker, G.R., Denis, J.-L., Breton, M., Gutberg, J., Embuldeniya, G., Carswell, P., Dunham, A., McKillop, A., Kenealy, T., Sheridan, N., & Wodchis, W. (2018). Mechanisms, contexts and points of contention: operationalizing realist-informed research for complex health interventions. *British Medical Research Methodology, 18*(178). https://doi.org/10.1186/s12874-018-0641-4

Sheehan, L., O'Donnell, C., Brand, S., Forrester, D., Addis, S., & El-Banna, A. (2018). *Signs of safety: Findings from a mixed-methods systematic review focussed on reducing the need for children to be in care.* What Works Centre for Children's Social Care.

Sihto, & Van Aerschot, L. (2021). Care Poverty Within the Home Space: Exploring the Emotional Experiences of Unmet Care Needs. *Frontiers in Sociology, 6*. https://doi.org/10.3389/fsoc.2021.637799

Sipilä, J., & Österbacka, E. (2013). *Enemmän ongelmien ehkäisyä, vähemmän korjailua? Perheitä ja lapsia tukevien palvelujen tuloksellisuus ja kustannusvaikuttavuus [More prevention, less repairing? The profitability and cost-effectiveness of child and family services]*. Publications 11/2013. Ministry of Finance.

Siverns, K., & Morgan, G. (2019). Parenting in the context of historical childhood trauma: An interpretive meta-synthesis. *Child Abuse & Neglect, 98*, Article 104186. https://doi.org/10.1016/j.chiabu.2019.104186

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M. et al. (2021). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ, 374*, Article 2061. https://doi.org/10.1136/bmj.n2061

Sorsa, V-P. (2017). Riittävätkö rahat yhteiseen hyvinvointiin? [Can we afford collective well-being?] In J. Kananen (Ed.), *Kilpailuvaltion kyydissä: Suomen hyvinvointimallin tulevaisuus* (pp. 69-96). Gaudeamus.

Stratton, P., Bland, J., Janes, E., & Lask, J. (2010). Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: The SCORE. *Journal of Family Therapy, 32*(3), 232–258. https://doi.org/10.1111/j.1467-6427.2010.00507.x

Stratton, P. (2005). *The Evidence Base of Family Therapy and Systemic Practice.* Association for Family Therapy, UK. https://www.nfft.no/media/1258/the-evidence-base-of-family-therapy-and-systemic-practice-960.pdf

Sundell, K., Soydan, H., Tengvald, K., & Anttila, S. (2010). From Opinion-Based to Evidence-Based Social Work: The Swedish Case. *Research on Social Work Practice, 20*(6), 714-722. https://doi.org/10.1177/1049731509347887

Sundell, K., Ferrer-Wreder, L., & Fraser, M. W. (2014). Going Global: A Model for Evaluating Empirically Supported Family-Based Interventions in New Contexts. *Evaluation & the Health Professions, 37*(2), 203–230. https://doi.org/10.1177/0163278712469813.

Svenlin, A.-R. (2020). *Kannatteleva ja jännitteinen tukiperhetoiminta. Lastensuojelun tutkiperhetoiminnan käyttöteoria ja tukisuhteet. [Enabling and Complex Support Family Intervention. Practice Theory and Relationships of Support Family Practice in Child Welfare]* [Academic Dissertation, University of Jyväskylä].

Tapola-Haapala, M. (2011). *Sosiaalityön ammattilaiset refleksiivisinä toimijoina. Tutkimus erikoistumiskoulutuksesta Margaret Archerin teorian valossa [Social Work Professionals as Reflexive Agents – An Examination of Specialized Studies in Social Work in the Light of Margaret Archer's Theory]*[Academic Dissertation, University of Helsinki].

Thyer, B. A. (2013). Evidence-Based Practice or Evidence-Guided Practice: A Rose by Any Other Name Would Smell as Sweet [Invited Response to Gitterman & Knight's "Evidence-Guided Practice"]. *Families in Society, 94*(2), 79-84. https://doi.org/10.1606/1044-3894.4283

Tilbury, C., Hughes, M., Bigby, C., & Fisher, M. (2021). Research End-User Perspectives about Using Social Work Research in Policy and Practice. *The British Journal of Social Work, 51*(4), 1186–1202. https://doi.org/10.1093/bjsw/bcab022

Tilley, N. (2016). EMMIE and engineering: What works as evidence to improve decisions? *Evaluation, 22*(3), 304-322. https://doi.org/10.1177/1356389016656518

Titmuss, R. (1968). *Universalims versus Selection. Commitment to Welfare*. Allen and Unwin.

Titmuss, R. (1974). *Social Policy. An Introduction*. Allen and Unwin.

Tomm, K. (1988). Interventive interviewing: Part III. Intending to ask lineal, circular, strategic, or reflexive questions? *Family Process, 27*(1), 1–15. https://doi.org/10.1111/j.1545-5300.1988.00001.x

Trowler, I., & Goodman, S. (2012). Conclusion and Reflections on Our Journey. In S. Goodman, & I. Trowler (Eds.), *Social Work Reclaimed. Innovative*

Frameworks for Child and Family Social Work Practice (pp.160–169). Jessica Kingley Publishers.
Turnell, A., & Murphy, T. (2017). *Signs of Safety. Comprehensive Briefing Paper* (4th Edition). Elia International Ltd. https://knowledgebank.signsofsafety.net/resources/introduction-to-signs-of-safety/signs-of-safety-comprehensive-briefing-paper/signs-of-safety-comprehensive-briefing-paper-en/signs-of-safety-comprehensive-briefing-paper?
Turnell, A., & Edwards, S. (1999). *Signs of Safety: A safety and solution oriented approach to child protection casework*. WW Norton.
Tuomioja, E. (1996). *Pekka Kuusi. Alkoholipoliitikko. Sosiaalipoliitikko. Ihmiskuntapoliitikko [Pekka Kuusi. The researcher of alcohol policy, social policy, and humankind]*. Tammi.
Urk, F., Grant, S., & Bonell, C. (2016). Involving stakeholders in programme theory specification: discussion of a systematic, consensus-based approach. *Evidence & Policy, 12*(4), 541–557. https://doi.org/10.1332/174426415X14474260456850
Van Aerschot, L., Mathew Puthenparambil, J., Olakivi, A., & Kröger, T. (2021). Psychophysical burden and lack of support: Reasons for care workers' intentions to leave their work in the Nordic countries. *International Journal of Social Welfare, 00*, 1– 14. https://doi.org/10.1111/ijsw.12520
Vinnerljung, B., Sundell, K., Löfholm, C., & Humlesjö, E. (2006). Former Stockholm child protection cases as young adults: Do outcomes differ between those that received services and those that did not? *Children and Youth Services Review, 28*, 59-77. https://doi.org/10.1016/j.childyouth.2005.02.009
Webster-Stratton, C., & Reid, M.J. (2010). The Incredible Years Parents, Teachers, and Children Training Series: A multifaceted treatment approach for children with conduct disorders. In J.R. Weisz & A.E. Kazdin (Eds.), *Evidencebased psychotherapies for children and adolescents* (pp. 194–210). Guilford Press.
Weiss, R. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation, 76*, 41-55. https://doi.org/10.1002/ev.1086
Westhorp, G. (2018). Understanding mechanism in realist evaluation and research. In N. Emmel, J. Greenhalg, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing Realist Research* (pp. 41-57). Sage.
Whittaker, C. E., Forrester, D., Killian, M., & Jones, R. K. (2016). Can we reliably measure social work communication skills? *International Journal of Child and Family Welfare, 17*(1/2), 47–63.
Wiener, N. (1961). *Cybernetics*. MIT Press.
Wilkinson, P., Muqweni, E., Broadbent, A., Bishop, R., Akister, J., & Bevington, D. (2014). *Social work – working for families. A service transformation by Cambridgeshire county council children's services*. Cambridgeshire and Peterborough NHS Trust.

Wilson, S., Hean, S., Abebe, T., & Heaslip, V. (2020). Children's experiences with Child Protection Services: A synthesis of qualitative evidence. *Children and Youth Services Review, 113*, Article 104974. https://doi.org/10.1016/j.childyouth.2020.104974

Wong, G., Westhorp, G., Manzano, A., Greenhalg, J., Jagosh, J.. & Greenhalg, T. (2016). RAMESES II reporting standards for realist evaluations. *BMC Medicine, 14*, Article 96. https://doi.org/10.1186/s12916-016-0643-1

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J.. & Pawson, R. (2013). RAMESES publication standards: realist syntheses. *BMC Medicine, 11*, Article 21. https://doi.org/10.1186/1741-7015-11-21

Yliruka, L.. & Tasala, T. (2022). *Lastensuojelun systeemisen toimintamallin kansallinen tilannekuva vuonna 2021 [The situation of the Systemic practice model for child protection in Finland in 2021]*. Policy brief 9/2022. Finnish Institute for Health and Welfare.

Other references

Ministry of Social affairs and Health (2020) STM päätti ensimmäisistä yliopistotasoisen sosiaalityön tutkimuksen tutkimusrahoituksista [STM decided on the first university-level social work research funding]. Press release 31.12.2020. https://valtioneuvosto.fi/-/1271139/stm-paatti-ensimmaisista-yliopistotasoisen-sosiaalityon-tutkimuksen-tutkimusrahoituksista-1

Ministry of Social affairs and Health (2021). STM päätti rahoittaa yhdeksää yliopistotasoisen sosiaalityön tutkimusta [Ministry of Social affairs and Health decided to fund nine university-level social work research projects]. Press release 4.8.2021. https://stm.fi/-/stm-paatti-rahoittaa-yhdeksaa-yliopistotasoisen-sosiaalityon-tutkimusta

Ministry of Social affairs and Health (2022). STM rahoittaa yhdeksää sosiaalityön yliopistotasoista tutkimusta vuonna 2022 [Ministry of Social affairs and Health funds nine university-level social work research projects in 2022]. Press release 1.7.2022. https://stm.fi/-/stm-rahoittaa-yhdeksaa-sosiaalityon-yliopistotasoista-tutkimusta-vuonna-2022

ToT training material (2017). II module 11.-12.9.2017. Unpublished. Finnish Institute for Health and Welfare.

Ministry of Finance (2020). *Valtiovarainministeriön kestävyysvajelaskelmien menetelmäkuvaus* [Ministry of Finance's method description of the calculations of the fiscal sustainability gap]. https://vm.fi/documents/10623/20793078/Valtiovarainministeri%C3%B6n+kest%C3%A4vyysvajelaskelmien+menetelm%C3%A4kuvaus.pdf/b98cf12f-bfd4-085f-159e-c2bcf5792293/

Yle News 28.9.2017. Sari Virtanen. *Itä-Lontoon opeilla lasten huostaanotot vähenivät Mäntsälässä ja Pornaisissa – nyt suuret suomalaiskunnat aikovat seurata perässä*

[Mäntsälä and Pornainen reduced the number of care orders by applying lessons learned from East London – now the bigger cities in Finland are following their example]. https://yle.fi/uutiset/3-9852664

APPENDICES

Appendix 1a

4.10.2017

Tiedote asiakkaille lastensuojelua koskevasta tutkimuksesta

[kunta] on mukana pilotoimassa uudenlaista lastensuojelun työskentelytapaa (ns. systeeminen lastensuojelun toimintamalli). Mallia pilotoidaan valtakunnallisesti osana Lapsi- ja perhepalvelujen muutosohjelmaa (LAPE) vuosina 2017–2018.

Terveyden ja hyvinvoinnin laitos (THL) ja [kunta] tutkivat mallin pilotointia. Tutkimuksessa kerätään tietoa sekä uuden että vanhan mallin mukaan työskentelevistä lastensuojelutiimeistä. Tarkoitus on vertailla sitä, mitä hyvää tai kehitettävää eri työskentelytavoissa on. Lisäksi vertaillaan sitä, miten hyvin eri työskentelytavoilla on voitu auttaa lastensuojelussa asioivia lapsia ja heidän vanhempiaan.

Tutkimusaineisto

Asiakkaita koskeva tutkimusaineisto kerätään pääosin lomakkeilla. Lomakkeita on kolmenlaisia:

- työntekijän arvio lapsesta ja vanhemmista
- lapsen tai nuoren itsearviointilomake
- vanhemman itsearviointilomake

Lomakkeissa arvioidaan lapsen, vanhemman ja perheen hyvinvointia, palvelutarpeita ja kokemuksia palveluista. Tutkimuksessa ollaan kiinnostuneita mm. seuraavista asioista:

- Missä asioissa kaipaatte apua ja tukea? Missä asioissa pärjäätte hyvin?
- Mitä perheellenne kuuluu työskentelyn alussa? Entä 4–6 kuukautta myöhemmin?
- Mitä piditte lastensuojelun työskentelystä?

Lisäksi osa asiakkaista kutsutaan yksilöhaastatteluun, jossa keskustellaan näistä samoista asioista. Haastatteluun osallistuminen on vapaaehtoista.

Lomakkeiden käyttö asiakastyössä ja tutkimuksessa

Lomakkeissa kysyttäviä tietoja on tarkoitus käyttää osana lastensuojelussa tehtävää palvelutarpeen arviointia ja asiakassuunnitelmaa.

Itsearviointilomakkeiden täyttäminen on vapaaehtoista. Lomakkeiden täyttämättä jättäminen tai niiden tutkimuskäytön kieltäminen eivät vaikuta perheesi saamiin palveluihin.

Osa kysymyksistä voi olla vaikeita tai niistä voi tulla mieleen asioita, joista olisi hyvä jutella työntekijän kanssa. Tämän vuoksi lomakkeet on tarkoitus käydä yhdessä työntekijän kanssa läpi joko samalla kun niihin vastataan tai myöhemmin.

Voit myös päättää, että ainoastaan tutkija saa nähdä lomakkeen vastaukset. Tällöin voit täyttää lomakkeen kotona, sulkea sen saamaasi kirjekuoreen ja lähettää kuoren suoraan THL:n tutkijalle.

Työntekijä merkitsee jokaiseen tutkijalle toimitettavaan lomakkeeseen vastaajan tutkimustunnuksen, jotta tutkija osaa yhdistää eri lomakkeisiin merkityt vastaukset toisiinsa. Vastaajan nimeä tai henkilötunnusta ei yhdistetä tutkimusaineistoon.

Tutkimusaineistoa käsitellään luottamuksellisesti. Yksittäistä henkilöä ei voi tunnistaa tutkimuksesta.

Työntekijän lomakkeiden tutkimuskäyttö perustuu kaupungilta saatuun tutkimuslupaan. Itsearviointilomakkeiden tutkimuskäyttö perustuu vastaajan suostumukseen. Lapsen kohdalla suostumus pyydetään lapsen huoltajilta.

Asiakas- ja tutkimusrekisterit

Työntekijän nähtäväksi annetut lomakkeet ovat osa perheestäsi kertyviä lastensuojelun asiakastietoja ja ne kuuluvat [kunta] lastensuojelun asiakasrekisteriin. Tutkijalle luovutetut lomakkeet ovat osa tutkimusaineistoa, jolle on laadittu tieteellisen tutkimuksen rekisteriseloste. Myös haastatteluaineistolle on laadittu oma rekisteriseloste. Voit halutessasi pyytää rekisteriselosteet nähtäväksesi omalta sosiaalityöntekijältäsi tai tutkijoilta.

Tutkimuksen toteuttajat ja valtakunnallinen

Tutkimusaineistoa käsittelevät ensisijaisesti tutkijat Elina Aaltio ja Nanne Isokuortti. Aineistoa käytetään THL:n ja [kunta] tutkimusprojekteissa sekä Aaltion ja Isokuortin väitöstutkimuksissa. Haastatteluaineiston keruuseen ja käsittelyyn osallistuu myös Helsingin yliopiston maisterivaiheen opiskelijoita opinnäytetöillään. [kunta] kerättävä aineisto on osa isompaa valtakunnallista tutkimusta, jossa kerätään tietoa systeemisen lastensuojelumallin kokeilusta eri puolilla Suomea.

Tutkimukseen osallistuminen on vapaaehtoista. Kaikilta tutkimukseen osallistuvilta kerätään kirjallinen suostumus.

Kiitos osallistumisestasi!

Ystävällisesti

Elina Aaltio
VTM
Tutkija, Terveyden ja hyvinvoinnin laitos
Väitöskirjatutkija, Jyväskylän yliopisto
[sähköposti ja puhelinnumero]

Nanne Isokuortti
VTM, sosiaalityöntekijä
[työtehtävä]
Väitöskirjatutkija, Helsingin yliopisto
[sähköposti ja puhelinnumero]

Appendix 1 b

TERVEYDEN JA HYVINVOINNIN LAITOS

TIEDOTE TUTKIMUKSESTA 10.11.2017

Systeemistä toimintamallia pilotoiville [kunta] lastensuojelun sosiaalityöntekijöille ja johtaville sosiaalityöntekijöille

Tutkimuksen tavoitteet ja aineiston hyödyntäminen

Useat suomalaiset kunnat [kunta] mukaan lukien pilotoivat systeemistä lastensuojelun toimintamallia (nk. Hackneyn malli) osana Lapsi- ja perhepalvelujen muutosohjelmaa (LAPE) vuosina 2017–2018. Pilotoinnin yhteydessä toteutetaan arviointitutkimus, jonka toteuttaa Terveyden ja hyvinvoinnin laitoksen tutkija ja Jyväskylän yliopiston väitöskirjatutkija Elina Aaltio sekä [työtehtävä] ja Helsingin yliopiston väitöskirjatutkija Nanne Isokuortti (allekirjoittaneet).

Tutkimuksen pohjalta tuotetaan systeemisen toimintamallin pilotointia käsittelevä raportti Terveyden ja hyvinvoinnin laitokselle sekä 1-2 tieteellistä artikkelia. Tutkimusaineistoa hyödynnetään Aaltion ja Isokuortin väitöskirjoissa, joihin artikkelit sisältyvät. Artikkeleissa arvioidaan systeemisen toimintamallin vaikuttavuutta ja mallin käyttöönottoa Suomessa tarkastelemalla LAPE-hankkeessa toteutettavia pilotteja. Aaltion väitöskirja käsittelee lastensuojelun vaikuttavuutta ja Isokuortin väitöskirja kansainvälisten sosiaalityön käytäntöjen levittämistä ja käyttöönottoa.

Aineistonkeruun toteuttaminen ja aineiston käsittely

Kaikista systeemisen mallin pilotointiin osallistuvista pilottikunnista kerätään tutkimusaineisto, joka muodostuu pilottitiimien työntekijöiltä kerättävästä kyselyaineistosta ja esimiehiltä kerättävästä lomake- ja työpaja-aineistosta. Lisäksi valituista kunnista kerätään laajempi aineisto, jonka avulla tutkitaan systeemisen mallin vaikuttavuutta. [Kunta] on yksi näistä kunnista.

[Kunta] kerättävä erillisaineisto koostuu pilotoivien tiimien ja kontrollitiimien työntekijäkyselyistä, asiakkaiden hyvinvointia kartoittavista lomakkeista (lapsen, vanhemman ja työntekijän täyttämät lomakkeet) sekä haastatteluaineistosta (työntekijöiden ja asiakkaiden haastattelut).

Muutoksen seuraamisen mahdollistamiseksi tutkimusaineistoa kerätään pilotoinnin käynnistyessä vuodenvaihteessa 2017–2018 ja uudelleen loppukeväästä-alkusyksystä 2018. Tutkimusaineistoa kerätään sekä pilottitiimistä että pilottiin osallistumattomista tiimeistä (ns. kontrollitiimit). Kontrollitiimeinä toimii sellaisia lastensuojelutiimejä, jotka eivät pilotoi systeemistä lastensuojelumallia keväällä 2018.

Tutkimuksen määrällisen aineiston keruusta ja analyysista vastaa Elina Aaltio. Haastatteluaineiston keruusta vastaa Nanne Isokuortti. Allekirjoittaneet luovuttavat keräämänsä aineistot ja niiden analyysien tulokset anonymisoituna toisilleen kummankin tutkijan käyttöön.

Lisäksi aineistonkeruuseen tulee osallistumaan Jyväskylän yliopiston sosiaalityön maisterivaiheen sekä sosiaalityön erikoistumiskoulutuksessa mukana olevia opiskelijoita. Opiskelijat osallistuvat erityisesti laadullisen aineiston keräämiseen. Tutkimusta varten haetaan [Kunta] asianomaiset tutkimusluvat.

Aineiston keruuseen osallistuville työntekijöille annetaan erikseen lisäohjeita lomakkeiden täyttämistä varten. Haastattelujen järjestämisestä sovitaan myös erikseen.

Tutkimusaineistoa käsitellään luottamuksellisesti. Aineisto raportoidaan siten, ettei yksittäistä henkilöä voi tunnistaa tutkimuksesta.

Aikataulu

Asiakkaiden hyvinvointia kartoittava lomakeaineisto kerätään kahdessa erässä. Ensimmäinen aineistonkeruujakso toteutetaan vuoden 2018 alussa pilottitiimien käynnistettyä asiakastyönsä. Toinen aineistonkeruujakso toteutetaan loppukeväällä tai alkusyksyllä 2018. Kumpikin aineistonkeruujakso kestää noin kaksi kuukautta, minkä aikana täytetään lomakkeet osasta asiakastapauksia. Otoskoon määrä ja lomakkeiden täyttöprosessi käydään tiimien kanssa läpi erikseen joulukuussa 2017.

Haastatteluaineiston keruu käynnistyy loppuvuodesta 2017, jolloin järjestetään pilottiin osallistuvien tiimien sosiaalityöntekijöille ja tiimien johtaville sosiaalityöntekijöille puolistrukturoidut fokusryhmähaastattelut. Nämä haastattelut toistetaan vuoden 2018 aikana.

Myös asiakkailta kerätään haastattelututkimusaineisto vuoden 2018 aikana. Asiakkaiden etsimiseen toivotaan työntekijöiden apua. Aineistonkeruun ajankohdat tarkentuvat pilotin aikataulun mukaan.

Seuraavassa osiossa on tietoa työntekijöiden ja johtavien sosiaalityöntekijöiden haastatteluista. Määrällisen aineiston sekä asiakkailta kerättävän aineiston keruusta informoidaan erikseen myöhempänä ajankohtana.

Työntekijöiden ja johtavien sosiaalityöntekijöiden fokusryhmähaastattelujen toteutuksesta

Haastattelijoina toimivat Jyväskylän yliopiston maisterivaiheen sosiaalityön opiskelijat Jonna Haarala (jonna.r.haarala(at)student.jyu.fi) ja Tuija Kauppinen (tuija.i.kauppinen@student.jyu.fi). Haarala toteuttaa syksyn 2017 haastattelut ja Kauppinen vuoden 2018 haastattelut. Yhteen haastattelukertaan on hyvä varata aikaa 3 tuntia. Haastattelut toteutetaan tiimien työskentelytiloissa. Haastattelut sisältävät kysymyksiä kahdesta teemasta: 1) oman työn ja työlle saatavan tuen arviointi ja 2) näkemykset systeemisestä toimintamallista. Tiimit saavat lisää tietoa haastatteluajoista ja -paikoista myöhemmin.

Tutkimukseen osallistuminen on vapaaehtoista. Kaikilta tutkimukseen osallistuvilta kerätään kirjallinen suostumus.

Kiitos jo etukäteen yhteistyöstänne.

Ystävällisesti

Nanne Isokuortti
VTM, sosiaalityöntekijä
[työtehtävä]
Väitöskirjatutkija, Helsingin yliopisto
[sähköposti ja puhelinnumero]

Elina Aaltio
VTM
Tutkija, Terveyden ja hyvinvoinnin laitos
Väitöskirjatutkija, Jyväskylän yliopisto
[sähköposti ja puhelinnumero]

Appendix 2.

Items in research questionnaires

Effectiveness data

KINDL-R

Item in Finnish (7–13- / 14–17-vuotiaat)	Item in English (7-13- / 14-17-year-olds)
Viime viikon aikana…	During the past week…
1.1 tunsin oloni sairaaksi	I felt ill
1.2 minulla oli päänsärkyä tai vatsakipua / minulla oli kipuja	I had a headache or tummy-ache / I was in pain
1.3 olin väsynyt ja voimaton / olin väsynyt ja uupunut	I was tired and worn-out / I was tired and worn-out
1.4 tunsin oloni vahvaksi ja kestäväksi	I felt strong and full of energy
2.1 minulla oli hauskaa ja nauroin paljon	I had fun and laughed a lot
2.2 minulla oli tylsää	I was bored
2.3 olin yksinäinen	I felt alone
2.4 tunsin oloni pelokkaaksi	I was scared
3.1 olin ylpeä itsestäni	I was proud of myself
3.2 tunsin oloni hyväksi	I felt on top of the world
3.3 olin tyytyväinen itseeni	I felt pleased with myself
3.4 minulla oli paljon hyviä ideoita	I had lots of good ideas
4.1 tulin hyvin toimeen vanhempieni kanssa	I got on well with my parents
4.2 viihdyin hyvin kotona / viihdyin kotona	I felt fine at home
4.3 meillä oli pahoja riitoja kotona	we quarrelled at home
4.4 vanhempani kielsivät minua tekemästä joitakin asioita / tunsin että vanhempani rajoittivat tekemisiäni	my parents stopped me from doing certain things / I felt restricted by my parents
5.1 leikin ystävieni kanssa / tein asioita ystävieni kanssa	I played with friends / I did things together with my friends
5.2 muut lapset pitivät minusta / ystäväni pitivät minusta	other kids liked me / I was a "success" with my friends
5.3 tulin hyvin toimeen ystävieni kanssa	I got along well with my friends
5.4 tunsin olevani erilainen kuin muut	I felt different from other children / I felt different from other people

6.1 pärjäsin hyvin koulutehtävissä	doing my schoolwork was easy
6.2 pidin oppitunneista / koulu oli mielestäni mielenkiintoinen	I enjoyed my lessons / I found school interesting
6.3 olin huolissani tulevaisuudestani	I worried about my future
6.4 pelkäsin saavani huonoja arvosanoja tai numeroita	I worried about bad marks or grades / I worried about getting bad marks or grades
Vastausvaihtoehdot: ei kertaakaan / harvoin / joskus / usein / koko ajan	*Response options: Never / Seldom / Sometimes / Often / All the time*

© Kid-KINDLR / children / Ravens-Sieberer & Bullinger/ 2000 and © Kiddo-KINDLR / adolescents / Ravens-Sieberer & Bullinger/ 2000

SCORE-15

Item in Finnish	Item in English
1. Perheessäni keskustelemme meille tärkeistä asioista	In my family we talk to each other about things which matter to us.
2. Perheenjäsenet eivät useinkaan puhu totta toisilleen	People often don't tell each other the truth in my family.
3. Perheessämme kuunnellaan kaikkia.	Each of us gets listened to in our family.
4. Eri mieltä oleminen tuntuu perheessämme hankalalta	It feels risky to disagree in our family.
5. Meidän on vaikeaa selviytyä jokapäiväisistä ongelmista	We find it hard to deal with everyday problems.
6. Luotamme toisiimme	We trust each other.
7. Perheessämme on kurja tunnelma	It feels miserable in our family.
8. Kun suutumme toisillemme, jätämme toisemme tarkoituksella huomiotta	When people in my family get angry they ignore each other on purpose.
9. Näytämme menevän kriisistä toiseen perheessäni.	We seem to go from one crisis to another in my family.
10. Kun joku meistä on poissa tolaltaan, hänestä pidetään perheessäni hyvää huolta.	When one of us is upset they get looked after within the family.
11. Asiat näyttävät aina menevän pieleen perheessäni.	Things always seem to go wrong for my family.
12. Perheenjäsenet ovat ilkeitä toisilleen.	People in the family are nasty to each other.
13. Perheenjäsenet puuttuvat liikaa toistensa asioihin.	People in my family interfere too much in each other's lives.
14. Syytämme toisiamme, kun asiat menevät pieleen.	In my family we blame each other when things go wrong.
15. Pystymme löytämään uusia keinoja selvitä vaikeista asioista.	We are good at finding new ways to deal with things that are difficult.

Vastausvaihtoehdot: Kuvaa meitä erittäin hyvin / kuvaa meitä hyvin / kuvaa meitä osittain / ei kuvaa meitä kovin hyvin / ei kuvaa meitä ollenkaan	*Response options: Describes us very well / well / partly / not well / not at all*

© AFT SCORE group www.aft.org.uk/view/score.html

Need for outsider help according to the child / parent

Item in Finnish	Item in English
Mieti perheesi tämänhetkistä tilannetta kokonaisuudessaan. Kuinka paljon sinun mielestäsi tarvitsette ulkopuolista apua?	Think about your family's overall situation at present. How much outside help do you think you need?
Vastausvaihtoehdot: ei lainkaan = 0, todella paljon =10	*Response options: none = 0, very much = 10*

Service user feedback according to the child / parent

Item in Finnish	Item in English
Voin kertoa lastensuojelun sosiaalityöntekijälle avoimesti kaikenlaisista asioista.	I can openly share all kinds of issues with my social worker
Lastensuojelun sosiaalityöntekijä on kuunnellut ajatuksiani ja toiveitani.	My social worker has listened to my thoughts and wishes.
Olen saanut lastensuojelusta tarvitsemaani apua.	I have received help from child protection.
Olen tavannut lastensuojelun sosiaalityöntekijää tarpeeksi usein.	I have met my social worker often enough.
Vastausvaihtoehdot: täysin eri mieltä = 1, täysin samaa mieltä = 5	*Response options: strongly disagree = 1, strongly agree = 5*

Need for outsider help according to the social worker

Item in Finnish	Item in English
Kun arvioit tilannetta kokonaisuudessaan, millaiseksi arvioit lapsen tämänhetkisen lastensuojelun tarpeen?	Based on your overall judgement, how much in need of child protection is the child at present?
Vastausvaihtoehdot: ei lainkaan lastensuojelun tarvetta = 0, lapsen hyvinvointi vakavasti vaarantunut =10	*Response options: no need = 0, the child's safety is seriously endangered = 10*

The number of meetings

Item in Finnish	Item in English
Viimeisen kuuden kuukauden aikana, kuinka monta kertaa	During the past six months, how many times
-lapsen asioista vastaava sosiaalityöntekijä on tavannut lasta ilman vanhempia?	-a social worker has had a meeting with the child without the parents
-lapsen asioista vastaava sosiaalityöntekijä on tavannut lasta ja vanhempaa/vanhempia yhtä aikaa?	-a social worker has had a meeting with the child and the parent(s) together?
-lapsen asioista vastaava sosiaalityöntekijä on tavannut vanhempaa/vanhempia ilman lasta?	-a social worker has had a meeting with the parent(s) without the child?
-muu lastensuojelutiimin työntekijä on tavannut lasta ilman vanhempia?	-another practitioner within the team has had a meeting with the child without the parents
-muu lastensuojelutiimin työntekijä on tavannut lasta ja vanhempaa/vanhempia yhtä aikaa?	-another practitioner within the team has had a meeting with the child and the parent(s) together?
-muu lastensuojelutiimin työntekijä on tavannut vanhempaa/vanhempia ilman lasta?	-another practitioner within the team has had a meeting with the parent(s) without the child?

SPM fidelity (in Study I)

Item in Finnish	Item in English
Onko tämän lapsen kohdalla käytetty seuraavia työskentelytapoja?	Has following methods been applied in the case at hand?
-Sukupuu	-Genogram
-Systeemisten hypoteesien tekeminen	-Formulating systemic hypotheses
-Lapsi ja/tai vanhemmat ovat itse olleet läsnä systeemisessä viikkokokouksessa	-Having the family / some of the family members present in a weekly team meeting
-Systeemisen tiimin perheterapeutti on tavannut lasta ja/tai vanhempia	-Meeting between a family therapist and the family / some of the family members
Vastausvaihtoehdot: kyllä / ei	*Response options: yes / no*

Case characteristics

Item in Finnish	Item in English
Lapsen ikä	Child's age
Lapsen sukupuoli *Vastausvaihtoehdot: tyttö / poika / other*	Child's sex *Response options: girl / boy / other*
Lapsen äidinkieli *Vastausvaihtoehdot: suomi / ruotsi / muu*	Child's first language *Response options: Finnish / Swedish / other*
Mikä seuraavista kuvaa parhaiten lapsen asumista käytännössä? *Vastausvaihtoehdot: Lapsi asuu biologisten vanhempiensa kanssa / Vanhemmat eronneet, lapsi asuu toisen vanhemman kanssa / Vanhemmat eronneet, lapsi asuu vuoroin toisen, vuoroin toisen vanhemman kanssa / Lapsi on sijoitettu kodin ulkopuolelle*	What of the following best describes child's living arrangements? *Response options: Child living with birth parents / Parents separated, child living with one parent / Parents separated, child living alternately with each parent / Out-of-home care*
Onko lapsi tietojesi mukaan ollut tätä ennen lastensuojelun asiakkaana? *Vastausvaihtoehdot: ei / kyllä, kuinka monta kuukautta?*	To your knowledge, has the child previous involvement in child protection? *Response options: no / yes, how many months?*
Kuinka kauan lapsi on ollut sinun asiakkaanasi?	For how long you have been the case-holding social worker for this child?
Vanhemman ikä	Parent's age
Vanhemman äidinkieli *Vastausvaihtoehdot: suomi / ruotsi / muu*	Parent's first language *Response options: Finnish / Swedish / other*
Vanhemman koulutus *Vastausvaihtoehdot: ei peruskoulutusta / peruskoulu / keskiasteen koulutus / korkea-asteen koulutus*	Parent's education *Response options: no education / basic education / upper secondary level / tertiary education*
Vanhemman työmarkkinatilanne *Vastausvaihtoehdot: ansiotyössä tai yrittäjä / kotona lapsen kanssa / opiskelija / työtön tai työkyvytön / muu*	Parent's labour market status *Response options: gainfully employed or self-employed / home with a child / student / unemployed or incapacitated / other*
Asema lapseen nähden *Vastausvaihtoehdot: biologinen vanhempi / isä- tai äitipuoli / muu*	Relationship with the child *Response options: biological parent / step-parent / other*
Asema lapsen toiseen vanhempaan nähden *Vastausvaihtoehdot: avio- tai avoliitossa / eronnut / eroprosessi käynnissä / muu*	Relationship with the other parent *Response options: cohabiting or married / separated / separating / other*
Asuu arvioitavan lapsen kanssa *Vastausvaihtoehdot: kyllä / ei*	Currently living with the child *Response options: yes / no*

Child's wellbeing according to the practitioner [based on a measure developed in the Association for Finnish Local and Regional Authorities, not public]

Signs of abuse or neglect [based on a measure developed in the Association for Finnish Local and Regional Authorities, not public]

Context data

Administrative survey

Team structure

Item in Finnish	Item in English
Pilotoivan tiimin kokoonpano (n)	The composition of the SPM team
-sosiaalityöntekijöitä (n)	-social workers (n)
-sosiaaliohjaajia/perhetyöntekijöitä (n)	-child or family practitioners (n)
-koordinaatori (n)	-coordinator (n)
-konsultoiva sosiaalityöntekijä (n)	-consultant social worker (n)
-kliinikkoja (n)	-clinicians (n)
-muita (n), keitä?	-others (n), who?
Tiimin käytettävissä (% viikkotyöajasta)	Availability for the SPM team (% of weekly working hours)
-kliinikko	-clinician
-koordinaattori	-coordinator
-konsultoiva sosiaalityöntekijä	-consultant social worker

Social worker survey

Team meetings

Item in Finnish	Item in English
Kuinka monen oman asiakkaasi asiaa on käsitelty systeemisessä viikkokokouksessa perheterapeutin kanssa? (n)	How many of your cases your team has discussed in the weekly meetings during the implementation period in total? (n)

Systemic practice

Item in Finnish	Item in English
Oletko käyttänyt pilotoinnin aikana seuraavia menetelmiä?	Have you used the following methods during the implementation?
-systeemisten hypoteesien asettaminen	-formulating systemic hypotheses
-sukupuu	-genogram
Vastausvaihtoehdot: En ole käyttänyt lainkaan / Olen käyttänyt yhden asiakkaan kanssa / Olen käyttänyt useamman asiakkaan kanssa / En osaa sanoa tai en tiedä, mikä tämä menetelmä on	*Response options: Not at all / With one service user / With more than one service user / I can't say or I do not recognise this technique*
Kuinka monen oman asiakkaasi kanssa olet arviosi mukaan voinut työskennellä aiempaa intensiivisemmin pilotoinnin käynnistymisen jälkeen? (n)	Since the beginning of the implementation, in how many cases you have been able to work more intensively than previously? (n)

SPM clarity

Item in Finnish	Item in English
Tiimin kokoonpano ja tiimin jäsenten roolit systeemisessä tiimissä	The composition of the team and the different roles
Systeemisen tiimikokouksen / viikkokokouksen rakenne	The structure of the weekly team meetings
Systeemisen tiimin työskentelytavat	The methods used in the weekly team meetings
Systeemisen ajattelun ja perheterapeuttisen työskentelyn lähtökohdat	The basics of systemic thinking and family therapy
Hypoteesien tekeminen	Formulating hypotheses
Sukupuun tekeminen	Drawing a genogram
Lastensuojelun erityiskysymysten (kuten päihteet, mielenterveys, väkivalta) käsittely systeemisesti	Dealing specific child protection issues (such as substance abuse, mental health, violence) systemically
Lapsen turvallisuuden ja riskien arviointi systeemisessä työskentelyssä	Assessing safety and risk in systemic practice
Vastausvaihtoehdot: 1 = ei lainkaan selvä, 5 = riittävän selvä	*Response options: 1 = not clear at all, 5 = clear enough*

Satisfaction with the training

Item in Finnish	Item in English
Koulutus antoi hyvät valmiudet systeemisen mallin implementointiin.	The training prepared me to implement the systemic model
Koulutuksesta sai selkeän käsityksen, miten mallia on tarkoitus soveltaa päivittäisessä asiakastyössä.	The training gave me a clear understanding of how to implement the systemic model in daily practice
Koulutus oli riittävän konkreettista.	The training was concrete enough
Koulutusmateriaali oli hyödyllistä.	The training material was useful
Koulutuksen pituus oli sopiva.	The length of the training was suitable
Mallin käytännön soveltamista koskevaa materiaalia oli tarpeeksi.	There was enough material on how to put the systemic model into practice
Koulutus oli hyvin toteutettu.	The training was well delivered
Vastausvaihtoehdot: 1 = ei lainkaan selvä, 5 = riittävän selvä	*Response options: 1 = not clear at all, 5 = clear enough*

Participant responsiveness

Item in Finnish	Item in English
Systeemisessä tiimissä tapahtuva yhteinen reflektointi on auttanut minua tekemään työtäni aiempaa paremmin.	The collective reflection during the weekly team meetings has helped me to do my work better.
Olen saanut systeemisestä viikkokokouksesta kaipaamaani tukea asiakastyön tekemiseen.	I have received the necessary support for practice from the team meetings.
Perheterapeutin osallistuminen viikkokokouksiin on auttanut ymmärtämään asiakasperheen toimintaa uudella tavalla.	The clinician has helped me to understand the family from a new perspective.
Perheterapeutin osallistuminen viikkokokouksiin on auttanut suunnittelemaan asiakasperheen kanssa työskentelyä uudella tavalla.	The clinician has helped me to plan how to proceed with the family in a new way.
Haluan, että tiimimme jatkaa systeemisen mallin mukaista työskentelyä.	I want our team to continue using the systemic model.
Voisin suositella systeemistä mallia kollegoilleni.	I could recommend the systemic model to my colleagues.
Mallin pilotointi on kuormittanut minua lisää.	The implementation of the systemic model has been an additional burden.
Vastausvaihtoehdot: 1 = ei lainkaan selvä, 5 = riittävän selvä	*Response options: 1 = not clear at all, 5 = clear enough*

Caseload

Item in Finnish	Item in English
Kuinka monen lapsen asioista vastaat lastensuojelun sosiaalityöntekijänä?	As a case-holding social worker, what is the total number of child protection cases you are currently dealing with?

Appendix 3.

Authorship contribution statements

Original paper II

Aaltio, Elina & Kannasoja, Sirpa (accepted). Achieving agreement on service needs in child protection. Comparing children's, mothers', and practitioners' views over time and between approaches. *Journal of Children's Services*.

The article "Achieving agreement on service needs in child protection. Comparing children's, mothers', and practitioners' views over time and between approaches" evaluated the extent of agreement and discrepancy between three groups of informants and the effect of agreement on child protection outcomes.

Aaltio was responsible for conceptualization (i.e., the original idea, formulation of overarching research goals and aims) and study design (i.e., quasi-experimental study design, data collection methods). Aaltio was responsible for the writing. Kannasoja commented the manuscript.

Aaltio was responsible for quantitative analysis. Kannasoja contributed in the analysis, and Aaltio and Kannasoja interpreted the findings and edited the article together.

Aaltio requested a research permit from research sites and informed consent from all research participants prior to the data collection.

Espoo/Jyväskylä 8.6.2022

Elina Aaltio Sirpa Kannasoja

Original paper III

Isokuortti, Nanne & Aaltio, Elina (2020). Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland. *Children and Youth Services Review*, 119. https://doi.org/10.1016/j.childyouth.2020.105647

The article ”Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland” evaluated fidelity to the SPM and the possible influencing factors with mixed methods.

Nanne Isokuortti collected and analysed qualitative data (i.e., nine focus group interviews and 20 individual interviews). Aaltio commented on interview protocols when preparing the data collection as well as qualitative findings when writing the article. Isokuortti was responsible for conceptualization (i.e., the conceptual framework and formulation of overarching research aim) and most of the writing (i.e., Introduction and its sub-sections 1.1-3, qualitative results, Discussion and Conclusions). She also edited and revised Aaltio’s writing concerning quantitative data and methods section and quantitative results.

Elina Aaltio collected and analysed quantitative data (i.e., surveys, n = 56, and administrative data). Isokuortti commented on survey when preparing the data collection as well as commented on quantitative findings when writing the article. Aaltio contributed to the conceptualization (i.e., adapting conceptual framework with realist evaluation), writing of the article (i.e., the Section 1.3 on the RSW evaluations, quantitative results) as well as edited and revised the whole draft, in particular Discussion, based on Isokuortti’s writing.

Isokuortti and Aaltio operationalised the core components of SPM, constructed the fidelity thresholds, wrote the Methods and Materials section, and interpreted the findings together. Regarding the Methods and Materials section, Isokuortti was responsible of describing the qualitative data collection and analysis and Aaltio of quantitative data collection and analysis.

The authors seeked research permits mutually. The authors had individual research plans but collaborated in planning research.

Helsinki 22.1.2021

Nanne Isokuortti Elina Aaltio

Original paper IV

Aaltio, Elina & Isokuortti, Nanne (2021). Developing a programme theory for Systemic Practice Model in children's social care. Key informants' perspectives. *Child & Family Social Work,* 27(3), 444-453. https://doi.org/10.1111/cfs.12896

The article "Developing a programme theory for the Systemic Practice Model in children's social care. Thematic analysis of key informants' perspective" formulated an initial programme theory for SPM model based on the perceptions of key informants involved in the national development and dissemination of the model.

Elina Aaltio was responsible for conceptualization (i.e., the original idea, formulation of overarching research goals and aims) and design of methodology (i.e., realist evaluation).

Nanne Isokuortti contributed to the conceptualization (i.e., formulation of specific research goals and aims) and design of methodology (i.e., data collection methods).

Elina Aaltio and Nanne Isokuortti designed data collection protocols and collected the data (i.e., workshop data, key informant interviews) together.

Aaltio was responsible for thematic analysis of the workshop data and for most of the writing (Children's social care in Finland, Reclaiming Social Work and its adaptation in Finland, Methodological framework, Data analysis methods, the descriptions of the key components and CMO configurations in the Results section, Discussion). Isokuortti contributed in the analysis, wrote Participants and procedure and completed Results (i.e., references to formal theory and comparisons with service as usual). Isokuortti was responsible for the style and structure of the article as a whole. The authors wrote Introduction, Abstract and Conclusions together. Both authors edited each others' writing and revised the article together.

Aaltio and Isokuortti requested an informed consent from all research participants prior to the data collection.

Espoo/Helsinki 8.1.2021

Elina Aaltio Nanne Isokuortti

ORIGINAL PAPERS

I

EVALUATING THE EFFECTIVENESS OF THE SYSTEMIC PRACTICE MODEL OF CHILDREN'S SOCIAL CARE – A PILOT STUDY ON CHILD- AND FAMILY-LEVEL OUTCOMES

by

Aaltio, Elina

Children and Youth Services Review

https://doi.org/10.1016/j.childyouth.2022.106732

accepted

Abstract

The Systemic Practice Model (SPM) is a Finnish adaptation of the Reclaiming Social Work (RSW) model, which incorporates systemic ideas and practice into children's social care. This study is the first attempt to evaluate the effectiveness of the RSW model outside England. The study employed a quasi-experimental repeated measures design. Questionnaires assessing child- and family-level outcomes and the quantity (i.e., number of meetings) and quality (i.e., service user feedback) of the practice were administered to social workers, children and parents twice over a six-month follow-up period. Outcome data comprise 112 cases (SPM cases n=56, service-as-usual cases n=56) from 18 child protection teams (SPM teams n=9, service-as-usual teams n=9) at three sites. The overall need for child protection decreased across the sample during the follow-up period. While the intensiveness of practice was higher in the SPM group, no significant differences were found between the study groups in practice outcomes or service user satisfaction with child protection. The limitations and implications of the study for future research are discussed.

Keywords: Systemic Practice Model, Effectiveness, outcomes, pilot study, children's social care

1. Introduction

During the past decade, children's social care and child protection systems have been criticised as inefficient and dysfunctional. For instance, the demands of bureaucracy have reduced practitioners' capacity to work directly with service users in England (Munro, 2011). To improve services, Munro (2011) suggested that, to work effectively with children, young people and families, the child protection system should value and develop professional expertise. In Finland, similar demands were expressed after the high-profile death of an 8-year-old girl under a child protection plan in 2012. Following this event, the Finnish government commissioned a report on the state of children's social care (Kananoja et al., 2013). The problems identified included excessive bureaucracy, high staff turnover, poor management, insufficient early interventions, and high caseloads. In addition, service-user participation and a comprehensive understanding of families' needs were poor. As part of a broader programme to reform child and family services, the Finnish Government funded the development and dissemination of the Systemic Practice Model (SPM), a Finnish adaptation of the *Reclaiming Social Work* (RSW) model originally developed in England (Goodman & Trowler, 2012). The aim of these models is to incorporate systemic ideas and practice into children's social care. The SPM has multiple goals in relation to improving the quality and quantity of practice and the wellbeing of children and families in children's social care services (Lahtinen et al., 2018). In Finland, 31 municipal children's services sites implemented the SPM between Autumn 2017 and Summer 2018.

The purpose of this study was to assess the child- and family-level outcomes of the SPM during the first stage of its implementation. Despite the promising results of

the RSW in improving the quality of local practice (Bostock et al., 2017; Bostock et al., 2019; Forrester et al., 2013), no evidence exists on the effectiveness of the original RSW model or its adaptations either within or outside England. The present study addresses this research gap. This pilot study compared the effects of systemic practice in an SPM group with changes in a service-as-usual control group in three sites participating in the national implementation project in 2018. The pilot study also tested the outcome measures, data collection strategy and study design to gain information for a future full evaluation of the SPM.

For this study, hypotheses on the expected outcomes of the SPM were initially formulated based on a description of the SPM (Lahtinen et al., 2018) and then discussed in two SPM-related steering groups, one in charge of the national evaluation and one in charge of the national training of trainers. The main hypothesis was that systemic practice should improve family dynamics and the subjective well-being of a child, and decrease child abuse and neglect and the overall need for child protection. The systemic approach was expected to achieve these goals by improving the quantity and quality of child protection practice. The service was to be relationship-based. Social workers were expected to meet children and families more frequently than usual, and families should feel that they had been part of the process, could trust their social worker, and had received help.

The research questions were: What changes were found regarding the child's wellbeing, safety (i.e., signs of abuse or neglect, need for child protection), family dynamics, and the quality and quantity of services during the follow-up? What differences, if any, were observed between the two study groups, i.e., families supported by the SPM and families receiving service as usual?

1.1 The Finnish and English systems of children's social care

The Finnish child protection system has been described as a family service system with 'a best interest of the child' focus whereas England has a hybrid system that is risk-oriented but also supports vulnerable families (Gilbert et al., 2011). The Finnish system is deregulated, with more professional discretion in decision-making than the English system (Berrick et al., 2015). An international survey (Berrick et al., 2016) comparing time use and institutional support found that decision-making more frequently involved coworkers or a team in Finland than in England. However, Finnish social workers had over double the caseload and experienced higher work pressure than their English counterparts.

1.2 Reclaiming Social Work

The original RSW model was developed in the London Borough of Hackney, starting in 2007. It is not licensed or manualised, but a description of the model's general ideas, values, theories and systemic practice has been published by its co-founders (Goodman & Trowler, 2012). According to Trowler & Goodman (2012, pp. 14-25), the RSW changed the whole services system by establishing a shared value base (e.g., collaboration and respect) and creating structures and systems that facilitate effective working with families, improve leadership and the recruitment of competent staff.

The RSW included a training programme designed to enhance the skills and knowledge base of the staff, and introduced the use of small multi-disciplinary units. These systemic units consist of a social worker, a children's practitioner, a family therapist, a unit coordinator providing administrative support, and a consultant social worker, who leads the unit and has overall case responsibility. Units hold weekly meetings in which they discuss all the children allocated to the unit. These meetings serve as the main mechanism for case supervision. Practitioners are trained to embed the systemic approach in their social work, while the family therapist ensures that the approach is maintained (Pendry, 2012). The key idea of systemic practice is that problems are embedded in relationships and that family members' actions are a response to others' actions. In this relational frame, the social worker is not expected to fix the family but instead resolve the problem with members of the family's system.

Based on their observational study, Forrester et al. (2013) outlined six features that distinguish the original Hackney units from conventional teams: shared work, in-depth case discussion, a shared systemic approach, skills development, special roles, and low caseloads. In other agencies in England, the RSW has not been implemented in quite the same way. For example, systemic units have been replaced by larger systemic teams with three to four case-holding social workers and, in some cases, without a clinician (Bostock et al., 2017). The RSW is sometimes used alongside another practice model, Signs of Safety (Baginsky et al., 2020).

1.3 The Finnish Systemic Practice Model

The Finnish adaptation of the RSW model was outlined in 2016 during a workshop organised by the Finnish Institute for Health and Welfare (THL). The core values and ideas of the Systemic Practice Model are based on the original RSW model (Goodman & Trowler, 2012). In the SPM, teams consist of a consultant social worker, a family therapist, a unit coordinator, social workers and occasionally also child or family practitioners (Lahtinen et al., 2017). The recommended team size was between six and eight members in total. The content of the SPM was fleshed out in more detail during the Training of Trainers programme, which included an introduction to systemic thinking and family therapy, the composition and functioning of a systemic team, the structure of systemic weekly meetings, principles of systemic practice, and an introduction to selected methods and techniques. The training applied the description of systemic family therapy proposed by Leeds Family Therapy Research Center and reported in Lorås et al. (2017). According to this definition, systemic family therapy involves 11 specific competences: focusing on the system rather than the individual; seeing behavioral patterns within systems as circular and always evolving, identifying connections between circular patterns of behaviour and connections between beliefs and behavior within systems; paying attention to narratives and language; applying the ideas of constructivism and social constructionism; considering the importance of cultural context; taking a reflexive stance toward power differentials; acknowledging that reality is constructed between the therapist and the service user; applying self-reflexivity; and focusing on strengths and solutions (Lorås et al., 2017). However, it was acknowledged that child protection is not therapy. As applied in the child protection context, systemic practice involved collaborating and sharing expertise

with families and other professionals, and formulating goals and creating meanings together with families while acknowledging the social worker's authority position.

The SPM training programme did not follow the RSW training curriculum. Hence, the content of these models differed. While both models shared similar systemic ideas and techniques (e.g., genograms, formulating hypotheses), the Finnish training programme included new tools and ideas (such as the possibility to invite families to weekly team meetings) and lacked some of the content of the original RSW model (e.g., social learning theory). In addition, the concrete measures that were taken to disseminate the SPM were targeted at the team level whereas the original RSW was a reform of the whole system.

Based on an initial process evaluation (Author's own, 2019), the majority of the social workers participating in the first stage of SPM implementation were willing to continue with the model. However, the implementation process was in many ways challenging, as, for example, caseloads were high (on average, 37 children per caseholding social worker) and the training lacked specificity. A study (Author's own, 2020a) evaluating the implementation fidelity of the SPM revealed high variability across sites and individual teams. Shortcomings in training, lack of resources and leadership, and imprecision of the SPM all hindered its implementation. Overall, however, practitioners were very satisfied with the SPM and valued the contribution of a family therapist in team meetings. Based on these findings, it was expected that the SPM would already have had positive effects on practice even if it had not been fully implemented when the data for this study were collected.

1.4 Previous outcome evaluations of the RSW model and its adaptations

Previous evaluations of the RSW model have concluded that its implementation increases the quality of children's services compared to the conventional approach. During implementation, social workers in the original RSW units in Hackney expressed greater satisfaction with their work environment, social-work processes and work-related wellbeing, while the number of looked-after children decreased by 30 percent (Cross et al., 2010). In their realist-informed mixed-method evaluation conducted in Hackney, Forrester et al. (2013) found that the practitioners in the RSW units spent more time with families, provided them with more intensive help, demonstrated high levels of direct practice skills, and made high-quality assessments compared to those in conventional teams. In a study that quantified and paired observations of supervision (n = 14) with observations of direct practice (n = 18), Bostock et al. (2019) found a statistically significant association between systemic supervision and high quality practice with families. However, as no full evaluation of the effectiveness of the RSW model has been conducted, the evidence base on the effectiveness of the RSW remains limited (Author's own, 2020b).

2. Material and methods

2.1 Study design and conditions

This study used a quasi-experimental, repeated measures design. The time between the baseline and follow-up measurements was approximately six months. The aim was to compare changes in child protection outcomes and process indicators between SPM teams and service-as-usual (SAU) teams.

The development and dissemination of the SPM was organised by the Finnish National Institute for Health and Welfare (THL). THL also conducted a national evaluation of the dissemination and local implementation of the SPM. The decision to adopt RSW in Finland was made within a government-funded project (LAPE 2016-2018) aimed at reforming Finnish child and family services as a whole. The author of the study did not participate in any way in the decision-making process or LAPE preparation phase. It was not until the decision of the dissemination of the SPM had been made that, as part of the national evaluation process, the author prepared the research plan for this study.

During 2017-2018, THL trained fifty-eight (n=58) social workers or family therapists to work in pairs as local trainers. The Training of Trainers programme consisted of seven training days and four supervision sessions. By the summer of 2018, these trainers had trained fifty-two (n = 52) local teams at 31 sites in use of the systemic practice model. While the local teams implementing the SPM covered all stages of children's social care from assessment to out-of-home care, the majority of the teams (n = 22) dealt with child protection. Local training consisted of six days of team training and subsequent group supervision sessions provided by the trainers. Next, teams implemented these components in their daily practice by holding systemic weekly meetings and applying systemic orientation and methods in their casework.

Caseloads in the SPM teams were high, with an average of 38 children per social worker, compared to the 20 children per social worker recommended by THL. However, the high caseloads of the pilot teams were comparable with those of the SAU group, where social workers were individually responsible for 37 children on average.

2.2 Sample

To compare systemic practice with SAU, two samples of families were drawn, one supported by social workers in SPM teams and the other supported by social workers in SAU teams. Data were collected from nine (n = 9) SPM teams and nine (n = 9) SAU teams at three (n = 3) municipal children's services sites.

The sites were selected from all the sites that had signed up to participate in the national implementation of the SPM by autumn 2017. The selection criteria were: 1) the organization had enough children's services teams to be able to assign some for SPM training (SPM teams), leaving others to continue providing services as usual (SAU teams); and 2) the organization would be able to implement the SPM in March 2018 at latest, enabling baseline and follow-up data to be collected by the end of 2018. Three suitable sites were identified and all were willing to participate in the study. One site, with four (n = 4) SPM teams and four (n = 4) SAU teams, was in the metropolitan area of Finland. The second site, with four (n = 4) SPM teams and two (n = 2) SAU teams, was a medium-sized regional center in Central Finland, and the third

site, with one (n = 1) SPM team and three (n = 3) SAU teams, was a medium-sized regional-center in Eastern Finland .

The SAU group comprised teams working with a similar population at the same site. These teams held their usual team meetings, which focused on case allocation and general administrative tasks. The SAU teams did not receive any particular training and did not implement any specific model to guide their practice during the study.

In two of the participating organisations, the SPM and SAU teams worked with child protection cases. In one organization, the SPM and SAU teams worked with both child protection cases and children in need of less demanding social care. In this organisation, only the child protection cases received systemic practice and participated in the study.

The socio-demographic characteristics of the study groups are presented in Table 1. The only significant difference between the study groups was that the children in the SAU group were on average older (U = 2023, p = .004).

Table 1. Baseline (T1) socio-demographic characteristics of the children in the SPM and SAU groups participating in the study.

	Total *N* = 112	SMP group *n* = 56	SAU group *n* = 56	Significant difference SPM vs. SAU group *p*
Age				
Median (Range)	13.0 (1-17)	11.0 (1-17)	13.5 (5-17)	.004
Sex				
Female	56 (50 %)	33 (59 %)	23 (41 %)	ns.
Male	56 (50 %)	23 (41 %)	33 (59 %)	
First language				
Finnish	103 (93 %)	54 (96 %)	49 (89 %)	ns.
Other	8 (7 %)	2 (4 %)	6 (11 %)	
Living arrangements				
Child living with birth parents	28 (25 %)	12 (21 %)	16 (29 %)	ns.
Parents separated, child living with one parent	65 (58 %)	36 (64 %)	29 (52 %)	
Parents separated, child living alternately with each parent	7 (6 %)	4 (7 %)	3 (5 %)	
Out-of-home care	8 (7 %)	2 (4 %)	6 (11 %)	
Other	4 (4 %)	2 (4 %)	2 (4 %)	

2.3 Procedure

The study used purposive sampling. Social workers were asked to fill in assessment questionnaires concerning the families they were actively working with at the time of the first data collection. To decrease selection bias, social workers were instructed not to exclude cases they might perceive as challenging, such as non-Finnish-speaking or un-cooperative families. Both new and older cases were included in the study due to the small number of new cases per team per month. Randomisation of cases was not feasible, as the SPM was being implemented under real-life circumstances, meaning that families were allocated to social workers based on their home address. Interfering in the official allocation protocols would have entailed high-level decisions beyond the scope of this project.

The first data collection (T1), a period of three months, was conducted at each site in spring 2018 after the SPM team had been trained in systemic practice and the weekly team meetings with a clinician had started. Prior to the data collection, the researcher visited each participating team twice to give oral and written instructions on how to select cases, fill in the research questionnaires and recruit service users to the study. First, the social workers were asked to fill in case assessment forms based on their direct observations, discussions with the family and any information they possessed on the child and the child's parents. They were then asked to deliver the research materials, i.e., the self-assessment forms, information letters and consent forms, to the child and the child's parents. In some cases, this task was performed by family practitioners. This was in line with practice at the time, as most face-to-face practice was done by family practitioners rather than social workers. The practitioners

were instructed to deliver and fill in the research forms as part of their everyday practice so that the information collected from and with the families would simultaneously support their casework. Children's forms were to be completed during a meeting with a social worker or a child practitioner whereas adolescents and parents could choose to complete the form alone or during a meeting.

The second stage of the data collection (T2) started approximately six months after the end of first stage. The social workers were asked to assess the same cases and deliver the research material to them. Participation in the research was voluntary for children and parents. All the children and parents who had been assessed by their social worker or family practitioner at T1 were asked to fill-in the self-assessment questionnaires at T2 irrespective of their participation at T1 or whether the follow-up assessment had been made.

Ethical approval was granted by the National Institute of Health and Welfare Research Ethics Committee (2017–09).

2.4 Instruments

Data were collected from three groups of informants (practitioners, children, and parents) with questionnaires on wellbeing and the need for child protection and on the number of meetings and service-user feedback. The social workers' questionnaire included items on child and parent demographics, child protection process indicators (i.e., duration of the client relationship, number of meetings, interventions), and selected outcome measures. The self-assessment questionnaires included subjective outcome measures and, at T2, feedback items on the child protection process. Two versions of the children's questionnaire were in use, one for 7- to 12-year-olds and one for 13- to 17-year-olds. Both versions included the same item content but with simpler wording for the younger participants.

2.5 Measures

Child subjective well-being was measured by KINDL-R, a generic instrument for assessing health-related quality of life in children and adolescents (Ravens-Sieberer & Bullinger, 1998). This study used two versions of the questionnaire, one for 7- to 12-year-olds and one for 13- to 17-year-olds. Both versions contain 24 items with 5-point response scale: never, seldom, sometimes, often, all the time. The respondent is asked how they have felt during the past week with respect to six dimensions: physical well-being (e.g., "I felt ill"), emotional well-being (e.g., "I felt alone"), self-esteem (e.g., "I was proud of myself"), family (e.g., "I got on with my parents"), friends (e.g., "I played with friends"), and everyday functioning in school or nursery (e.g., "I enjoyed my lessons"). The sub-scales were combined to produce a total score ranging from 0 to 100. Higher scores indicate higher well-being. [T1: $\alpha = .84$, T2: $\alpha = .84$]

Family dynamics according to the child and to the parent were measured using the SCORE-15, an instrument developed to monitor progress and outcome in systemic family therapy (Carr & Stratton, 2017). The SCORE-15 consists of 15 items describing aspects of family functioning (e.g., "It feels risky to disagree in our family", "We trust each other"). Responses were given on five point Likert-scales ranging from 1,

describes my family very well to 5, describes my family not at all. The total score is the mean score for all the items in the scale, and thus ranges from one to five. Low scores indicate better adjustment, as in the original measure. [Mothers T1: α = .90, T2: α = .89, children T1: α = .84, T2 α = .86]

Need for outsider help according to the child and parent was measured with a single question developed for this study: "Think about your family's overall situation at present. How much outsider help do you think you need?" Answers were given on a 11-point response scale from 0 = none, to 10 = very much.

Need for child protection according to the social worker was measured with a single question developed for this study: "Based on your overall judgement, how much in need of child protection is the child at present?" Answers were given on a 11-point response scale from 0 = no need, to 10 = the child's safety is seriously endangered.

Signs of abuse or neglect were measured in the social worker questionnaire with a set of 21 items concerning different types of abuse or neglect. The wording was positive (e.g. "The child's daily activities are done in safe surroundings", "The parent treats the child age-appropriately"). Responses were given on 4-point scale (0, No signs of abuse or neglect, 1, Some signs, 2, Serious signs, 3, Can't say). Excluding value 3, these variables were combined into a sum variable ranging from 0 to 42. The measure was originally developed in a previous research project aimed at creating a structural assessment form for assessing the need for child protection (Author's own, 2015). For the present study, the phrasing and number of items was modified based on previous experience and consultation with national experts in child protection and child psychiatry. [T1: α = .88, T2 α = .89]

The number of meetings during the past six months was measured in the social worker questionnaire. Respondents were asked how many times a) a social worker and b) a child practitioner had had a meeting i) with the child without the parents, ii) with the parents without the child, or iii) with the child and the parent(s). Based on this information, the responses were recoded into four variables, two for social workers and two for child practitioners: the number of meetings the social worker/child practitioner had had with the child or with any/all family members (including all meetings with the child, parent(s) or both). These measures indicated the intensiveness of face-to-face practice.

Items on *service user feedback* were included in the follow-up self-assessment questionnaires for children and parents. A trustful relationship was measured with one question "I can openly share all kinds of issues with my social worker" and participation in the process with one question "My social worker has listened to my thoughts and wishes." Responses were given on a 5-point Likert-scale ranging from 1 = strongly disagree to 5 = strongly agree. Overall satisfaction with practice was measured with these and two further statements ("I have received help from child protection", "I have met my social worker often enough"). The score for the four variables were combined to produce a total score ranging from 4 to 20, with higher scores indicating more satisfaction with the child protection process. [Mothers α = .81, children α = .76]

SPM fidelity was measured in the social worker questionnaire with four questions. Respondents were asked whether one or several of the following methods included in the SPM had been applied in the case at hand: 1) genogram, 2) formulating

systemic hypotheses, 3) having the family/some of the family members present in a weekly team meeting, and 4) a meeting between a family therapist and the family/some of the family members. Since the idea was to tailor systemic practice case by case, practitioners were expected to apply at least one of these methods, but not necessarily all of them, in each case. Hence, a dichotomous variable was computed indicating whether at least one of the listed methods had been applied in a given case. Given the small sample size, a scale variable indicating the dose of systemic practice was not used in this study.

2.6 Sample Size and Attrition

A total of 65 cases were assessed by social workers at both T1 and T2. The rate of attrition between T1 and T2 was 42 percent. In 10 cases, the case had been closed or referred to another team before T2. The remaining attrition was related to staff changes. These resulted in gaps in the data collection due to poor motivation to participate in the study amid the difficulties of implementating the SPM or to experiencing a stressful work situation.

Missing data analyses showed that the data missing from the social workers' forms were not related to study group or the child's age, sex, language or living arrangements. The children's questionnaire was completed at both measurement points by 50 children, or half of all the eligible children aged 7-17 (n = 100). The data missing from the children's forms were not related to study group or child demographics. A total of 43 mothers completed the parent self-assessment questionnaire at both measurement points. Non-Finnish-speaking mothers were less likely to participate at both measurement points (X^2 (1) = 4,09, p = .043). The data missing from the mothers' forms were not related to study group, child demographics or mother's age, education, employment status, relationship with the father (i.e., marriage, cohabitation or divorced) or relationship with the child (i.e., birth parent or step-mother). A total of nine (n = 9) fathers completed the parent self-assessment questionnaire at both measurement points. Here, the missing data were related to living arrangements: in cases where children lived with both birth parents, fathers were more active than in cases in which parents were separated or a child had been placed in out-of-home care (X^2 (3) = 13.9, p = .003). Additionally, fathers who were cohabiting with or married to the child's mother participated more often than divorced fathers (n=52)(X^2 (1) = 11.9, p = .001). Fathers who did not participate throughout the study were older (M = 44.9, SD = 9.3) than participating fathers (M = 38.3, SD = 4.9, t (17.7) = 3.254, p = .004). Fathers' missing data were not related to study group or other child or father demographics. Due to the small number of fathers in the study groups (n = 3 in SPM group and n = 6 in SAU group), fathers' subjective outcome and feedback measures were omitted from the analysis.

Additional attrition analysis conducted to explore whether missing data were associated with the primary outcome measures at T1 showed that the missing data were related to only one of the outcome variables at baseline: the amount of neglect and abuse assessed by social worker was higher at T1 among those participating at both measurement points (M = 9.2) compared to those participating only at T1 (M = 6.8) (t(108.50) = 2.28, p = .025). This finding indicates that the cases participating at

both measurement points were more severe than those participating only at the baseline. While it is possible that a number of the less severe cases had already been succesfully closed by the time of the T2 data collection, this could not be confirmed as no information on the reason for dropping out was available in 37 cases. Only six of the ten cases for which the reason for attrition was known were reported as closed before the follow-up. Three of the ten cases had been referred to out-of-home care and in one case the family had relocated.

2.7 Statistical analyses

The analysis were carried out in the following steps. First, missing data analysis was peformed using Chi-square-test for nominal variables and independent sample t-tests for continuous variables. Second, baseline differences between the two groups were tested using t-test or the Mann-Whitney U-test for continuous variables and Chi-square-test for nominal variables. Third, to answer the research question on changes in outcomes, the effect of time and group, and the time*group interaction term for the variables of interest was tested using repeated measures analysis of variance. In cases where the assumption of normality was not met, additional non-parametric tests, i.e., the Mann-Whitney U-test and Wilcoxon signed-rank test, were conducted. Fourth, to examine the differences in the frequency of meetings and service user feedback between the study groups at T2, the Mann-Whitney U-test was applied. To elaborate the association between the number of meetings and feedback variables, Pearson's Product-Moment correlation was used. The SPSS Statistics 26 package was used for all analyses.

3. Results

3.1 Preliminary analysis

The baseline characteristics of the study groups with respect to the child protection process indicators and outcome scores are shown in Table 2. The only significant difference between the groups was in re-referrals, with more re-referred cases in the SAU group ($X^2(1) = 7.80$; $p = .005$).

Table 2. Baseline characteristics in the SPM and SAU groups.

	Total (N=112) *n (%)*	*SPM group (n=56)* *n (%)*	*SAU group (n=56)* *n (%)*	Significant difference SPM vs. SAU group *p*
Re-referred case	67 (60)	26 (47)	41 (73)	.005
	Total *M (SD)*	*SPM group* *M (SD)*	*SAU group* *M (SD)*	*p*
Length of current episode, months	15.0 (22.4)	15.1(21.5)	14.9 (23.4)	ns.
Total length of previous involvement in child protection before current episode, months	39.1 (35.7)	45.3 (39.4)	35.0 (33.0)	ns.
Need for child protection (0-10) according to the SW	6.2 (2.4)	6.6 (2.1)	5.9 (2.6)	ns.
Signs of abuse or neglect according to the SW	8.2 (5.9)	8.0 (5.4)	8.4 (6.5)	ns.
Subjective well-being according to the child	69.8 (12.9)	72.3 (12.6)	68.2 (13.0)	ns.
Family dynamics according to the child	2.0 (0.6)	2.0 (0.6)	2.1 (0.6)	ns.
Family dynamics according to the mother	2.1 (0.6)	2.2 (0.7)	2.1 (0.6)	ns.

3.2 Changes in children, parents and families in the whole sample and groups

Table 3 presents the means and standard deviations of the child- and family-level outcome variables of the SPM and SAU group for cases that participated in the study at both measurement points (T1 and T2).

3.2.1 Child's well-being

The *child's well-being according to the child* showed no effect of time (Wilks' $\Lambda = 0.98$, $F(1, 39)=0.77$, $p = .39$, $\eta_p^2 = .019$) or of time*group (Wilks' $\Lambda = 0.94$, $F(1, 39)=2.53$, $p = .12$, $\eta_p^2 = .061$).

3.2.2 Family dynamics

Family dynamics according to the child showed no significant effect of time (Wilks' Λ = 1.00, $F(1, 40)=0.02$, $p = .9$, $\eta_p^2 = .000$) or effect of time*group (Wilks' $\Lambda = 1.00$, $F(1, 40) = 0.03$, $p = .87$, $\eta_p^2 = .001$). *Family dynamics according to the mother* showed a main effect of time (Wilks' $\Lambda = 0.82$, $F(1, 41) = 9.14$, $p = .004$, $\eta_p^2 = .182$). The mean score decreased from 2.2 to 2.0, indicating an improvement in family functioning as assessed by mothers. However, no significant effect of time*group was observed (Wilks' $\Lambda = 1.00$, $F(1, 41)=0.00$, $p = 1.0$, $\eta_p^2 = .000$), indicating that both approaches were equally effective in changing family dynamics.

3.2.3 Need for help

Need for help according to the child showed no significant effect of time,(Wilks' $\Lambda = 1.00$, $F(1, 38) = 0.07$, $p = .94$, $\eta_p^2 = .000$), or effect of time*group (Wilks' $\Lambda = 1.00$, $F(1, 38) = 0.17$, $p = .69$, $\eta_p^2 = .004$). *Need for help according to the mother* also showed no significant effect of time (Wilks' $\Lambda = 0.94$ $F(1, 40) = 2.60$, $p = .11$, $\eta_p^2 = .061$) or of time*group (Wilks' $\Lambda = 0.93$, $F(1, 40) = 3.02$, $p = .09$, $\eta_p^2 = .07$).

3.2.4 Need for child protection

Need for child protection according to the social worker showed a main effect of time (Wilks' $\Lambda = 0.64$, $F(1, 63) = 34.8$, $p < .001$, $\eta_p^2 = .356$). Across the sample, the mean decreased from 6.5 to 4.6, indicating an improvement in the child's situation. However, no significant effect of time*group was observed, meaning that both approaches were equally effective in reducing the need for child protection (Wilks' $\Lambda = 0.97$, $F(1 ,63) = 1.74$, $p = 0.19$, $\eta_p^2 = .027$).

3.2.5 Observed abuse or neglect

S*igns of abuse or neglect according to the social worker* showed no significant effect of time (Wilks' $\Lambda = 0.95$, $F(1, 63) = 3.50$, $p = .07$, $\eta_p^2 = .053$), or of time*group (Wilks' $\Lambda = 0.95$, $F(1, 63) = 3.64$, $p = .06$, $\eta_p^2 = .055$).

Since the assumption of normality was not met in three cases, i.e., *need for help according to the child*, *need for child protection*, and *signs of abuse or neglect*, additional non-parametric tests were conducted and the results compared with those obtained from the parametric test. The results were consistent with the results reported above.

Table 3. Means and standard deviations of the outcome variables in the SPM and SAU groups at T1 and T2.

Variable	SPM group		SAU group	
	T1	T2	T1	T2
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Subjective well-being, child	75.2 (11.5)	74.0 (9.5)	68.0 (12.4)	72.2 (13.4)
Family dynamics, child	1.9 (0.6)	1.9 (0.6)	2.0 (0.5)	2.1 (0.6)
Need for help, child	3.0 (2.4)	2.9 (2.3)	3.3 (2.8)	3.5 (3.0)
Family dynamics, mother	2.2 (0.7)	2.0 (0.6)	2.2 (0.5)	2.0 (0.5)
Need for help, mother	5.9 (2.8)	4.6 (2.2)	5.5 (2.9)	5.5 (3.2)
Need for child protection, social worker	7.0 (1.8)	4.6 (2.5)	6.1 (2.4)	4.6 (2.9)
Signs of abuse or neglect, social worker	9.1 (6.1)	6.3 (5.6)	9.3 (7.3)	9.3 (8.1)

3.3 Effect of fidelity on the outcomes

Additional analysis was conducted to evaluate the effect of fidelity on the outcomes. Three fidelity groups were compared: i) cases in the SPM group where none of the key SPM methods had been applied, ii) cases in the SPM group where at least one of the key SPM methods had been applied, and iii) SAU cases. The results confirmed, first, that none of the key SPM methods had been applied in the SAU cases. Second, no effects of time*group were observed between the three groups in any of the outcome measures. These findings indicate that, in this sample, the application of SPM methods did not result in more effective practice.

3.4 Changes in the child protection process

The intention behind introducing the SPM was to achieve better outcomes by changing the nature and intensiveness of the child protection process. It was expected that social workers would meet children and families more frequently than usual and encourage family members to participate in the process, thereby creating a trustful relationship between the social worker and family.

Meeting children without the presence of other family members was infrequent in both groups (Table 4). However, the social workers in the SPM group had more private meetings with children than the social workers in the SAU group. Analysis of the total of all meetings between a case holding social worker and families showed that the number of meetings in the SPM cases was more than double that in the SAU cases. Moreover, the mean total number of all meetings between any child protection practitioner (social worker or family practitioner) and a family in the SPM group was nearly threefold that in the SAU group. Half (52%) of the families in the SAU group had met their child protection practitioner less than once a month, compared to only

one in ten in the SPM group (9%). These results reveal that systemic practice was more intensive than service as usual.

However, no differences were detected in service user experiences between two approaches (Table 5). In general, mothers were more satisfied than children with the child protection process. Surprisingly, the number of private meetings between the child and a social worker correlated negatively with the child's sense of trust ($r(41) = -.34$, $p = .025$). Moreover, the number of such meetings correlated negatively with mothers' sense of participation ($r(37) = -.50$, $p = .001$), sense of trust ($r(38) = -.43$, $p = .006$) and overall satisfaction with services ($r(38) = -.47$, $p = .002$). In contrast, the total number of all types of meetings did not correlate with any of the feedback variables. To elaborate this further, the correlation analysis was run separately for the two study groups. In the SPM group, correlations were found between the number of private meetings with a child and the mother's sense of participation ($r(19) = -.49$, $p = .024$) and sense of trust ($r(19) = -49$, $p = .024$). However, no correlation was observed between the child's feedback variables and the number of meetings. In the SAU group, no correlation was found between the feedback variables and the number of meetings of any type.

Table 4. Means, standard deviations and medians for variables on the number of meetings during follow-up.

	SPM group		SAU group			
	M (SD)	*Mdn*	*M (SD)*	*Mdn*	*U*	*p*
Meetings, child and social worker	2.3 (2.2)	2.0	1.1 (1.2)	1.0	321.0	.016
Total n of meetings, any/all family member(s) and social worker	11.1 (6.5)	9.0	5.0 (2.3)	5.0	166.0	<.001
Total n of meetings, any/all family member(s) and any practitioner	22.6 (17.6)	16.0	8.5 (12.2)	5.0	115.5	<.001

Table 5. Means, standard deviations and medians for variables on feedback from children and parents at T2.

	SPM group		SAU group			
	M (SD)	*Mdn*	*M (SD)*	*Mdn*	*U*	*p*
Child						
Sense of participation	3.5 (1.2)	4.0	3.9 (1.0)	4.0	243.5	.268
Sense of trust	3.4 (1.3)	3.0	3.7 (1.1)	4.0	250.0	.350
Overall satisfaction	14.3 (3.9)	15.0	15.1 (2.9)	16.0	255.5	.432
Mother						
Sense of participation	4.3 (1.1)	5.0	4.5 (1.0)	5.0	257.0	.636
Sense of trust	4.3 (1.1)	5.0	4.6 (0.9)	5.0	257.0	.457
Overall satisfaction	17.0 (3.3)	18.0	17.9 (2.7)	19.0	238.5	.300

4. Discussion

This study evaluated child- and family-level outcomes of the SPM during the first stage of its implementation at three children's services sites sites in Finland. The study was exploratory in nature given that the development strategy of the SPM was collaborative and iterative, i.e., the content and the delivery of the model was defined and negotiated during the process based on discussions in various forums, e.g., steering group of the national Training of Trainers, Training of Trainers moduls, and national workshops organised for managers responsible for local implementation. Nevertheless, the steering group in charge of the national evaluation of the SPM hoped for some estimation of the potential benefits of the SPM. The pilot study was conducted to test the outcome measures, data collection strategy and study design to gain information for a future full evaluation. Given that experimental and quasi-experimental research is rarely applied in the field of social work research (Holosko, 2010; see also Author's own, 2020b; Olsson and Sundell, 2016; Sheehan et al., 2018), only a limited number of valid instruments and data collection strategies found to be effective are currently available for the purpose of conducting an outcome evaluation. Hence, the present findings can help to improve the study design of any future full evaluation of the SPM.

Despite its limitations, this pilot study showed that it is possible to measure and detect changes in wellbeing, family dynamics, safety, and service quality. The quasi-experimental study design also showed that it is possible to compare these changes between the SPM and service-as-usual teams under real-life conditions.

Across the whole sample, the results indicate a statistically significant decrease in the need for child protection assessed by a social worker during the 6-month follow-up. However, this positive change was not observed in the service user data: the need for outside help as assessed by children and mothers did not change over time. The mothers reported a slight improvement in their family dynamics whereas the children did not. In addition, the subjective well-being of the children participating in this study did not improve during the follow-up. No significant differences were detected between the study groups in child's wellbeing, safety or family dynamics.

However, the systemic teams differed from service as usual in terms of the intensiveness of the children's social care process. On average, a case holding social worker had met the child alone twice in the SPM group and once in the SAU group during the past six months. Although this could indicate a more child-focused orientation in the SPM group, it is noteworthy that meetings of this type were rare in both groups. On the other hand, the frequency of other meetings was relatively high in the SPM group, especially when compared to service as usual. Nevertheless, the intensiveness of the process did not improve families' satisfaction with the help received. The finding that mothers were more satisfied with the process than children could be linked to the finding that children seldom met their social worker. However, the number of private meetings between a child and a case holding social worker correlated negatively with the child's trust in her or his social worker. Even more surprisingly, this variable was negatively correlated with mothers' satisfaction with services. When the study groups were analysed separately, this correlation was no longer found in the SAU group. In the SPM group, the negative correlation between the social worker's meetings with the child and mothers' sense of trust and participation remained. One possible explanation for this could be that service users have reservations about social workers building a direct relationship with the child in a context where meeting children individually is not typically practised by social workers. If so, social workers should better explain the purposes of these meetings to both children and parents in order to create trust along with greater intensiveness

The present findings suggest that systemic practice was not more effective than service as usual in improving child- and family-level outcomes. Several factors may explain this overall finding. First, it is possible that the first stage of SPM implementation did not have sufficient leverage to induce the intended changes in practice in the SPM teams. Based on a fidelity analysis performed in a parallel sub-study (Author's own, 2020a), wide variation in the content, dose and coverage of the SPM was found. While some SPM teams had adopted systemic thinking and techniques, others had not. In addition, variation was found between individuals in the same team. The key barriers to quality implementation were lack of clarity on the concrete meaning of systemic practice, insufficient training, high caseloads and staff turnover, and lack of commitment to change by organisations. On the other hand, a positive learning environment and hands-on coaching facilitated the implementation of the SPM. Some teams had been able to apply systemic practice despite the barriers.

In general, the majority of the practitioners (79%) participating in the fidelity study wished to continue using the model despite the challenges presented by its implementation. Second, it is possible that the quality of service-as-usual equals that of the SPM, even if fully implemented. However, In England, Bostock et al. (2019) found an association between truly systemic case discussion in weekly team meetings and better quality practice. This would indicate that at least some components of the practice model, e.g., case supervision, should outperform service-as-usual. Before this hypotheses can be tested in Finland, the quality of SPM implementation should improve. Third, it is possible that the outcomes of systemic practice are not observable with the outcome measures used in this study. In the field of children's social care practice and research, no consensus exists on outcomes or how to measure them (Forrester, 2017; Hood, 2019). This makes finding suitable validated instruments challenging. This study used the SCORE-15, a measure designed especially to assess outcomes in systemic family therapy. However, since systemic social work is not therapy, this measure might be inappropriate outside of the therapy context. The main outcome measure, KINDL-R, is a generic instrument for assessing health-related quality of life in children and adolescents. It may be that this type of generic measure is not sensitive enough to detect the specific needs and experiences of children in need of protection.

5. Strengths and limitations

This pilot study was the first attempt to evaluate the outcomes of applying the SPM in Finland. The study used a quasi-experimental, repeated measures design, and collected outcome data from three groups of informants. However, it has several limitations. First, the sample size was small and attrition high. The attrition analysis showed that, in this sample, the cases participating at both measurement points were more severe than those participating only at the baseline. Hence, it is possible that some of the attrition is explained by successful case closures. Unfortunately, information on the reasons for dropout was only available in 10 out of 47 cases. The data collection strategy needs improvement before a full evaluation with larger samples can be conducted, and more meaningful comparisons between sites and fidelity groups can be made. Second, the interval between the baseline and follow-up measurement was approximately six months. This was considered to be a period during which the first positive changes in the wellbeing and living conditions of a child should become visible. In future research, a third measurement point should be used to assess long-term change. Third, neither the practitioners nor families in the study groups were randomly selected. Since the intervention in question is a team-level reform in the context of statutory children's services, random assignment would have entailed high-level decisions outside the scope of this study. Fourth, SPM fidelity was only assessed with a practitioner-reported indicator listing the application of key SPM tools and methods. Thus, more thorough information on changes in practitioners' reasoning and observations on direct practice would be needed to fully assess the level of SPM fidelity and its effect on outcomes. While the adoption of systemic thinking was assessed earlier in a parallel study (Authors' own, 2020a), this was done only on the team level and not the practitioner level. Since the team-level

analysis revealed variation among individuals in the same team, this information could not be analysed in relation to the outcome data. Despite the several challenges this research presented, it is a much-needed step in accumulating knowledge about a model that has gained notable nation-wide attention and popularity.

6. Implications for future research and innovations

An important lesson of the pilot study concerns the context in which the SPM is being implemented and evaluated. According to the findings, the implementation strategy of the first stage of SPM implementation did not have sufficient leverage to change practice in a way that would produce the desired outcomes. Based on a parallel study (Author's own, 2020a), the factors that hindered SPM implementation were lack of clarity about systemic practice, insufficient training, and inadequate resources and leadership.

As noted in the Discussion, social care, unlike health care, has few examples of experimental research. For this reason, there is no tradition to guide the design of interventional studies in which the implementation of an innovation must be controlled for in order to successfully compare study groups. The argument against the use of manuals or other concrete guidelines could be interpreted as demonstrating a preference not to 'tie the hands' of practitioners and prevent them from making tailored decisions. If the aim is to improve the quality of practice based on research and evidence, the stakeholders responsible for developing new interventions for social work should reconsider their reasoning. Some degree of compromise between participative approaches and more hierarchically operated study designs is needed. Similar tendencies have also been identified in other countries. For instance, a study (Baginsky et al., 2020) on the adoption of practice models (or practice frameworks) in England found that the majority of local authorities had adopted these models. Unfortunately, most of these were applying elements of different practice frameworks, making it difficult to investigate the correlation between a specific framework and better outcomes for children.

During the data collection, it became evident that some of the social workers implementing the SPM were heavily burdened, and that this impaired their motivation to participate in the study. Some of their frustration seemed to be related to the disordered nature of the SPM implementation process – an observation supported by the findings of a fidelity analysis (Author's own, 2020a). High caseloads and recruitment problems burdened practitioners and hindered them in changing their practice. In addition, concurrent national policy reforms and local organization reforms at two of the research sites prevented managers from supporting their teams in collecting data and implementing the SPM. Even when it is not possible – or even desirable – to create a sheltered environment for implementation, it is necessary to ensure that organisations have a reasonable chance to implement a new model with sufficiently high fidelity. To improve data collection, new strategies aimed at decreasing the burden on social workers need to be considered.

Conclusions

While there is strong interest in steering social work practice in accordance with research evidence (Hodge 2012; Julien-Chinn & Lietz, 2019), empirical evidence on the current effectiveness of many social work interventions and other psychosocial interventions directed at children and families is lacking (Authors own, 2020b; Breivik et al., 2021; Sheehan et al., 2018). Pilot studies can provide useful insights regarding the next steps in research and implementation. However, to produce robust study designs and strong evidence requires effort not only from researchers but also from the stakeholders responsible for disseminating and implementing these services. It is also essential to conduct a process evaluation in parallel with the outcome evaluation to assess changes in service delivery and reflect on these findings together with changes in outcomes.

References

Author's own (2015)

Authors' own (2019)

Authors' own (2020a)

Author's own (2020b)

Baginsky, M., Ixer, G., & Manthorpe, J. (2020). Practice Frameworks in Children's Services in England: An Attempt to Steer Social Work Back on Course? *Practice: Social Work in Action*, *33*(1), 1–17. https://doi:10.1080/09503153.2019.1709634.

Berrick, J., Peckover, S. Pösö, T., & Skivenes, M. (2015). The formalized framework for decision-making in child protection: A cross-country comparison. *Journal of European Social Policy*, *25*(4), 366–378. https://doi.org/10.1177/0958928715594540

Berrick, J., Dickens, J., Pösö, T., & Skivenes, M. (2016). Time, Institutional Support, and Quality of Decision Making in Child Protection: A Cross-Country Analysis. *Human Service Organizations: Management, Leadership & Governance*, *40*(5), 451–468. doi:10.1080/23303131.2016.1159637.

Bostock, L., Forrester, D., Patrizo, L., Godfrey, T., Zonouzi, M., Antonopoulou, V., Bird, H., & Tinarwo, M. (2017). *Scaling and deepening the Reclaiming Social Work model. Evaluation Report*. London: Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625227/Scaling_and_deepening_the_Reclaiming_Social_Work_model.pdf

Bostock, L., Patrizo, L., Godfrey, T., & Forrester, D. (2019). What is the impact of supervision on direct practice with families? *Children and Youth Services Review*, 105. https://doi:10.1016/j.childyouth.2019.104428.

Breivik, K., Eng, H., Kaiser, S., Karjalainen, P., Kurki, M., Kyrrestad, H., Laajasalo, T., Martinussen, M., Merikukka, M., Peltonen, K., Dahl Pettersen, S., Rasmussen, L.-M., Reedtz, C., & Rye, M. (2021). The First 1000 Days in the Nordic Countries: Psychosocial Interventions and Psychological Tests: A Review of the Evidence. Nordic Council of Ministers, Nordic Council of Ministers Secretariat. Nord 2021:037. https://pub.norden.org/nord2021-037/

Carr, A., & Stratton, P. (2017). The Score Family Assessment Questionnaire: A Decade of Progress. *Family Process*, *56*(2), 285–301. https://doi: 10.1111/famp.12280

Cross, S., Hubbard, A., & Munro, E. (2010). *Reclaiming Social Work. London Borough of Hackney. Children and Young People's Services. Part 1: Independent Evaluation*. https://core.ac.uk/download/pdf/30322069.pdf

Forrester, D., Westlake, D., McCann, M., Thurnham, A., Shefer, G., Glynn, G., & Killian, M. (2013). *Reclaiming Social Work? An Evaluation of Systemic Units as an Approach to Delivering Children's Services. Final report of a comparative study of practice and the factors shaping it in three local authorities*. University of Bedfordshire. https://uobrep.openrepository.com/handle/10547/594517

Forrester, D. (2017). Outcomes in children's social care, *Journal of Children's Services*, *12*(2–3), 144–157. https://doi.org/10.1108/JCS-08-2017-0036

Gilbert, N., Parton, N., & Skivenes, M. (2011). *Child Protection Systems: International Trends and Orientation*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199793358.001.0001

Goodman, S., & Trowler, I. (2012). *Social work reclaimed: Innovative frameworks for child and family social work practice*. Jessica Kingsley Publishers.

Holosko, M. (2010). What Types of Designs are We Using in Social Work Research and Evaluation. *Research on Social Work Practice, 20*(6), 665–673. https://doi.org/10.1177/1049731509339586

Hood, R. (2019). What to measure in child protection? *The British Journal of Social Work, 49*(2), 466–484. https://doi.org/10.1093/bjsw/bcy066

Julien-Chinn, F., & Lietz, C. (2019). Building learning cultures in the child welfare workforce. *Children and Youth Services Review*, 99, 360–365. https://doi.org/10.1016/j.childyouth.2019.01.023

Kananoja, A., Lavikainen, M., & Oranen, M. (2013). *Toimiva lastensuojelu. Selvitysryhmän loppuraportti,* [Functioning Child Protection. Final report of the clearing group] (Ministry of Social Affairs and Health publications 2013:19). Sosiaali- ja terveysministeriö [Ministry of Social Affairs and Health].

Lahtinen, P., Männistö, L., & Raivio, M. (2017). Kohti suomalaista systeemistä lastensuojelun toimintamallia: Keskeisiä periaatteita ja reunaehtoja [Toward the Finnish Systemic Practice Model for child protection. Core principles and preconditions] (Discussion paper 7/2017). National Institute for Health and Welfare.

Leathers, S., Melka-Kaffer, C., Spielfogel, J., & Atkins, M. (2016). Use of evidence-based interventions in child welfare: Do attitudes matter? *Children and Youth Services Review*, 70, 375–382. https://doi.org/10.1016/j.childyouth.2016.10.022

Lorås, L., Bertrando, P., & Ness, O. (2017). Researching Systemic Therapy History: In Search of a Definition. Journal of Family Psychotherapy, *28*(2), 134–149. https://doi.org/10.1080/08975353.2017.1285656

Munro, E. (2011). *The Munro Review of Child Protection: Final Repots. A child-centred system*. London: Department for Education.

Olsson, T., & Sundell, K. (2016). Research that Guides Practice: Outcome Research in Swedish PhD Theses Across Seven Disciplines 1997–2012. *Prevention Science*, *17*(4), 525-532. doi: 10.1007/s11121-016-0640-9

Pendry, N. (2012). Systemic Practice in a Risk Management Context. In S. Goodman & I. Trowler (Eds.), *Social Work Reclaimed. Innovative Framework for Child and Family Social Work Practice* (pp. 26–33). London and Philadelphia: Jessica Kingsley Publishers.

Ravens-Sieberer, U., & Bullinger, M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results. *Quality of Life Research*, 7, 399–407. doi: 10.1023/a:1008853819715.

Sheehan, L., O'Donnell, C., Brand, S., Forrester, D., Addis, S., & El-Banna, A. (2018). Signs of safety: Findings from a mixed-methods systematic review focussed on reducing the need for children to be in care. What Works Centre for Children's Social Care. https://whatworks-csc.org.uk/wp-content/uploads/Signs_of_Safety_a_mixed_methods_systematic_review.pdf

Trowler, I., & Goodman, S. (2012). A Systems Methodology for Child and Family Social Work. In S. Goodman & I. Trowler (Eds.), *Social Work Reclaimed. Innovative Framework for Child and Family Social Work Practice* (pp. 14–25). Jessica Kingsley Publishers.

II

ACHIEVING AGREEMENT ON SERVICE NEEDS IN CHILD PROTECTION. COMPARING CHILDREN'S, MOTHERS', AND PRACTITIONERS' VIEWS OVER TIME AND BETWEEN APPROACHES

by

Aaltio, Elina & Kannasoja, Sirpa

Journal of Children's Services

https://doi.org/10.1108/JCS-12-2021-0052

accepted

Abstract

Purpose - While research on service users' participation and their perceptions on the quality of services exists, there seems to be a research gap on agreement between family members' and practitioners' assessments of the family's situation. This study investigated agreement and its effect on outcomes by comparing the viewpoints of three groups of informants (children, mothers, and practitioners) in the context of statutory child protection in two study groups – one applying a systemic approach (SPM) and a service-as-usual control group (SAU).

Design/methodology/approach - A quasi-experimental repeated-measures study design was applied. Outcome data comprised 112 cases (SPM cases n=56, SAU cases n=56) at three sites. Data were collected from all participants at baseline and six-months later.

Findings – First, practitioners' analyses of a child's need for protection did not meet family members' expressed need for help. Second, child-mother agreement on the need for service intervention at T1 predicted a decrease in abuse and neglect from T1 to T2 as assessed by practitioners. In this sample, no differences were found between the two groups.

Originality/value – This study highlights the importance of making explicit the viewpoints of children, parents and practitioners in casework and research to gain better understanding of how their perspectives differ over the course of the process, and how possible disagreements at the beginning affect outcomes.

Introduction

Service user's viewpoints should be considered in children's social care when assessing the need for services and service outcomes. First, according to the United Nations Convention on the Rights of the Child (CRC), children have the right to express their views and participate in decision-making process affecting their lives. Any organisation claiming to operate rights-based practice should make efforts to enable children freely express their opinions and feel that their voices are heard (Falch-Eriksen *et al*., 2021). National legislation in many countries, such as in England (Carr, 2004) or Finland (Social Welfare Act 1301/2014), requires service user participation in social care services. Second, evidence exists that the family's participation can help to improve the quality and possibly also outcomes of children's services (Gladstone *et al*., 2012; Heimer *et al*., 2018). However, in several studies service users have reported negative experiences when participating in service design and delivery (Bekaert *et al*., 2021, Falch-Eriksen *et al*., 2021).

Service user engagement and participation can refer to collective involvement in service development or to users' individual involvement in cases that concern them personally (Diaz, 2020; Kiili *et al*., 2021). In this study, we use the terms engagement, participation and involvement interchangeably to refer to the active contribution of children and parents in their own social care cases.

New models or frameworks for practice have been developed to improve the quality and outcomes of child protection, such as Signs of Safety (Turnell and Murphy, 2017) and Reclaiming Social Work (RSW) (Goodman and Trowler, 2012). While the theoretical basis and key skills of these models differ, they both promote the importance of service users' participation in improving service quality and outcomes. This study focused on the Systemic Practice Model (SPM), a Finnish adaptation of RSW, that has been widely adopted in Finland since 2017. RSW and the SPM apply the ideas and methods of systemic family therapy in statutory child and family social work. In the SPM (Lahtinen *et al*., 2017; Author's own, in review a), service user participation, relationship-based practice and power sharing are some of the key principles.

While the idea of acknowledging the limits of practitioners' knowledge underpins systemic practice, child protection authorities have a legal responsibility to assess a child's need for protection and act accordingly. In Finland (Child Welfare Act 417/2007), this responsibility is assigned to a case holding social worker (henceforth practitioner). While a child's parents (or custodians) are primarily responsible for the child's wellbeing and development, children also have a right, as laid down by the CRC, to be heard on these issues. Hence, three viewpoints should be considered when assessing a child's need for protection and services: those of the child, the practitioner, and the parent(s). Research on service users' participation and their perceptions on the quality of services (Bekaert *et al*., 2021; Falch-Eriksen *et al*., 2021) has rather neglected the issue of agreement between these three parties and the effect of such agreement on service outcomes. Forrester *et al.* (2013) explored agreement between the family and practitioner by comparing the correlations between practitioners' and parents' ratings of a group applying RSW and a service-as-usual group. They found the level of agreement on key issues in the family was higher in RSW group. However,

the sample did not include children as informants. In addition, the association with outcomes was not analysed.

This study investigated agreement and its effect on outcomes in the context of statutory child protection by comparing the viewpoints of three parties (a child, a parent, and a practitioner) in two study groups – a group applying a systemic approach (SPM) and a service-as-usual control group (SAU). Specifically, we explored informant discrepancy, i.e., the mean level differences between the participants' scores on families' service needs and family dynamics, and the extent of agreement, i.e., the inter-rater correspondence. By applying a quasi-experimental repeated-measures study design, it is possible to i) investigate change in agreement longitudinally, here over a six-month period, and ii) compare changes between study groups. Finally, we explored the relationship between informant discrepancy at baseline and child protection outcomes at follow-up. The main hypothesis was that agreement between practitioners and two family members would increase over time as the three groups of informants explored the family's problems and service needs together. Achieving some form of agreement or understanding with service users on their problems and possible solutions is one of the goals of direct social work practice (Trevithick, 2012). A second hypothesis was that agreement would be higher in the SPM group, since practitioners in this group have received additional training and support for involving families and co-reflecting with them. A third hypothesis was that the higher the level of agreement, the better the child protection outcome.

The research questions were:

1) Is there a discrepancy between the informants regarding service needs and the family's situation? If so, what differences can be detected over time and between study groups?
2) To what extent do the informants agree on the family's situation and the need for child protection? What differences occur over time between the informants and between the two study groups?
3) Is agreement at T1 related to the outcome at T2?

Service user participation

Various reasons have been given for taking service users' personal views on their well-being and service needs into account in the context of child protection services. First, engaging families in decision-making may be a prerequisite for a successful intervention. Meta-analyses on a range of therapeutic interventions for children and adolescents (Shirk and Karver, 2003) and adults (Horvath and Symonds, 1991) have found consistent, although modest, associations between engagement and positive treatment outcomes. Similar associations are expected in child protection interventions; for example, Horwitz and Marshall (2015) state that families only benefit from interventions they are committed to. This statement has been empirically supported. Gladstone *et al*. (2012) found that parents' perceived engagement was associated with positive outcomes for parents, such as satisfaction with the service, subjective perceptions of improved parenting, and feeling that their children were safer. In turn, based on coded observations of practice and outcome data collected from parents, Forrester *et al*. (2019) found that practitioners' engagement skills seemed

less important in producing positive outcomes than the skills termed by the authors as 'good authority' i.e., purposefulness, clarity about concerns, and a focus on the child. Nonetheless, the authors state that service quality and respecting service users are important even if child protection could not resolve the family's problems.

Second, according to the CRC, authorities are obliged to hear children and make decisions that are in the child's best interest. To this end, practitioners need to canvass children's opinions and preferences (Falch-Eriksen *et al*., 2021). Moreover, parents and children may frame the situation differently, especially when children are at risk due to lack of parental care. Based on their analysis of child protection referrals and care plans, Heimer *et al*. (2018) found that while children's perceptions were heard at the beginning of the child protection process, parents' views became increasingly dominant in the process from assessment to investigation and the provision of care. When an intervention was based solely on parents' perceptions, it was likely to poorly match the problem. The authors concluded that involving children improves the quality of decision-making and intervention design.

Third, collecting information both from children and parents enables practitioners to compare these viewpoints and address any disagreement. Parent-child discrepancy provides valuable information for use when working with the family and designing an intervention. Goolsby *et al*. (2018), who evaluated a psychotherapy intervention for children with psychosocial deficits, found that parent-child discrepancy was associated with poorer treatment outcomes. To improve the effectiveness of treatment, the authors suggest that practitioners should address such discrepancies and examine the factors causing them, such as family functioning, awareness of problems, motivation to be involved in treatment, and conflicting perspectives.

While many argue for service user participation and family engagement, these ideals are not necessarily met in child protection practice. In UK, for example, Diaz (2020) found that while practitioners regard children's participation as important, they lack understanding and knowledge of how to implement meaningful participation, or are too busy to do that. As a result, contrary to national guidelines on good social work practice, children were not informed of the processes where decisions regarding them are made, and they felt as not being listened or supported. A systematic review (Falch-Eriksen *et al*., 2021) on children's participation in child protection found that their perceptions are consistently overlooked in various countries. Similarly, in their metasynthesis on family members' perceptions, Bekaert *et al*. (2021) found many studies reporting negative experiences by service users of practitioners' engagement with them. For instance, parents felt practitioners held narrow pre-conceived ideas about their problems and didn't take time to hear parents' interpretations of the problem or service plan.

Systemic approaches and child protection systems in England and Finland

The Finnish Systemic Practice Model (SPM) for child protection is an adaptation of the Reclaiming Social Work (RSW) model originally developed in England (Goodman and Trowler, 2012). The English child protective system has been typologised as a hybrid system that provides support for families but is also child protection oriented,

whereas the Finnish system is primarily family-service oriented (Gilbert *et al*., 2011). The Finnish system is less regulated than the English system (Berrick *et al*., 2015), but caseloads are higher in Finland than in England (Berrick *et al*., 2016).

In England, the majority of local authorities have adapted some framework or model to guide their practice (Baginsky *et al*., 2020). While Signs of Safety is the most popular practice model in England, the government-funded reform programme for child and family services in Finland 2016–2019 was inspired by RSW. RSW combines the methods and principles of systemic family therapy with statutory child protection practice. RSW incorporates reforms from leadership to case management and frontline casework. The aim of these changes and activities is to enable relationship- and strength-based practice, where the family is seen as a system interacting with other systems, such as the extended family and professionals. Systemic practice seeks to engage the family, reflect multiple perspectives and act differently in the light of these insights (Bostock *et al*., 2019). In previous research, RSW has been associated with improved quality of child protection practice. In a mixed-methods evaluation study conducted in the original site pioneering RSW, Forrester *et al*. (2013) observed that practitioners in RSW units were more empathic and received more positive feedback from parents compared to practitioners in service-as-usual teams. Moreover, agreement between practitioners and parents about the nature and seriousness of the family's issues was higher in the RSW units. Bostock *et al*. (2019) explored new sites implementing RSW and found an association between systemic case supervision and practitioner skills, i.e., better relationship-building skills and "good authority" (Forrester *et al*., 2019).

In the SPM (Lahtinen *et al*., 2017), child protection organisations form systemic teams comprising a consultant social worker, a family therapist, a unit coordinator, 1–3 social workers and occasionally also family practitioners. These teams hold weekly meetings, where the consultant social worker and family therapist provide systemic case supervision. The purpose of these sessions is to formulate hypotheses from multiple perspectives and plan the next steps for how to proceed with the family. Compared to service as usual, practitioners should engage more in mutual reflection, share responsibility and knowledge of cases, receive case supervision more frequently, and spend more time on direct practice (Author's own, in review). In terms of direct practice with families, practitioners are expected to spend more time with the families, and apply systemic tools and methods, such as genograms and formulating hypotheses. The use of these methods should provide insight and new perspectives on families' dynamics, history, patterns, and narratives for both practitioners and service users (Author's own). While systemic practice is inspired by the ideas central to systemic family therapy (Lorås et al., 2017), the intention is not to transform social work practice into therapy. Rather, systemic practice involves collaborating and sharing expertise with families and other professionals and formulating goals and creating meanings together with families while acknowledging practitioner's authority position. The SPM is outlined as a child- and family-based model where service users' needs are the starting point for casework (Lahtinen *et al*., 2017).

In previous research, practitioners implementing the SPM during the first stage of national dissemination reported difficulty in applying the model, but considered it had potential and especially valued the contribution of a family therapist in case

discussions (Isokuortti and Aaltio, 2020). A quasi-experimental study comparing child- and family-level outcomes found no differences between SPM and SAU groups in children's safety or wellbeing, or in family dynamics (Author's own, in review). However, the SPM practitioners met families more often than their service-as-usual counterparts. While these findings suggest that SPM did not outperform conventional practice in service outcomes, it has yet been analysed whether it could improve agreement between the parties involved in child protection case work.

Method

Study design and conditions

This study used a quasi-experimental, repeated measures design. Outcome data were collected from three groups of informants: children, parents and practitioners.

The SPM was developed and disseminated by the N.N. Institute, which also conducted a national evaluation of SPM implementation and collected the data for this study as part that project.

During 2017–2018, The N.N. Institute trained fifty-eight (n=58) local trainers in use of the SPM. By summer 2018, these had trained fifty-two (n=52) local teams at 31 sites. Local training comprised six days of team training and subsequent group supervision sessions led by the trainers. The training included an introduction to systemic thinking and family therapy, the functioning of a systemic team, team roles, and systemic weekly meetings, documentation, principles of systemic practice, and an introduction to systemic methods and techniques. After training, teams began to hold systemic weekly meetings and apply systemic ideas and methods in their casework. All teams participating in this study had a consultant social worker, a unit coordinator, and a systemic family therapist participating in their weekly team meetings. However, all teams but one had more than three social workers, i.e., the teams were larger than intended.

The Institute recommended that the municipalities participating in the project reduce caseloads in the SPM teams to 20 children per practitioner. Since most of the municipalities did not have the additional resources for this, caseloads in the SPM teams remained high. In the SPM teams participating in this study, caseload was on average 37 children per practitioner as in the SAU group.

Sample

The sample comprised a group of families supported by practitioners in SPM teams (n=9) and a group of families supported by practitioners in SAU teams (n=9). Data were collected from three municipal children's service sites. One was a large city in the metropolitan area of Finland and the other two were medium-sized regional centres, one in Central and one in Eastern Finland. All teams worked with child protection cases, i.e., with children who had been assessed as needing protection but could nevertheless remain in their parents' custody.

The SAU teams did not receive any additional training and their practice was not guided by any shared model or framework. The baseline characteristics of the

study groups are shown in Table I. The only significant difference between the study groups was that the children in the SAU group were on average older (U = 2023, p = .004).

Table I. Baseline (T1) socio-demographic characteristics of the children in the SPM and SAU groups.

	Total *N*=112	SPM *N*=56	SAU *N*=56	Significant difference SPM vs. SAU *p*
Age				
Median (Range)	13.0 (1–17)	11.0 (1–17)	13.5 (5–17)	.004
Sex				
Female	56 (50%)	33 (59%)	23 (41%)	ns.
Male	56 (50%)	23 (41%)	33 (59%)	
First language Finnish	103 (93%)	54 (96%)	49 (89%)	ns.

Procedure

Purposive sampling was used. The first data collection (T1) was conducted at each site in Spring 2018 over a three-month period. Practitioners were asked to fill in assessment questionnaires pertaining to the families they were currently actively working with. The idea was to deliver and fill in these questionnaires as part of everyday practice so that the information collected from and with the families would simultaneously support casework. Adolescents and parents could choose to complete the questionnaire alone or during a meeting. Children's questionnaires were to be completed during a meeting with a practitioner. The follow-up data (T2) were collected approximately six months later. The data collected from the practitioners were used with the permission of the municipalities participating in the study. Participation in the research was voluntary for the children and their parents. To be able to use children's forms, informed consent was asked from guardians and children over 12 years old. In the case of parents' forms, parents were asked for their informed consent. All the parents and children who had been assessed at T1 were asked to fill-in the self-assessment questionnaires at T2 irrespective of their participation at T1. Ethical approval was granted by the N.N. (2017–09). All personal information collected in this study has been treated as highly sensitive.

Instruments

Questionnaire data on wellbeing and the need for service intervention were collected from three groups of informants (practitioners, children, and parents). The practitioner's questionnaire included items on child and parent demographics and outcome measures. The self-assessment questionnaires included subjective outcome measures. Two versions of the children's questionnaire were used (7- to 12- and 13- to 17-year-olds). Both versions contained the same item content but with simpler wording for the younger participants.

Measures

Key outcome measures

Need for service intervention was measured with a single-item question developed for this study. In the practitioner's questionnaire, the item was "Based on your overall judgement, how much in need of child protection is the child at present?" Answers were given on a 11-point response scale (from 0 = no need to 10 = the child's safety is seriously endangered). In the children's and parents' forms, the question was "Think about your family's overall situation at present. How much outsider help do you think you need?" Answers were given on a scale from 0 = none to 10 = very much.

Child's subjective well-being was measured by KINDL, a generic instrument for assessing health-related quality of life in children and adolescents (Ravens-Sieberer and Bullinger 1998). Both versions of the questionnaire contain 24 items with a 5-point response scale: never, seldom, sometimes, often, all the time. The respondent is asked how they have felt during the past week with respect to five dimensions: physical well-being (e.g., "I felt ill"), emotional well-being (e.g., "I felt alone"), self-esteem (e.g., "I was proud of myself"), family (e.g., "I got on with my parents"), friends (e.g., "I played with friends"), and everyday functioning in school (e.g., "I enjoyed my lessons"). For items asking about poor quality of life, the values are reversed. The sub-scales were combined to produce a total score ranging from 0 to 100. High scores indicate higher well-being. [T1: α = .84, T2: α = .84]

Signs of abuse or neglect were measured in the practitioner questionnaire with a set of 21 items on different types of abuse or neglect. The wording was positive (e.g., "The child's daily activities are done in safe surroundings", "The parent treats the child age-appropriately"). Responses were given on 4-point scale (0 = no signs of abuse or neglect, 1 = some signs, 2 = serious signs, 3 = can't say). Excluding value 3, responses were combined to form a sum variable ranging from 0 to 42. The measure was originally developed in a previous research project (Aaltio, 2015). For the present study, the wording and number of items was modified based on consultation with experts in child protection and child psychiatry. [T1: α = .88, T2: α = .89]

Other measures

Family dynamics according to the child and to the parent were measured using the SCORE-15, an instrument developed to monitor progress and outcome in systemic family therapy (Stratton et al, 2010). The SCORE-15 contains 15 items describing aspects of family functioning (e.g., "It feels risky to disagree in our family"; "We trust each other"). Responses were given on 5-point Likert-scales ranging from 1 = describes my family very well to 5 = does not at all describe my family. For items asking about poor family dynamics, the values are reversed. The total score is the mean score for all items, and ranges from one to five. Low scores indicate better adjustment. [Mothers T1: α = .90, T2: α = .89, children T1: α = .84, T2: α = .86]

Child-related case characteristics were measured in the practitioner's questionnaire with a set of 10 items on the child's physical, mental, cognitive, and social wellbeing and behaviour (e.g., "Does the child have mental health-related problems, such as

fatigue, exhaustion or tiredness?", "Does the child have any friends?"). Responses were given on a 4-point scale where 1 indicated no problems, 2 some problems, 3 serious problems and 4, can't say. The exact wording depended on the question. Responses 2 and 3 were counted as indicating problems.

Sample Size and Attrition

A total of 65 cases were assessed by practitioners at both T1 and T2 (attrition 42%). Missing data analyses showed that missing data were not related to study group or the child's age, sex, language or living arrangements. The children's questionnaire was completed at both measurement points by 50 children, i.e., 50 percent of all the children aged 7-17 (n=100) participating in the study. Missing data were not related to study group or child demographics. 83 mothers participated in the study at T1 and 43 completed the parent self-assessment questionnaire at both measurement points (attrition 48%). Non-Finnish speaking mothers were less likely to participate at both T1 and T2 (X^2 (1) = 4,09, p = .043). Missing data were not related to study group, child demographics or mother's age, education, employment status, relationship with the father (i.e., marriage, cohabitation or divorced) or relationship with the child (i.e., birth parent or stepmother). 25 fathers participated in the study at T1, and nine completed the parent self-assessment questionnaire at both measurement points. In all these cases the mother had also replied. To keep the sample more homogenous, we decided to use mothers' responses only.

Statistical analyses

The analyses were carried out in the following steps. First, baseline differences between the two groups were tested using t-test or the Mann-Whitney U-test for continuous variables and Chi-square-test for nominal variables. Second, to assess informant discrepancy (mean level differences) on the need for service intervention and the family dynamics, paired-sample t-tests were used to compare each pair of informants at T1 and T2. Third, to assess the change in this discrepancy between T1 and T2 in the whole sample and between the study groups, the difference between the group means on the need for service intervention and family dynamics were calculated. These new variables were then used to test the effects of group and time and group-by-time interaction by using repeated measures analysis of variance. Fourth, to answer the research question on the extent of agreement between the child, the child's mother and the practitioner at T1 and T2, separately, Pearson's Product-Moment correlation was used. Fifth, to examine if the extent of agreement had increased more in the SPM than SAU group, Fisher's Z-test was used. Sixth, to examine the effect of agreement on outcomes, three change scores (T2–T1) were calculated for the selected key outcome variables: the need for child protection according to the practitioner, signs of abuse or neglect as assessed by the practitioner, and the child's subjective well-being. Thereafter, multiple regression analysis was conducted. Child and mother reports separately and the interaction terms were used as predictors for change scores of all three outcome variables. Practitioner-child and practitioner-mother reports separately and the interaction terms were used as

predictors for change in signs of abuse or neglect and the child's well-being. SPSS Statistics 27 was used for all analyses.

Results

Preliminary analysis

The baseline characteristics of the study groups with respect to the child-related measures are shown in Table II. No significant between-group differences were found.

Table II. Child-related characteristics and outcome measures in the SPM and SAU groups at T1.

	Total *n (%)*	SPM *n (%)*	SAU *n (%)*	Significant difference SPM vs. SAU *p*
Some or many problems (according to the practitioner) with				
physical health	23 (35.4)	9 (32.1)	14 (37.8)	.634
mental health	55 (84.6)	21 (75.0)	34 (91.9)	.062
cognitive functions	36 (55.4)	13 (46.4)	23 (62.2)	.206
contacting other people	33 (50.8)	13 (46.4)	20 (54.1)	.543
being bullied	34 (52.3)	14 (50.0)	20 (54.1)	.746
controlling emotions	46 (70.8)	17 (60.7)	29 (78.4)	.121
neglecting one's own basic physical needs	38 (58.5)	17 (60.7)	21 (56.8)	.749
	Total *M (SD)*	SPM *M (SD)*	SAU *M (SD)*	*p*
Signs of abuse or neglect according to the practitioner	8.2 (5.9)	8.0 (5.4)	8.4 (6.5)	.752
Need for service intervention according to the practitioner	6.2 (2.4)	6.6 (2.1)	5.9 (2.6)	.122
Subjective well-being according to the child	69.8 (12.9)	72.3 (12.6)	68.2 (13.0)	.200

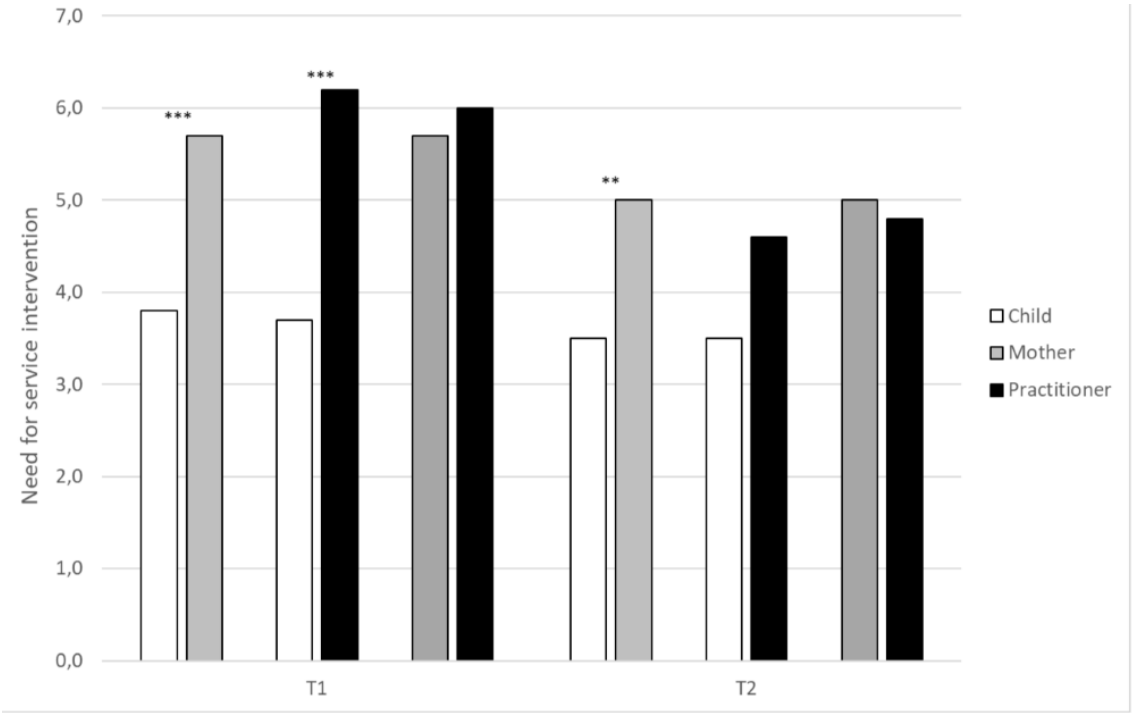
Discrepancies

The need for service intervention according to the child, the child's mother, and the practitioner

Across the sample, the results indicate a significant mother-child discrepancy on the need for service intervention at T1, $t(58) = -4.48$, $p < .001$, and at T2, $t(39) = -2.69$, $p = .01$. Mothers reported significantly higher scores than children at both measurement points. (Figure 1.) A significant child-practitioner discrepancy was found at T1, $t(65) = -5.44$, $p < .001$ but not at T2, $t(42) = -1.72$, $p = .093$. No significant differences between

mothers' and practitioners' assessments were found at T1, $t(80) = -0.70$, $p = .487$, or at T2, $t(41) = 0.34$, $p = .73$).

Figure 1. Need for service intervention according to a child, a mother, and a practitioner at T1 and T2.



In the whole sample, the difference in means between the practitioners' and children's groups decreased from 3.1 to 1.3 over the follow-up period. This change was significant (Wilks' $\Lambda = 0.80$, $F(1, 34)=8.57$, $p = .006$, $\eta_p^2 = .201$. In the SPM group, the mean difference decreased from 4.3 to 1.8, and in the service-as-usual group from 2.3 to 1.0. No statistically significant group-by-time interaction effect was observed (Wilks' $\Lambda = 0.98$, $F(1, 34)=0.74$, $p = .395$, $\eta_p^2 = .021$), indicating that the systemic approach and service-as-usual were equally effective in increasing the level of agreement between a practitioners and a children.

The difference in means between the practitioners' and mothers' groups decreased from 0.8 to -0.1, indicating that at T1, the family's need for child protection was rated higher by practitioners than mothers, whereas at T2, mothers were slightly more concerned about the situation than practitioners. However, no significant effect of time was observed (Wilks' $\Lambda = 0.91$, $F(1,34) = 3.20$, $p = .082$, $\eta_p^2 = .086$). In the SPM group, the mean difference decreased from 1.3 to –0.1. In the service-as-usual control group, the absolute difference remained the same (0.1), although practitioners were more concerned than mothers at T1 and mothers more concerned than practitioners at T2. There was no significant group-by-time interaction (Wilks' $\Lambda = .96$, $F(1,34) = 1.38$, $p = .248$, $\eta_p^2 = .039$).

The difference in means between the mothers and children decreased from 2.2 to 1.6, but no significant effect of time was found (Wilks' $\Lambda = 0.94$, $F(1, 30) = 1.87$, $p = .181$, $\eta_p^2 = .059$). In the SPM group, the difference between the means decreased from

2.5 to 1.2 and in the service-as-usual control group from 2.1 to 1.8. No statistically significant group-by-time interaction effect was observed (Wilks' Λ = .97, $F(1,30)$ = 0.79, p = .382, η_p^2 = .026).

The family dynamics according to children and mothers

Paired-sample t-tests found no significant differences between children's and mothers' assessment of their family's dynamics at either T1 (children: M=2.0, mothers: M=2.1, $t(61)$ = –0.64, p = .524), or T2 (children: M=2.0, mothers: M=2.0, $t(39)$ = 1.47, p = .149). The finding indicates that on average, children and mothers gave similar assessments of the quality of their family's dynamics. The difference in means between children and mothers was 0.0 at T1 and 0.1 at T2 both in the whole sample and in each study group. These differences showed no significant effect of time (Wilks' Λ = .95, $F(1,31)$ = 1.76, p = .195, η_p^2 = .054) or effect of group-by-time interaction (Wilks' Λ = 1.00, $F(1,31)$ = 0.14, p = .712, η_p^2 = .004).

The extent of agreement

The need for service intervention according to children, mothers, and practitioners

In the whole sample, practitioners' assessments of the need for service intervention did not correlate with those of children at either T1 or T2 (see Table III). Similarly, no association was found between practitioners' and mothers' assessments at T1 or T2. Children's scores showed a medium-level (r = .33) correlation (Cohen 1992) with mothers' scores at T1 but not at T2. These findings indicate that the extent of agreement between practitioners and family members was poor. However, child-mother agreement was moderate at baseline but not at follow-up.

To explore the difference in correlations between the SPM and SAU groups at T2, correlations between the same variables were calculated by study group. None of the correlations were statistically significant.

Family dynamics according to children and mothers

In the whole sample, a significant association was found between children's and mothers' assessments of family dynamics at both T1 and T2 (see Table III). The strength of the correlations (r > .50) (Cohen 1992) indicates a high extent of child-mother agreement on their family's dynamics.

At T2, the correlation was $r(21)$ = .58, p = .003 in the SAU group and $r(15)$ = .52, p = .034 in the SPM group. The difference in the correlations was not significant, z = -.28, p = .39.

Other study variables

We also explored the associations between the need for service intervention, family dynamics and two variables used in outcome evaluation, i.e., *the child's subjective wellbeing* and practitioner-assessed *signs of abuse or neglect.* At both measurement points, children's subjective well-being correlated with children's assessment of family dynamics (see Table III). This correlation was negative owing to the scoring protocol. Hence, these findings indicate that poor family dynamics was associated with children's poor subjective well-being. At T1, mothers' assessments of their family's dynamics correlated with their children's subjective well-being, but this association was no longer found at T2. A significant association was found between practitioner-assessed signs of abuse and neglect and the need for service intervention at both T1 and T2. Interestingly, practitioners' assessments of abuse or neglect did not correlate with children's subjective well-being. A medium correlation (r = .33) (Cohen 1992) was found between abuse or neglect and child-assessed family dynamics at T2.

When the study groups were examined separately, significant correlations were found between mothers' assessments of the need for help and their family's dynamics: $r(21) = .61$, $p = .002$ in the SPM group and $r(22) = .51$, $p = .011$ in the SAU group. Significant correlations were also found between mothers' and children's assessments of their family's dynamics: $r(15) = .52$, $p = .034$ in the SPM group and $r(21) = .58$, $p = .003$ in the SAU group. However, these between-group differences in the correlations were not significant.

Table III. Correlations for outcome variables at T1 and at T2.

Variable		1	2	3	4	5	6
1. Need for service intervention, practitioner	T1	–					
	T2	–					
2. Need for service intervention, child	T1	.12	–				
	T2	–.14	–				
3. Need for service intervention, mother	T1	.02	.33*	–			
	T2	.12	.22	–			
4. Family dynamics, child	T1	.10	.30*	.11	–		
	T2	.20	.07	.13	–		
5. Family dynamics, mother	T1	.07	.28*	.50**	.50**	–	
	T2	.20	.12	.54**	.57**	–	
6. Subjective well-being, child	T1	–.07	–.23	–.09	–.58**	–.43**	–
	T2	–.05	–.09	–.27	–.43**	–.27	–
7. Abuse or neglect, practitioner	T1	.53**	.02	–.08	.20	.07	–.08
	T2	.40**	–.14	–.13	.33*	.09	.12

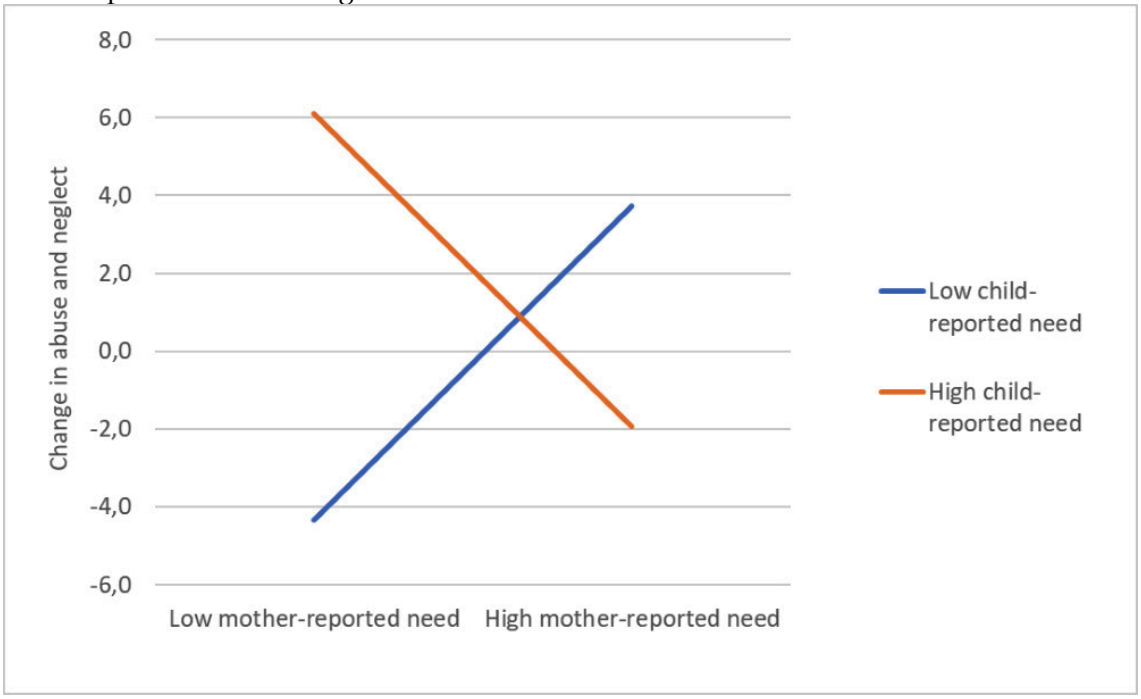
*$p < .05$. **$p < .01$.

Effect of agreement on outcomes

Multiple regression was conducted to explore the relationship between the informants' views on the need for service intervention at T1 and change from T1 to T2 in three outcomes: need for service intervention according to practitioners, signs of abuse or neglect according to practitioners, and children's subjective well-being.

In the whole sample, children's and mothers' assessments of service need did not, separately, predict change in outcomes. However, their interaction did predict change in abuse or neglect. (Table IV.) Based on the analysis, greater agreement between child and mother on the need for help at T1 predicted a decrease in the amount of abuse or neglect. Child-mother agreement on no need for help predicted a greater decrease in abuse and neglect than child-mother agreement on the need for help (Figure 2). Disagreement between children and their mothers predicted an increase in the amount of abuse or neglect reported by the practitioner.

Figure 2. Interaction between child-reported and mother-reported need at T1 predicted change in practitioner-reported abuse or neglect



A similar regression was conducted for each study group. In the SPM group, no significant relationships were found. In the SAU group, a significant relationship between child-mother discrepancy and signs of abuse and neglect was found (B = –2.283, SE B = 0.84, β = –0.52, t = –2.73, p = .013).

In the whole sample, children's and practitioners' assessments on the need for service intervention at T1, both separately and by the interaction, did not predict change in the outcome variables. No significant differences within the study groups were detected. Similarly, mother's and practitioner's views and their interaction did not predict changes in outcome variables in the whole sample or in the study groups.

Table IV. Multiple regression analysis on the relationship between child-mother discrepancy (mean-centred variables) on need for service intervention at T1 and change in scores (T2–T1) of three outcome variables.

	B	*SE B*	β	*p*
Need for service intervention according to practitioner				
Model 1				
child-reported need	0.22	0.57	.07	.706
mother-reported need	–0.10	0.49	–0.04	.846
R^2 = .004, Δ $F(2, 34)$ = .074, p = .928				
Model 2				
child-reported need	0.25	0.61	0.08	.681
mother-reported need	–0.11	0.50	–0.04	.830
Child x mother interaction	–0.11	0.58	–0.03	.852
R^2 = .005, Δ $F(1, 33)$ = .035, p = .852				
Signs of abuse or neglect according to practitioner				
Model 1				
child-reported need	–0.99	1.16	–0.15	.400
mother-reported need	1.59	1.00	0.28	.121
R^2 = .072, Δ $F(2, 34)$ = 1.31, p = .283				
Model 2				
child-reported need	–0.07	1.13	–0.01	.954
mother-reported need	1.26	.94	0.22	.188
Child x mother interaction	–2.77	1.08	–0.41	.015
R^2 = .226, Δ $F(1, 33)$ = 6.57, p = .015				
Child's subjective well-being				
Model 1				
child-reported need	2.94	2.10	0.25	.172
mother-reported need	1.40	1.82	0.14	.45
R^2 = .095, Δ $F(2, 30)$ =1.58, p = .223				
Model 2				
child-reported need	2.62	2.32	0.22	.268
mother-reported need	1.42	1.85	0.14	.448
Child x mother interaction	0.81	2.31	0.07	.727
R^2 = .10, Δ $F(1, 29)$ = 0.12, p = .727				

Discussion

This study explored whether the viewpoints of children, mothers and practitioners converge over time, and, if so, whether this change differs between a group applying the systemic approach to child protection compared to a service-as-usual control group. The effect of agreement on child protection outcomes was also explored.

The first hypothesis was that the agreement would increase across the whole sample over time. This was expected, given that one task of a practitioner is to achieve at least some form of agreement with the service user before proceeding to shared decision-making and planning how to overcome problems (Trevithick, 2012). In this sample, the level of agreement between practitioners and children on the need for service intervention increased over time. At baseline, practitioners gave graver assessments than children of the child's situation. However, six months later the difference between these groups was no longer significant. No differences were found between mothers' and practitioners' assessments.

Interestingly, while children and mothers tended to agree on the quality of their family's dynamics, mothers saw more need for outsider help at both measurement points. This finding can be interpreted in several ways. First, mothers differ from children in their awareness of the service system and support that might be available. Second, as they are responsible for taking care of the child, mothers need to assess the balance between their personal resources and the difficulties within the family that they are trying to tackle. The differences between the role of a child and that of a carer in the family mean that children's and carers' perspectives and experiences will also differ (Miller and Bentovim, 2003).

The correlations found between the study variables revealed that practitioners' analyses of a family's need for service intervention was not associated with either children's or mothers' expressed need for help. At T2, an association was found between children's assessments of their family's dynamics and practitioner-reported abuse or neglect. Otherwise, practitioners' and family members' views did not converge during the process. These findings could indicate less successful progress than expected in this sample. The casework with families had not yet resulted in shared understandings of problems. Several explanations can be offered to interpret the disparity between practitioners and service users' views. First, families might seek to protect their private sphere. Children's loyalty to parents may prevent them from discussing their family life with social workers (Pölkki et al., 2012; Wilson et al., 2020). As Trevithick (2011) points out, "some service users do not want us to know what they know – and this may be a rational or aware stance that they are adopting." Second, children's poor relationship with their social worker, or lack of support or information needed to meaningfully participate in child protection processes may impair their ability or willingness to be involved (Diaz, 2020). Furthermore, the psychological and emotional stress caused by the process of investigating abuse, and mistrust of social workers may hinder children from disclosing abuse and neglect (Wilson et al., 2020). Third, in some cases, violence experienced at home or by peers in the school environment may become 'normalised' (Allnock & Atkinson, 2019; Paterson-Young, 2021). Similarly, some parents who have a history of childhood trauma themselves may have difficulties in distinguishing between acceptable boundary setting and

abusive behavior in parenting (Siverns & Morgan, 2019). Thus, a situation that may seem normal to a child or a parent may rise concern for a social worker.

The second hypothesis was that the agreement would be higher in the SPM group, as the practitioners had received special training and support, including mutual reflection, to increase families' involvement in case work. However, no between-group differences in discrepancies were found. Unfortunately, exploring between-group differences in correlations was not possible. In this sample, the systemic approach did not promote child-mother or family member-practitioner agreement any better than service-as-usual. In their meta-analyses, Bekaert *et al*. (2021) and Falch-Eriksen *et al*. (2021) conclude that to engage and support families, practitioners need more time, reasonable caseloads, and adequate budgets. Factors facilitating participation were active listening, access to the practitioner, private discussions with the practitioner, and practitioner effort to build a trusting relationship with the child (Falch-Eriksen *et al*., 2021). In both groups, high caseloads may have affected practitioners' possibilities to facilitate families' engagement.

Finally, we explored the effect of agreement on outcomes. The results showed that child-mother disagreement on the need for service intervention predicted negative and child-mother agreement positive outcomes in terms of abuse or neglect. Interestingly, this decrease was larger when both child and mother reported that the family was not in need of outsider help. Despite this, the child protection practitioners seem to have carried out a successful intervention. The present findings suggest that child protection practice is less successful when one family member sees the need for service intervention and the other one does not.

Strengths and limitations

This study is among the few in the context of children's social services that has collected quantitative data from three groups of informants and compared their viewpoints. Outcomes were assessed by two informants, the child and the practitioner. This decision was based on the legal responsibility of social workers to assess children's need for protection, and on children's legal right to express their views and participate in the decision-making process.

The study has its limitations. First, the small sample size limits the representativeness of the study, and the attrition between the two measurement points made between-group comparisons difficult. However, at T1, apart from the mean higher age of the SAU group children, no between-group differences in background variables were found. Attrition was not associated with study group. Thus, the study groups were similar enough to allow reasonable comparisons between them. Second, the low participation of fathers resulted in their exclusion. In future research, more attention should be paid to recruiting fathers.

Third, in the absence of any single indicator of what constitutes success in a child protection intervention (Forrester, 2017), we used a combination of measures designed and validated in other settings (i.e., SCORE-15, KINDL) and non-validated measures designed for this study (need for service intervention, abuse and neglect). Validating these measures for this population was beyond the scope of this study. Thus, to

develop and validate suitable outcome measures for use in the child protection context requires more research.

Finally, according to a previous report (Isokuortti & Aaltio, 2020), the systemic approach had not yet been fully implemented during the study period. This means that had differences between the study groups been found, this would have limited the conclusions. More collaboration between implementers and researchers is clearly needed to improve implementation quality and success in data collection.

Conclusions and implications

The findings from this sample do not suggest that the SPM, as it was implemented during the first stage of national dissemination, is likely to outperform service as usual in improving agreement between service users and practitioners. Further research is needed to explore the effectiveness of systemic practice on agreement.

More generally, the present findings regarding the whole sample have implications for future practice. First, practitioner assessments of the family's situation might not reflect child or parent assessments. Hence, the situation should be explored from multiple perspectives using multiple sources of information to identify alternative hypotheses and strategies for the next steps.

Second, possible disagreements between family members should be addressed in practice since child-parent disagreement might have a negative effect on child protection outcomes. To do this, practitioners should first form a good relationship both with a child and parents. Based on previous research on shared decision-making, practitioners need better understanding, additional training, and more time to engage service users (Diaz, 2020). As further described by Stabler et al. (2020), practitioners need to work collaboratively with the family to create open dialogue and motivate family to engage with the process, and build trust and reduce shame to discuss difficult issues. This has implications for organizations since they need to provide supervision and organization culture to support good practice. Moreover, service users suffer from continuous changes of their social workers and their lack of time. Hence, organizations need to find strategies to decrease practitioners' high caseloads and reduce rapid turnover of the staff (Diaz, 2020).

Third, findings indicate that child-parent agreement predicts a successful intervention both when they agree or disagree on for need professional help. One hypothesis could be that child-parent agreement in general provides a basis for collaboration between practitioner and family. On the other hand, agreement may also hide possible risk factors within the family if it relates to child's loyalty towards the parent or 'normalisation' of problems. Thus, both in research and in practice, the mechanisms explaining the agreement should be further explored and critically appraised.

To conclude, information should be more often collected from multiple sources allowing comparison between practitioners' and family members' viewpoints. In terms of future research, along with quantitative measures enabling statistical analysis, qualitative research, such as individual interviews, is needed to understand the underlying mechanisms that explain identified agreements and disagreements at

a given timepoint and changes in practitioner's, child's and parent's viewpoints over time.

References

Allnock, D., and Atkinson, R. (2019), "'Snitches get stitches': School-specific barriers to victim disclosure and peer reporting of sexual harm committed by young people in school contexts", *Child Abuse & Neglect*, Vol. 89, pp.7–17. https://doi.org/10.1016/j.chiabu.2018.12.025

Baginsky, M., Ixer, G. and Manthorpe, J. (2020), "Practice Frameworks in Children's Services in England: An Attempt to Steer Social Work Back on Course?", *Practice*, Vol.33 No.1, pp.1–17. https://doi:10.1080/09503153.2019.1709634.

Bekaert, S., Paavilainen, E., Scheke, H., Baldacchino, A., Jouet, E., Zablocka-Zytka, L., Bachi, B., Bartoli, F., Carra, G., Cioni, RM., Crocamo, C. and Appleton, JV. (2021) "Family members' perspectives of child protection services, a metasynthesis of the literature", *Children and Youth Services Review*, Vol.128 Article 106094. doi:10.1016/j.childyouth.2021.106094

Berrick, J., Peckover, S. Pösö, T., and Skivenes, M. (2015), "The formalized framework for decision-making in child protection: A cross-country comparison", *Journal of European Social Policy*, Vol. 25 No.4, pp.366–378. doi:10.1177/0958928715594540

Berrick, J., Dickens, J., Pösö, T. and Skivenes, M. (2016), "Time, Institutional Support, and Quality of Decision Making in Child Protection: A Cross-Country Analysis", *Human Service Organizations: Management, Leadership & Governance*, Vol. 40 No.5, pp.451–468. doi:10.1080/23303131.2016.1159637.

Bostock, L., Forrester, D., Patrizo, L., Godfrey, T., Zonouzi, M., Antonopoulou, V., Bird, H. and Tinarwo, M. (2017), "Scaling and deepening the Reclaiming Social Work model: Evaluation report", Department for Education, London. Available at: https://www.gov.uk/government/publications/scaling-and-deepening-the-reclaiming-social-work-model (accessed 14 October 2019)

Bostock, L., Patrizo, L., Godfrey, T. and Forrester, D. (2019), "What is the impact of supervision on direct practice with families?", *Children and Youth Services Review*, Vol. 105 Article 104428. https://doi:10.1016/j.childyouth.2019.104428

Carr, S. (2004) "Has service user participation made a difference to social care services?", position paper 3, The Social Care Institute for Excellence, London, March 2004.

Carr, A. and Stratton, P. (2017), "The Score Family Assessment Questionnaire: A Decade of Progress", *Family Process*, Vol. 56 No. 2, pp.285–301. https://doi: 10.1111/famp.12280Cohen, J. (1992), "A power primer", Psychological Bulletin, Vol. 112 No.1, pp.155–159.

Diaz, C. (2020), Decision making in child and family social work: perspectives on participation. Policy Press, Bristol.

Falch-Eriksen, A., Toros, K., Sindi, I. and Lehtme, R. (2021), "Children expressing their views in child protection casework: Current research and their rights going forward", *Child & Family Social Work*, Vol. 26 No. 3, pp.485–497. https://doi.org/10.1111/cfs.12831

Forrester, D., Westlake, D., McCann, M., Thurnham, A., Shefer, G., Glynn, G. and Kilian, M. (2013), "Reclaiming social work? An Evaluation of Systemic Units as an Approach to Delivering Children's Services", University of Bedfordshire,

Tilda Goldberg Centre for Social Work and Social Care. Luton. Available at: http://uobrep.openrepository.com/uobrep/handle/10547/594517 (accessed 18 June 2017)

Forrester, D. (2017), "Outcomes in children's social care", *Journal of Children's Services*, Vol. 12 No. 2-3, pp.144-157. https://doi.org/10.1108/JCS-08-2017-0036

Forrester, D., Westlake, D., Killian M., Antopolou, V., McCann, M., Thurnham, A., Thomas, R., Waits, C., Whittaker., C and Hutchison, D. (2019), "What is the relationship between Worker Skills and Outcomes for Families in Child and Family Social Work", *The British Journal of Social Work*, Vol. 49 No. 8, pp. 2148-2167. https://doi.org/10.1093/bjsw/bcy126

Gilbert, N., Parton, N., and Skivenes, M. (2011) *Child Protection Systems: International Trends and Orientation*, Oxford University Press, New York. DOI: 10.1093/acprof:oso/9780199793358.001.0001

Gladstone, J., Dumbrill, G., Leslie, B., Koster, A., Young, M. and Ismaila, A. (2012), "Looking at engagement and outcome from the perspectives of child protection workers and parents", *Children and Youth Services Review*, Vol. 34 No. 1, pp. 112-118. https://doi.org/10.1016/j.childyouth.2011.09.003

Goodman, S. and Trowler, I. (Eds.) (2012), *Social Work Reclaimed. Innovative Framework for Child and Family Social Work Practice*, Jessica Kingsley Publishers, London and Philadelphia.

Goolsby, J., Rich, B., Hinnant, B., Habayeb, S., Berghorst, L., De Los Reyes, A., and Alvord, M. (2018), "Parent-Child Informant Discrepancy is Associated with Poorer Treatment Outcome", *Journal of Child and Family Studies*, Vol. 27 No. 4, pp. 1228-1241. https://doi.org/10.1007/s10826-017-0946-7

Heimer, M., Näsman, E. and Palme, J. (2018), "Vulnerable children's rights to participation, protection, and provision: The process of defining the problem in Swedish child and family welfare", *Child & Family Social Work,* Vol. 23 No. 2, pp. 316-323. https://doi.org/10.1111/cfs.12424

Horvath, A. O., and Symonds, B. D. (1991), "Relation between working alliance and outcome in psychotherapy: A meta-analysis", *Journal of Counseling Psychology*, Vol. 38 No. 2, pp. 139–149. https://doi.org/10.1037/0022-0167.38.2.139

Isokuortti, N. and Aaltio, E. (2020) Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland", *Children and Youth Services Review*, Vol. 119, 105647.

Kiili, J., Itäpuisto, M., Moilanen, J., Svenlin, A.-R. and Malinen, K.E. (2021), "Professionals' views on children's service user involvement", *Journal of Children's Services*, Vol. 16 No. 2, pp. 145–158. https://doi.org/10.1108/JCS-10-2020-0069

Lahtinen, P., Männistö, L. and Raivio, M. (2017), "Kohti suomalaista systeemistä lastensuojelun toimintamallia: Keskeisiä periaatteita ja reunaehtoja" ["Towards the Finnish Systemic Practice Model for child protection. Core principles and preconditions"], Discussion paper 7, Finnish Institute for Health and Welfare, Helsinki.

Miller, L.B. and Bentovim, A. (2003), "Assessing Families: The Family Assessment of Family Competence, Strengths and Difficulties", Bell M. and Wilson K. (Eds), *The Practitioner's Guide to Working with Families,* Palgrave, London, pp.57-84.

Paterson-Young, C. (2021), "Exploring how children subjected to violence in the home cope with experiences in Secure Training Centres", Child Abuse & Neglect, Vol. 117, Article 105076. https://doi.org/10.1016/j.chiabu.2021.105076

Pölkki, P., Vornanen, R., Pursiainen, M. and Riikonen, M. (2012), "Children's Participation in Child-protection Processes as Experienced by Foster Children and Social Workers", Child Care in Practice, Vol. 18, No.2, 107–125. https://doi.org/10.1080/13575279.2011.646954

Ravens-Sieberer, U. and Bullinger, M. (1998), "Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results", *Quality of Life Research*, Vol. 7 No.5, pp. 399–407. doi: 10.1023/a:1008853819715.

Shirk, S. R., and Karver, M. (2003), "Prediction of treatment outcome from relationship variables in child and adolescent therapy: A meta-analytic review", *Journal of Consulting and Clinical Psychology*, Vol. 71 No. 3, pp. 452–464. https://doi.org/10.1037/0022-006X.71.3.452

Siverns, K., & Morgan, G. (2019). Parenting in the context of historical childhood trauma: An interpretive meta-synthesis. Child Abuse & Neglect, 98, Article 104186. https://doi.org/10.1016/j.chiabu.2019.104186

Stabler, L., O'Donnell, C., Forrester, D., Diaz, C., Willis, S. and Brand, S. 2019. Shared decision-making: What is good practice in delivering meetings? Involving families meaningfully in decision-making to keep children safely at home: A rapid realist review. Technical Report.

Stratton, P., Bland, J., Janes, E., and Lask, J. (2010), "Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: The SCORE", *Journal of Family Therapy,* Vol. 32 No. 3, pp. 232-258. https://doi.org/10.1111/j.1467-6427.2010.00507.x

Trevithick, P. (2012), *Social work skills and knowledge. A practice handbook*, Third Edition, McGraw-Hill Education, Open University Press, Maidenhead

Trevithick, P. (2011), "Understanding defences and defensiveness in social work", *Journal of Social Work Practice*, Vol. 25 No. 4, pp. 389-412, DOI: 10.1080/02650533.2011.626642

Turnell, A. and Murphy, T. (2017), *Signs of Safety. Comprehensive Briefing Paper*, 4th edition, East Perth, Elia International Ltd. https://knowledgebank.signsofsafety.net/resources/introduction-to-signs-of-safety/signs-of-safety-comprehensive-briefing-paper/signs-of-safety-comprehensive-briefing-paper-en/signs-of-safety-comprehensive-briefing-paper?

III

FIDELITY AND INFLUENCING FACTORS IN THE SYSTEMIC PRACTICE MODEL OF CHILDREN'S SOCIAL CARE IN FINLAND

by

Isokuortti, N. & Aaltio, E. 2020

Children and Youth Services Review, *119*, Article 105647.

https://doi.org/10.1016/j.childyouth.2020.105647

Reproduced with kind permission by Elsevier.

Children and Youth Services Review 119 (2020) 105647

Contents lists available at ScienceDirect

Children and Youth Services Review

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

Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland

Nanne Isokuortti [a,*], Elina Aaltio [b,c]

[a] Faculty of Social Sciences (Social Work), University of Helsinki, Finland
[b] Department of Social Sciences and Philosophy (Social Work), University of Jyväskylä, Finland
[c] Department of Children, Young People and Families, Finnish Institute for Health and Welfare, Finland

ARTICLE INFO

Keywords:
Systemic Practice Model
Implementation
Fidelity
Influencing factors
Children's social care
Systemic family therapy

ABSTRACT

Given that multiple countries have recently adopted social work practice models in children's services, it is striking that only a few studies have systematically analysed both the level of fidelity and potential implementation barriers and facilitators. The aim of this study is to provide an in-depth analysis of how and why the Reclaiming Social Work (RSW) model works in different settings. The study context was the implementation in Finland of an adaptation of the model, the Systemic Practice Model (SPM). This mixed-methods study evaluates 1) fidelity to the SPM and 2) the possible influencing factors. The results reveal high variability in the extent of fidelity in 23 implementation sites, and even among individual teams within the same site. A lack of clarity concerning systemic social work practice, insufficient training, and inadequate resources and leadership hindered the implementation, whereas coaching and positive experiences of the SPM were facilitating factors. In particular, the involvement of a clinician qualified in systemic family therapy was crucial in embedding the new approach. The relationship between the level of fidelity and the influencing factors worked both ways (e.g., low coverage was associated with a decrease in participant responsiveness, and vice versa). Given the complexity of children's social care as an implementation environment, careful preparation and ongoing support are crucial in the implementation of practice models.

1. Introduction

In recent years, multiple countries have adopted social work practice models (or practice frameworks) to improve outcomes for children and families (Gillingham, 2018). Practice models are embedded in a particular theory and practice approach, which guide all stages and aspects of social work (Baginsky et al., 2020; Barbee et al., 2011). Despite their popularity, a number of evaluations related to such models describe challenges in their implementation (e.g., Antle et al., 2012; Laird et al., 2018). Ultimately, failure in this regard compromises the model's effectiveness (Durlak & DuPre, 2008). The present study analyses the initial nation-wide implementation of the Systemic Practice Model (SPM) in Finland. The SPM is an adaptation of the Reclaiming Social Work (RSW) model (Goodman & Trowler, 2012), developed in an English child and family social work agency. The aim of the model is to deliver systemic social work practice in children's services.

The purpose of this study is to assess implementation fidelity (i.e., the extent to which the intervention is delivered as intended) and to formulate hypotheses concerning the possible influencing factors, which could be tested and refined in future evaluations. The undertaking of a comprehensive fidelity assessment gives researchers and practitioners a better understanding of how and why the model works in different contexts and the extent to which its outcomes can be improved (Carroll et al., 2007). A detailed implementation analysis is essential particularly when transporting interventions from one cultural context to another (Sundell et al., 2014).

The present study addresses several gaps in the existing research. First, it represents the first attempt to evaluate the fidelity of the RSW model, including its adaptations outside England. Second, it complements previous implementation studies on practice models in combining fidelity assessment and the analysis of influencing factors (Carroll et al., 2007). The objectives are:

1. to describe the level of fidelity of the SPM by measuring the details of its content, dose (in other words frequency and duration) and coverage;

* Corresponding author at: University of Helsinki, Faculty of Social Sciences (Social Work), Unioninkatu 37, P.O. Box 54, 00014 University of Helsinki, Finland.
E-mail addresses: nanne.isokuortti@helsinki.fi (N. Isokuortti), elina.m.aaltio@jyu.fi (E. Aaltio).

https://doi.org/10.1016/j.childyouth.2020.105647
Received 4 June 2020; Received in revised form 22 October 2020; Accepted 22 October 2020
Available online 28 October 2020
0190-7409/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

2. to identify possible implementation barriers and facilitators based on participant experiences.

1.1. Fidelity and influencing factors

Although multiple terms are used in the literature, in the context of evaluation research fidelity usually refers to the extent to which implementers adhere to the programme as it was designed by the developers (Carroll et al., 2007). It is necessary to measure fidelity because failure to deliver the core components as intended ultimately influences the intervention outcomes (Durlak & DuPre, 2008). Consequently, fidelity is considered one of the key outcomes of implementation (Proctor et al., 2010). However, Toomey et al. (2020) point out that fidelity is increasingly viewed as a more multi-faceted concept, which focuses not only on the delivery of a programme but also on interrelationships between domains such as delivery and receipt (see e.g., Bellg et al., 2004). In essence, merely assessing fidelity does not provide information on the factors that influenced the implementation outcomes. Mindful of this, we chose the Conceptual Framework for Implementation Fidelity (CFIF) (Carroll et al., 2007) as an evaluation framework because it facilitates the joint analysis of fidelity and influencing factors. It also incorporates the different measurement areas and acknowledges their relationships. In general, frameworks broaden current understanding of fidelity, guide its assessment and enhancement, and support the structuring and standardising of research, all of which enhance comparability and the synthesising of findings (Toomey et al., 2020).

According to Carroll et al., (2007), fidelity (or adherence) may be measured in terms of the content, coverage (or reach), frequency and duration of the intervention: the last two measures could be included in the concept of dose. Evaluators should assess the extent to which the content of the intervention, in other words the 'active ingredients', have been administered to the participants as often and for as long as intended. The degree to which the intended content, dose and coverage have been delivered is the degree of implementation fidelity achieved for that model.

Various influencing factors may facilitate or impede the level of fidelity, and they may influence each other (facilitation strategies could enhance the quality of the delivery, for example). The factors (or moderators) included in the CFIF include intervention complexity (a description of the intervention and its real nature), facilitation strategies (e.g., training, manuals), the quality of delivery (the appropriateness of the process) and participant responsiveness (engagement with the model). Responsiveness refers both to the service users and to those responsible for delivering the model (Carroll et al., 2007). The modified version used in this study (Hasson, 2010) has two additional factors: context (e.g., the surrounding social systems as well as historical and concurrent events) and recruitment (procedures used to attract potential intervention participants). Our assessment of the influencing factors focuses on intervention complexity, facilitation strategies, participant responsiveness and organisational factors. The focus is on a provider context (von Thiele Schwarz et al., 2019), in this case social work teams. Because this evaluation was of the initial implementation of the SPM, and it relied on survey and interview data, we excluded the assessment of the quality of delivery. We also excluded the assessment of client recruitment because that factor is not applicable in the context of a statutory child protection service, given its involuntary characteristics. To avoid conceptual confusion, we use the term 'influencing factor' as well as 'barriers and facilitators' instead of 'moderators' in this article. Fig. 1 below illustrates the application of the CFIF in this study.

Although the balance between fidelity and adaptation is continuously addressed in the literature, many studies take 'one or the other position' without explicitly considering the level of 'appropriate' adaptation (von Thiele Schwarz et al., 2019; Miller et al., 2020; Toomey et al., 2020). The authors of these studies therefore recommend addressing the issue of fidelity and adaptation throughout the process, as well as exploring coexistence potential by explicating the core components of the intervention. Even the core components may be changed or removed in some instances, but such actions should be carefully planned and measured so as to avoid unintentional deviation (Stirman et al., 2013; Miller et al., 2020). Accordingly, we define the core components of interest in Section 2.2, and in terms of measures in Section 2.5.

In sum, evaluation of an implementation may prevent the drawing of potentially false conclusions about its effectiveness in that it provides insights into the process and the factors that influence the outcomes (Carroll et al., 2007). Moreover, careful documentation of the process enhances the generalisability of the findings to other sites (Crea et al., 2009).

1.2. The Systemic Practice Model (SPM) and its implementation

The SPM is a Finnish adaptation of the RSW model, which incorporates systemic family therapy into child and family social work. The RSW has not been manualised, but the general ideas, values and theories underpinning it have been published by its co-founders (Goodman & Trowler, 2012). The overarching idea is to form small, multi-disciplinary units including a social worker, a systemically trained family therapist, a consultant social worker leading the team and a unit coordinator. Team members receive training in systemic thinking and methodology, and in the use of tools with families as well as within the weekly team meetings that serve as the main mechanism for case supervision. To our knowledge, thus far the RSW model has been implemented only in its country of origin.

The SPM was disseminated and implemented by the Finnish Institute of Health and Welfare (THL) in 2017–2018, funded by the Ministry of Social Affairs and Health. To support agencies in its implementation, THL published a paper describing the general idea of the model (Lahtinen et al., 2017), and organised national training of trainers (ToT). It also organised three national workshops for child protection managers, the aim being to inform them about systemic practice and its

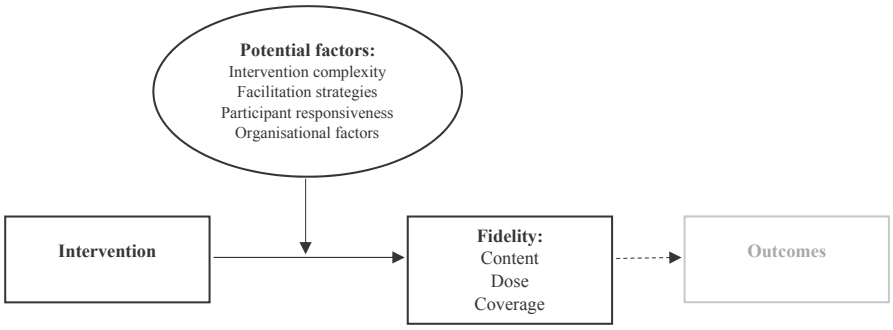


Fig. 1. The assessment of fidelity and the influencing factors in the present study in accordance with the modified CFIF (Hasson, 2010, originally from Carroll et al., 2007).

implementation.

1.3. Implementation evaluations of practice models for child and family social work

To date there have been few studies assessing the fidelity of practice models such as Signs of Safety (SoS) from Australia and Solution-Based Casework (SBC) from the United States. Researchers have identified high variability in the implementation of these models (Antle et al., 2012; Sheehan et al., 2018). Antle et al. (2012), for example, demonstrated with regard to SBC that a higher level of fidelity was associated with better case outcomes, whereas Sheehan et al. (2018) concluded in their review that there was limited evidence of whether SoS has been implemented well. Roberts et al. (2019) recently developed a fidelity-measurement tool for SoS, but to date there are no published systematic fidelity assessments.

A number of evaluations have identified barriers and facilitators in the implementation of practice models (Lambert, Richards, & Merrill, 2016; Pipkin, Sterrett, Antle, & Christensen, 2013; Sanclimenti, Caceda-Castro, & Desantis, 2017; Sheehan et al., 2018). The most substantial influencing factors include a supportive leadership and organisational climate, high-quality training and coaching, as well as alignment with other organisational systems and initiatives. Some studies also list high staff turnover as a significant implementation barrier (Roberts et al., 2019; Sheehan et al., 2018). Despite the positive experiences of professionals and service users with the models (Antle et al., 2012; Sheehan et al., 2018), the evaluations highlight the complexities involved in their implementation, particularly in large public systems. Above all, previous research emphasises the need to engage the whole organisation, including its senior leaders, in the change effort to support implementation (Lambert et al., 2016; Pipkin et al., 2013; Sanclimenti et al., 2017; Sheehan et al., 2018).

1.4. Evaluations of the Reclaiming Social Work (RSW) model

It has been concluded from previous evaluations of the RSW model that its implementation increases the quality of children's services compared to service as usual. According to a mixed-method evaluation carried out in the original children's services site in Hackney (Cross et al., 2010), social workers in new RSW units were more satisfied with the work environment, the social work processes and work-related wellbeing than practitioners who were not using the model. In addition, the number of looked-after children had decreased by 30 per cent during the period 2005/2006–2008/2009. Forrester et al. (2013) compared the RSW units in Hackney with two other sites in their realist-informed mixed-method evaluation. They concluded that practitioners in the RSW units spent more time with families, made high-quality assessments, demonstrated high levels of direct practice skills, and provided more intensive help for families. On the basis of their evaluation in other sites implementing the RSW, in turn, Bostock et al. (2019) quantified and paired observations of supervision ($n = 14$) with observations of direct practice ($n = 18$) and found a statistically significant association between the quality of case supervision and the quality of social work practice with families. However, it should be noted that the overall evidence base concerning the effectiveness of the RSW and other practice models is still limited (Isokuortti et al., 2020).

According to both mixed-methods evaluation focusing on five sites (Bostock et al., 2017) and qualitative evaluations focusing on one site (Laird et al., 2017, 2018), there is extensive variability in how a team structure is adopted, team meetings are run, and systemic practice is implemented. Bostock et al. (2019) report a variation in the quality of systemic case supervision in observed team meetings ($n = 14$) from non-systemic to 'encouraging' and 'fully systemic'. Laird et al. (2017, 2018) and Morris et al. (2018), in turn, found limited changes in social work practice in one agency that did not implement any structural changes.

Bostock et al. (2017) concluded that a systemically trained consultant social worker as a leader, systemic case discussion, clinician input and dedicated administrative support are vital in ensuring good systemic practice. Bostock et al. (2019) further stress the importance of having a clinician present to ensure the full incorporation of systemic concepts and practice in the supervision. Laird et al. (2018) emphasise the need to implement the whole model as intended given that a reduction in the amount of training for managers prevented the establishment of shared values and the changing of procedures and practice. Laird et al. (2017, 2018) further note that the caseloads of social workers should to be aligned with systemic practice to allow enough time for them to learn the new approach. Both Bostock et al. (2017) and Laird et al. (2018) found that recruitment challenges and staff turnover impeded implementation, and they stress the importance of supportive leadership and the engagement of the whole system in the change. Notwithstanding the challenges, the practitioners were satisfied with the RSW model, which they believed had improved their practice through its collaborative, reflexive and purposeful approach (Bostock et al., 2017). In addition, most families had positive perceptions of children's social care based on systemic practice (Bostock et al., 2017; Morris et al., 2018).

The initial evaluation of the model in Finland reflects the English implementation experiences. The Finnish case shows that high caseloads in particular impeded the uptake of systemic practice, although the social workers had generally positive perceptions of the model itself (Aaltio & Isokuortti, 2019). Nevertheless, there is still a need for a joint analysis of the interrelationship between the fidelity of the model and the possible factors that influence it, which is the focus of this study.

2. Materials and methods

2.1. Study design

A mixed-methods approach was used in the present study to allow an in-depth analysis of implementation fidelity and the participants' perceptions of the potential factors that influence it. We adopted a concurrent transformative design, which incorporates simultaneous qualitative and quantitative data collection and analysis followed by mutual interpretation of the findings based on a specific theoretical perspective (Creswell & Plano Clark, 2008). The primary data we used in our investigation of fidelity were quantitative, which we complemented with qualitative data. Our analysis of the influencing factors, in turn, was based on qualitative data complemented with quantitative data. Complementing the quantitative fidelity assessment with the qualitative analysis gave us a more detailed view on the use of systemic social work practice. Furthermore, comparing the qualitative interview findings with the survey results allowed us to compare the participants' perceptions on influencing factors in the three sites with data from a larger sample, as well as to preliminarily test some of these assumptions with the quantitative data.

Our study was further influenced by realist evaluation (Pawson & Tilley, 1997), which is a theory-driven form of evaluation that purports to formulate and refine theories explaining why a programme does or does not work. Realist evaluation is rooted in the realist philosophy of science. The point is that change is seen as a result of underlying mechanisms that are not always observable, thus the purpose is to identify these mechanisms by theorising and testing hypotheses based on these theories with empirical data. It is also understood that the functioning of mechanisms depends on the context. Hence, the aim is to understand what works, for whom, and under which circumstances.

Our aim is to formulate hypotheses about possible relationships between context and fidelity based on our mixed-methods data. We acknowledge that the model may work in some settings but not in others, hence our intention is to analyse how different factors, especially the context, may influence implementation (see also von Thiele Schwarz et al., 2019 on fidelity and context). The context referred to in realist evaluation includes the material, the social, the psychological, the

organisational, the economic and the technical (Greenhalgh et al., 2017). Consequently, team-level facilitation strategies and participant responsiveness as listed in the CFIF are different 'types of context', whereas intervention complexity refers to the characteristics of the model. To avoid conceptual confusion, we use the term "organisational factors" for what Hasson (2010) refers to as "context". As a result of our analysis, we have formulated hypotheses of how different factors, especially the context, influence implementation. The study was conducted in parallel with an outcome evaluation.

2.2. Operationalising the core components of the SPM

The original RSW model comprises a whole system reform, which involves structural changes (e.g., forming systemic units) and adapting the 7S framework to achieve effective practice. The approach is based on shared values, such as collaboration and respectful work, and skills derived from systemic family therapy and social learning theory (Goodman & Trowler, 2012). Forrester et al. (2013) outlined six features that distinguished RSW units in Hackney from conventional social work teams: (1) shared work, (2) in-depth case discussion, (3) a shared systemic approach, (4) skills development, (5) special roles and (6) low caseloads (pp. 88–102).

The Finnish adaptation was inspired by all the above-mentioned features, but the set of skills and theories differed. The most significant differences from the original model are the following: a larger team size, a focus on team-level change instead of training and coaching both practitioners and managers, as well as introducing new methods (such as inviting families to team meetings) and tools (such as the 'collaborative helping map' or 'three houses') from outside the RSW curriculum. In addition, ideas and their operationalisation evolved during the initial implementation. As an illustration, some implementers began to invite families to the team meetings during the implementation period, whereas others wished to restrict the meetings to professional groups to support learning and reflection. Consequently, the implementers did not share a mutual understanding about which of the ideas and SPM components were more important than others, or how to put them into practice (Aaltio & Isokuortti, 2019). Following the initial evaluation the researchers suggested to the national stakeholders that a series of workshops be held aimed at formulating a first SPM programme theory: this is currently under review.

On the basis of the findings from the initial evaluation and the stakeholder discussions, we identified the following three core components of SPM: (1) a team structure comprising a consultant social worker, between one and three social workers, a clinician (i.e., a qualified family therapist) and a coordinator (in total a maximum of eight members); (2) the holding of weekly reflective team meetings; and (3) systemic social work practice. First, the clinician helps the social workers with their systemic thinking and the use of family-therapy techniques, and the coordinator assists them with administrative tasks thereby giving them time for intensive casework. The consultant social worker provides practice leadership, whereas social workers take responsibility for the cases. Second, the purpose of the weekly team meetings is to reflect and find multiple perspectives on family cases by applying systemic thinking and similar techniques. The intention is that these systemic supervision sessions will help social workers to reflect on their cases and to plan interventions to support families. Third, to enable them to work with families in a systemic way, social workers are trained to follow a systemic approach and to apply the relevant techniques in their practice. The identified key techniques were genograms, formulating systemic hypotheses and circular questions, of which only first two were part of the initial training. Adopting a systemic approach entails frequent face-to-face work with families to exploit the full potential of these techniques. In conclusion, our fidelity assessment is based on these components (see Section 2.5).

2.3. Procedure

Ethical approval was granted by the National Institute of Health and Welfare Research Ethics Committee (2017–09). The research data was collected between five and twelve months after the commencement of the implementation at the site. The lead author designed the interview protocols, on which the second author gave comments, and conducted interviews with all social workers in June-September 2018 as well as with sixteen service users in July-November 2018. A research assistant conducted four interviews with service users at site three (two children and two parents), following the lead author's guidance and a shared interview protocol (see Section 2.5.2.1). All the participants were given information sheets, and they signed a consent form concerning their participation and the audio recording. The service users were offered a cinema voucher as an incentive, and the social workers participated in the interviews during their work hours. The interviews with the social workers were conducted in team-based focus groups, whereas the service users were interviewed individually. All the interviews were conducted face-to-face apart from one: a parent at site two was interviewed by phone. The social worker focus-group meetings lasted between 87 and 130 min, and individual interviews with service users between 26 and 63 min. After the interviews, the lead author discussed the initial findings with the second author based on the notes. The lead author was responsible for the qualitative analyses, including assessing the adoption of systemic practice based on the interviews (see Section 2.5.1.6). The interviews were recorded and transcribed verbatim.

The second author collected administrative data in February 2018 as part of the national evaluation conducted by the Finnish Institute for Health and Welfare (THL) in 2017–2019. This data was gathered from managers via administrative forms in a national workshop supporting the implementation. The forms included questions regarding the team structure and the resources of SPM teams, as well as the implementation schedule in each site. If information was missing, the supervisors were asked to complete the forms via email. In addition, the second author designed a survey (see Section 2.5), which was refined based on reflective discussions with the steering group of the national evaluation and the lead author. The survey was conducted in September-October 2018. The invitation and the link to the survey, and two follow-up reminders, were sent via email. The email addresses of the social workers were collected from managers after permission to conduct the research given at each site. All the participants gave their informed consent prior to their inclusion in the study.

2.4. Sampling and participants

The model was implemented between the autumn of 2017 and the summer of 2018 among 52 teams in 31 municipal children's service sites located in 14 counties around Finland. All these sites were asked to participate in the quantitative research and to give permission for survey data to be collected from practitioners. Permission was received from 27 implementation sites, of which one decided to withdraw due to implementation difficulties and three failed to deliver contact details for sending the survey. The four sites that did not grant permission to conduct the research informed the researcher that they had postponed the implementation, or could not name a responsible person to be contacted regarding the research permission and the participants. The 23 sites that eventually participated in the research covered 74 per cent of all known sites. Within these sites, the SPM was implemented among 39 teams focusing primarily on child protection and covering 75 per cent of all SPM teams. The survey was sent to all social workers in these teams (response rate 44%, $n = 56$), whereas administrative data was collected at a national workshop and through emails from managers.

Qualitative interviews were conducted with social workers and service users in three purposefully selected sites, in which the outcome data (see Section 2.1) was also gathered. All the sites were large (<100 000 habitants) enough to include multiple child protection teams to allow

comparison. However, they varied geographically: site one is situated in Southern, site two in Eastern and site three in Central Finland. These sites were also convenient choices given their willingness to allow more extensive data collection. Nine child protection teams in these sites implemented the SPM. All the social workers in these teams ($N = 44$) were invited to be interviewed, of whom 32 agreed (participation rate 73%). Table 1 presents the characteristics of the social-worker participants.

Service users ($n = 20$) were selected with the help of the social workers, because not all families served by the team had experience of systemic practice (see Section 3.1). The participants included six 12–17-year-old children (five girls and one boy) and 14 parents (12 mothers and two fathers). All of them had been involved in child protection prior to the implementation. Given that we did not aim to collect data from all the families involved in the teams we did not calculate the participation rate.

2.5. Measures

The fidelity measures were based on the authors' operationalisation of the core components of SPM (see Section 2.2). Table 2 gives an overview of the measures used to indicate the level of fidelity (see also Section 2.5.1) and questions concerning influencing factors, which were analysed from the interview (social workers and service users) and survey (social workers) data (see Section 2.5.2).

2.5.1. Fidelity

The fidelity thresholds were constructed as follows. First, we defined the content of the high-fidelity category for each measure. Given that the evaluation focused on the initial stage, our threshold for high fidelity was set below a perfect performance but on a level that clearly indicates the delivery of the model (see also Section 1.1). Next, we defined the category of medium fidelity such that it included signs of promising efforts to deliver the model but excluded cases that indicated only slight changes in teams or practices. The low-fidelity category included cases that indicated minor or no changes. Given the subjective nature of this process, the researchers formulated several versions of the fidelity thresholds, which were jointly discussed and refined based on the preliminary analyses.

Table 1
The characteristics of the social workers participating in the interviews and the survey.

Characteristics	Social worker interviews ($N = 32$)	Social worker survey ($N = 56$)
	M (SD)	*M (SD)*
Team size (i.e., a number of team members)	7.9 (1.1)	8.0 (1.5)
Caseload	Range 22–53	40.3 (15.6)
The length of the implementation period at the time of the data collection, months	5.6 (0.9)	7.7 (1.8)
	n (%)	*n* (%)
Model still in use at the time of the data collection		
Yes	22 (69)	43 (83)
No	3 (9)	3 (6)
Unsure	7 (22)	6 (12)
Has received SPM training		
Yes	28 (88)	53 (95)
No	4 (12)	3 (5)
Has received post-training coaching		
Yes	15 (47)	38 (69)
No	17 (53)	11 (20)
Unsure	0 (0)	6 (11)
Working in a team with a structure that fulfils the fidelity criteria		
Yes	0 (0)	15 (27)
No	32 (100)	41 (73)

2.5.1.1. Adoption of the team structure. The managers completing the administrative form were asked to report the number and type of practitioners and other professionals working in each SPM team. An ideal SPM team would have the following characteristics: 1) a consultant social worker, 2) a maximum of three social workers 3) a clinician, 4) a coordinator and 5) a maximum of eight team members. The data was recoded by counting how many of these characteristics a team fulfilled each team being given a score from zero to five. Each team was further categorised in the fidelity categories as follows: high-fidelity implementation applied to teams fulfilling all five criteria; medium-fidelity implementation applied to cases in which the team had a clinician and a coordinator, but had not effected other changes; and if the team had a clinician but did not fulfil any other criteria, the fidelity level was low.

2.5.1.2. The extent of the clinician's work in teams. The managers were asked for what percentage of their weekly working hours the clinician(s) were available to the team. The intention was for each team at least to have a clinician present in the weekly meetings, which would mean a contribution of 10 per cent of weekly working hours. Hence, 10 per cent constitutes the high-fidelity threshold, 1–9 per cent medium-fidelity, and no input low-fidelity.

2.5.1.3. The extent of the coordinator's work in teams. The managers were asked for what percentage of their weekly working hours the coordinator(s) were available to the team. The coordinator was expected to take notes in the weekly meetings and to help social workers with other administrative tasks. Hence, 50–100 per cent of weekly working hours would meet the high-fidelity criteria, 21–49 per cent medium-fidelity and 0–20 per cent low-fidelity.

2.5.1.4. The number of cases discussed in the team meetings. The social workers completing the survey were asked how many cases in total they had discussed with their SPM team in the weekly meeting during the implementation period. This number was divided by the number of months the implementation had lasted (excluding one summer month), according to the respondent. High fidelity required covering a minimum of four cases per social worker per implementation month, medium fidelity 2–3 cases and low fidelity one case or less.

2.5.1.5. The use of key techniques. The social workers were asked if they had used genograms and hypotheses during the implementation. There were four response categories: "Not at all", "With one service user", "With more than one service user" and "I can't say / I do not recognise this technique". The fidelity threshold for this indicator was considered high if the respondent had used both key techniques with multiple service users, medium if he or she had used either genograms or hypotheses with multiple service users, and low if he or she had used these tools with one service user, or not with any.

2.5.1.6. The adoption of systemic thinking and the relevant techniques. Depending on how the interviewees described their practice, the teams were graded on a scale ranging from zero to 10 (0 = no change, 10 = major change) based on a detailed scale (see Table 3). The scale was constructed in collaboration with the second author in line with the social workers' descriptions in the interviews concerning the adoption of systemic thinking and techniques. The interview protocol is described in Section 2.5.2.1. From each transcript, the lead author assessed the extent to which the participants described their a) systemic thinking using related terms and ideas, b) their use of systemic thinking in practice with families, and c) their use of systemic techniques in practice. All the teams were further categorised in low-, medium- or high-fidelity groups (low: grades 0–3, medium: 4–6, high: 7–10). The interviewees in three of the teams showed considerable variation in their adoption level and

Table 2
An overview of the uses of quantitative and qualitative data for the evaluation of fidelity and the influencing factors, based on the modified CFIF (Hasson, 2010).

FIDELITY			
Core component	Indicator	Fidelity thresholds	Data source
Team structure	Content Adoption of the team structure -Consultant SW -Max. 3 SWs -Coordinator -Clinician -Max. 8 team members	High: All structural changes completed Medium: Involving the clinician and coordinator, team size too large Low: Involving the clinician, no coordinator, team size too large	Administrative data
	Dose The amount of a clinician's work in teams, hours per week	High: min. 10% Medium: 1–9% Low: No input	Administrative data
	Dose The amount of a coordinator's work in teams, hours per week	High: 50–100% Medium: 21–49% Low: 0–20%	Administrative data
Team meetings	Coverage The number of cases discussed in a team meeting	High: min. 4 cases Medium: 2–3 cases Low: max. 1 case per SW per implementation month	SW survey
Systemic practice	Content Use of key techniques	High: Use of both techniques with multiple service users Medium: Use of one technique with multiple service users Low: Use of techniques in a single case or none	SW survey
	Content Adoption of systemic thinking and techniques	High: 7–10 Medium: 4–6 Low: 0–3 grade in scaling the systemic practice	SW interviews
	Dose Intensive casework	The proportion of cases in which the SW can work intensively High: 80–100% Medium: 21–79% Low: 0–20%	SW survey

INFLUENCING FACTORS		
Factor	**Question**	**Data source**
Intervention complexity	How complex is the model?	SW interviews and surveys
Facilitation strategies	What strategies (e.g., manuals, guidelines, training and coaching) were used to support the implementation and how were they perceived by the social workers?	SW interviews and surveys
Participant responsiveness	How did the social workers and families engage with the model (e.g., satisfaction, enthusiasm, perception of outcomes of the intervention)?	SW interviews and surveys Service user interviews
Organisational factors	What organisational-level factors affected the implementation?	SW interviews and surveys

Note. SW, social worker.

were therefore assessed individually. Thus, their team grading was based on the median of the individual grades. Other more unified teams were given a joint grade without an individual analysis.

2.5.1.7. Intensive case work. The social workers were asked about the number of cases since the beginning of the implementation in which they had been able to work more intensively than previously. This number was then divided by the total number of cases with which the social worker was dealing at the time. The resulting measure was used to indicate the dose of systemic practice. This new variable was categorised in three fidelity groups, as follows: high fidelity, including social workers who reported working more intensively with 80 per cent of their cases or more; medium fidelity, including respondents who had been working with between 21 and 79 per cent; and low fidelity, referring to those working with a maximum of 20 per cent of their cases.

2.5.2. Influencing factors

2.5.2.1. Focus groups with social workers. The social workers were asked semi-structured questions concerning their views on the SPM and their implementation experiences. The interview protocol covered the following themes: 1) Describing the model at the site, 2) Experiences of the model and its implementation, 3) Experiences of implementation support (e.g., training), 4) The work environment (e.g., well-being) and 5) Perceptions of the potential use of the model in the future. An example of an interview question addressing fidelity was: "Could you describe the systemic model you have implemented this year?". After they had done so the interviewees were shown the components of the model on paper and were asked probing questions about the delivery of each one. Examples of questions addressing influencing factors included: "What has gone well in the implementation of the model?"; "What has been challenging?"; "What would you do differently in the implementation?"; "What do you think of the model at the moment?"; "On a scale from one to five, how motivated are you to use the model?"; "How clear does the model seem to you?"; and "What factors have supported you in the implementation?" Further probing questions were asked about the practitioners' responses, including their perceptions of the training.

2.5.2..2. Individual interviews with service users. The children and

Table 3
Qualitative assessment of the adoption of systemic thinking and techniques.

GRADE	DESCRIPTION
Low 0–3: No or little change in adopting systemic thinking in social work practice	**The interviewees:**
0	Do not describe any change in thinking and say that they have not changed their practice
1	Mention some systemic principles, but do not describe their relationship with social work practice/view the connection as distant
2	Mention some systemic principles, but have not reflected on how the principles relate to their own practice
3	Mention some systemic principles and have reflected on how the of principles relate to their own practice
Medium 4–6: Signs of systemic thinking and practising the techniques	
4	Mention some systemic principles and techniques as well as planning to use the techniques in practice
5	Briefly describe a change in thinking towards a systemic approach and mention obtaining ideas from the systemic team meeting for their own practice
6	Briefly describe a change in thinking and have purposefully attempted to apply systemic ideas or techniques in practice (e.g., used a timeline or a question that the clinician proposed)
High 7–10: Evidence of the application of systemic thinking and techniques in practice	
7	Describe a change in thinking with a few practical examples as well as indicating the use of systemic techniques or the active application of systemic ideas in practice
8	Describe a change in thinking with some practical examples as well as indicating the practising of systemic thinking and key systemic techniques (e.g., have used a genogram)
9	Elucidate a change in thinking with several practice examples and systemic terms, as well as indicating the use of the techniques several times in practice, with good results
10	Elucidate a change in thinking with several practical examples and systemic terms, and indicate the continuous use of the techniques in practice, with good results

parents were asked semi-structured questions addressing their current perceptions of child protection practice, whether practitioners had used systemic techniques with them, and their views on these techniques. The interview protocol covered the following themes: 1) Background (e.g., involvement with child protection services), 2) Meetings with social workers, 3) Experiences of systemic social work practice, 4) Participation and communication with social workers and 4) Satisfaction with and expectations of child protection services. Questions addressed to service users focusing on perceptions of practice included: "Could you describe your meetings with your social worker?" and "On a scale from one to five, how satisfied are you with the child protection services at the moment?" Questions focusing on systemic techniques included: "Have social workers drawn a genogram with you to discuss your family members, and if so, what did you think about it?"

2.5.2.3. SPM clarity. Eight items were used to measure how the social workers responding to the survey perceived the clarity of the SPM. The main question was "In adopting the new approach in your practice, how clear do you find the following aspects of the systemic model?" The specific aspects were: "The composition of the team and the different roles"; "The structure of the weekly team meetings"; "The methods used in the weekly team meetings"; "The basics of systemic thinking and family therapy"; "Formulating hypotheses"; "Drawing a genogram"; "Dealing specific child protection issues systemically" and "Assessing safety and risk in systemic practice". The social workers gave their responses on a five-point scale ranging from one, clear enough to five, not clear at all. A sum variable indicating *overall SPM clarity* was calculated, ranging from eight to 40 ($M = 19.48$, $SD = 6.40$).

2.5.2.4. Satisfaction with the training. The following seven items were used to measure how satisfied the social workers were with the training: "The training prepared me to implement the systemic model"; "The training gave me a clear understanding of how to implement the systemic model in daily practice"; "The training was concrete enough"; "The training material was useful"; "The length of the training was suitable"; "There was enough material on how to put the systemic model into practice"; and "The training was well delivered". The responses were given on a five-point Likert scale ranging from one, Strongly agree, to five, Strongly Disagree. Single items were recoded by combining categories 1 and 2, and 4 and 5 to measure the proportion of respondents agreeing or disagreeing with the statements. In addition, a sum variable indicating *overall satisfaction with the training* was calculated, ranging from seven to 35 ($M = 21.78$, $SD = 7.69$).

2.5.2.5. Experiences of the weekly team meetings. How the respondents experienced the team meetings was measured on two items - "The collective reflection during the weekly team meetings has helped me to do my work better"; and "I have received the necessary support from the team meetings" – rated on a five-point Likert scale ranging from one, Strongly agree, to 5 Strongly Disagree. A sum variable was calculated ranging from two to 10 ($M = 4.16$, $SD = 2.18$).

2.5.2.6. Experiences of support received from a clinician. Clinician support was also measured on two items - "The clinician has helped me to understand the family from a new perspective" and "The clinician has helped me to plan how to proceed with the family in a new way" – rated on a five-point Likert scale ranging from one, Strongly agree, to five, Strongly Disagree. A sum variable was calculated ranging from two to 10 ($M = 4.02$, $SD = 2.01$).

2.5.2.7. Satisfaction with the SPM and its implementation. W*illingness to continue using the SPM* was measured on one item, "I want our team to continue using the systemic model", as was *willingness to recommend the SPM to colleagues,* "I could recommend the systemic model to my colleagues", and the burden caused by the implementation, "The implementation of the systemic model has been an additional burden". In each case the responses were given on a five-point Likert scale ranging from one, Strongly agree, to five, Strongly Disagree.

2.5.2.8. Caseload. The respondents were asked to report the total number of child-protection cases they were currently dealing with.

2.6. Data analysis

2.6.1. Quantitative analysis

The quantitative data was organised and recoded (see Section 2.5), and the percentage frequencies were calculated. The Spearman correlation coefficient was calculated to examine the correlation between the variables, and a one-way analysis of variance (ANOVA) was used to

Table 4
The themes generated through the qualitative analysis with sub-themes, example citations and associated implementation categories based on the participants' perceptions.

Theme derived from the CFIF (i.e., influencing factor)	Sub-theme	Example citation	The example citation categorised as a barrier or a facilitator
Intervention complexity	A lack of clarity in the content of systemic practice	"Although we had the training, what is that systemic practice in the end, and how does it differs from what we already do?"	Barrier
	Varying learning experiences	"This jumping into the unknown [using new skills with families] was sometimes unpleasant."	Barrier
Facilitation strategies	Training	"It was very good, that training."	Facilitator
	Coaching	"Our clinician has actually taught us to use those circular questions."	Facilitator
Participant responsiveness	SW: positive experiences of the SPM	"Our clinician often provides the kinds of new perspectives that can really support your own practice."	Facilitator
	SW: negative experiences of the SPM	"If [the meeting] is all that jibber-jabber, then no, neither I nor my clients benefit from it."	Barrier
	Family experiences of the SPM	"Discussing our family situation in the meeting triggers thoughts. The professionals took different family members' roles and then shared what they thought about the situation. I think those were quite good thoughts."	Facilitator
Organisational factors	Caseloads	"SW4: Well in this context I think it's very difficult to implement. I: And by context you mean? SW4: Well, maybe time, caseloads."	Barrier
	Leaders' commitment to change	"At the same time [leaders] want us to do our work better. But then nothing… it eats me that the agency washes its hands completely of the resource situation."	Barrier
	Staff turnover	"People come and go. That is one of the biggest changes."	Barrier

Note. SW: social worker; I, interviewer.

examine the differences between the means of fidelity groups. The SPSS Statistics 24 package was used for the quantitative analyses.

2.6.2. Qualitative analysis

The lead author applied theoretical thematic analysis, which involves identifying the themes in a 'top-down' fashion to explore a theoretical framework (Braun & Clarke, 2013, p. 178), namely the CFIF in this article. The analysis proceeded in the following six steps: (1) reading and familiarisation, (2) coding the dataset, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes and (6) writing and finalising the analysis (Braun & Clarke, 2013). The ATLAS.ti program was used to code the transcripts.

Having done an initial reading of the interviews (phase 1), the lead author coded them all, guided by the CFIF (phase 2). The entire data set was coded, but with a particular focus on the influencing factors of interest (Table 2): whether they appeared in the data and what the interviewees were discussing concerning them. Given that the central organising concepts, i.e., themes (Braun & Clarke, 2013, 224), were generated from the CFIF (intervention complexity, facilitation strategies, participant responsiveness and organisational factors), the search for patterns concentrated on the related sub-themes (phase 3). All the codes added under the themes were collated and analysed to identify patterns. Visual mapping of themes and ATLAS.ti code reports were used in the identification process. The themes and sub-themes were reviewed multiple times to achieve coherence, and were discussed with the second author (phase 4). A keyword search was used to identify potentially overlooked parts. When all the themes had been defined (phase 5), all the citations connected to them were categorised as either barriers or facilitators (see Table 4). Finally, the relationships between the different themes were explored and the thematic map was finalised (phase 6). The analysis was deepened by means of data quantification and an analysis of code co-occurrence. The quantification focused on themes (i.e., influencing factors) and their categorisation as either barriers or facilitators (e.g., how many citations under the theme 'intervention complexity' were described as barriers or facilitators). The number of times a code was mentioned was counted. The unit of analysis was the entire sentence or paragraph as spoken by the interviewee. All occurrences were counted. The lead author translated the citations selected for this article into English (reported within quotation marks).

2.6.3. Mixed-methods analysis

Having conducted separate quantitative and qualitative analyses, we identified commonalities and discrepancies between our findings to complete the results in the interpretation phase (Creswell & Plano Clark, 2008). We also analysed the participants' perceptions on implementation barriers and facilitators, and formulated hypotheses concerning their potential linkages to the fidelity of the model. We tested the key hypotheses with the quantitative data.

3. Results

3.1. Fidelity

We observed considerable discrepancies in the different areas of fidelity measurement. Table 5 summarises the results pertaining to fidelity.

3.1.1. Team structure

All the teams involved included a clinician and a consultant social worker (see Table 6). However, we learned from the interviews that some teams were without a clinician after only a few months, and some received very little support from the consultant social worker. The clinicians' contributions to the team in terms of working hours varied from three to 110 per cent (several clinicians in one team) of working hours per week (M = 18%), whereas the variation in the coordinators' work was between three and 100 per cent (M = 26%). One fourth of the teams did not have a coordinator. When we looked at the indicators related to the systemic team as a whole we found that only six per cent of the teams met all three fidelity criteria (i.e., the right structure and high contributions from both the clinician and the coordinator), whereas 39 per cent achieved high scores on two indicators, and 41 per cent on only one indicator. In addition, 14 per cent of the teams did not achieve a high score for fidelity on any of these indicators.

3.1.2. Team meetings

The vast majority (86%) of the social workers achieved low scores in terms of the number of cases the social workers presented to the team in the weekly meeting during the implementation period: the responses varied from one case to 35 per social worker (M = 7). Proportional to the implementation months (one summer month excluded), this meant that

Table 5
A summary of the results concerning fidelity.

Areas to measure	Indicator	The level of fidelity achieved
Details of content	Adoption of the team structure	From all teams 45% scored high
Was each of the core components implemented as intended?	-Consultant SW -Max. 3 SWs -Coordinator -Clinician -Max. 8 members	32% medium 24% low adherence to the content
	Use of key systemic techniques with service users	From all SWs 36% scored high 34% medium 30% low adherence to the content
	Adoption of systemic thinking and techniques	From nine local teams 3 scored high 4 medium 2 low adherence to the content
Dose (frequency/duration)	The amount of a clinician's work in teams, hours per week	From all teams, 76% scored high 22% medium 3% low
Were the core components implemented as often and for as long as intended?		dose in the clinician's contribution
	The amount of a coordinator's work in teams, hours per week	From all teams, 21% scored high. 32% medium. 47% low dose in the coordinator's contribution
	Intensive casework	From all SWs, 2% scored high 20% medium 79% low dose of intensive casework
Coverage	The number of cases discussed in a team meeting during the intervention period	From all SWs, 5% scored high 9% medium 86% low in coverage.
What proportion of the target group participated in the intervention?		

the social workers, on average, brought only one case per month to the case discussion. At the time of the follow-up survey the intervention period had lasted from between two to 12 months ($M = 8$), depending on the team. We observed in the qualitative sample that the implementation was hampered in three teams before the follow-up data collection: one had no clinician, and two involved the clinician approximately only once a month.

3.1.3. Systemic practice

Only one third of the social workers scored highly on the use of key systemic techniques. The majority (79%) achieved low scores in terms of the frequency of family meetings, and one in four (25%) had not been able to increase the intensity with any of their cases. None of the social workers scored highly, whereas 27 per cent achieved low scores on both of these indicators. Our qualitative analysis of the adoption among social workers of systemic thinking and techniques revealed considerable variation (range 3–9) between the teams involved (see Table 7). Interestingly, we detected variability both between teams in the same sites as well among individuals in the same team.

Table 6
Results related to the structural content of the team.

Adoption of the team structure: involvement of professionals	Percentage (%) of all teams participating in the study ($n = 38$)
Consultant SW	100%
Clinician	100%
Coordinator	76%
Max. 3 social workers	62%
Max. 8 members in a team	84%
Teams fulfilling all team characteristics listed in the table	45%

3.2. Influencing factors

Among the citations in the interviews related to the influencing factors ($N = 877$), 37 per cent describe facilitators and 63 per cent barriers, thereby highlighting implementation-related challenges. Depending on the circumstances, the influencing factors might either impede or encourage the implementation (e.g., some social workers perceived the training as helpful, others disagreed). As anticipated, the factors also involved various interconnections.

3.2.1. Intervention complexity

Intervention complexity was coded as a barrier in 90 per cent of the interview citations in which it is mentioned ($N = 138$). We identified two themes: a lack of clarity in the content of systemic practice and varying learning experiences of the new approach. Although the team structure and the purpose of the team meetings were generally clear to the social workers, a number of them found the content of systemic practice in real-world settings, including the use of systemic techniques with families, somewhat confusing. As one of them stated, she had "no clear idea what should concretely happen in family meetings" when applying systemic practice. In comparison with the results of our survey, it seems that systemic practice was somewhat clear in theory. Systemic thinking and family therapeutic orientation were clear enough to 71 per cent of the respondents, and with regard to the techniques, 52 per cent were clear about formulating hypotheses, and genograms were clear to 75 per cent.

Several interviewees further remarked that the new way of working required plenty of learning. Some, for example, found participating in reflective case discussions in the team meetings more demanding than in the previous solution-oriented team discussions. Those who described applying systemic thinking and techniques in their practice demonstrated in-depth professional learning and related positive stress. Engaging with systemic practice had forced them to step out of their "comfort zone", thus changing their perceptions of social work practice and acquiring new communication skills. On the other hand, a few of them described feelings of frustration and incompetence when they tried to use difficult techniques in their work practices. Forging closer relationships and discussing childhood traumas also made some of them doubt their own practical skills.

3.2.2. Facilitation strategies

Two themes related to strategies intended to facilitate implementation included the team-based training and coaching in systemic practice. Facilitation strategies were described as facilitative in only 31 per cent of the citations in which they were mentioned ($N = 128$).

The training in particular divided opinions. Some social workers described it as "good" and "necessary", providing essential information about the model, whereas others said it was "superficial" and lacking in "structure" and "consistency", leaving them confused about the systemic practice in place prior to the implementation. All the interviewed teams had expected the training to offer concrete suggestions and guidance on techniques for engaging in systemic social work, an "understanding of what systemic practice actually is". Moreover, some social workers felt that the trainers belittled their current ways of working, thereby fuelling conflict between the new approach and service as usual. Some interviewees would have preferred the training to have lasted longer than

Table 7
Qualitative results concerning the adoption of systemic thinking and techniques

IMPLEMENTATION SITE AND TEAM	*M*	EXAMPLE CITATIONS
Site 1		
Team 1	9	SW5: "It is like their family system, and the purpose is that we professionals will withdraw at some point." SW3: "When we drew a genogram, we were really surprised that the father actually started to talk about those things [family history and relationships]. And we progressed with their case in a whole new way."
Team 2	6	SW2: "I also think that not-knowing assigns responsibility to the families that they will resolve [their own problems]." SW4: "I don't actually know what that systemic thinking is."
Team 3	8	SW1: "I think that drawing a genogram has had the most significant impact on my thinking." SW4: "I don't think that I would have achieved such good results as I did together with the team." SW3: "I don't think my thinking has changed at all [laughs]. If we think about what we have done, I think we have done systemic practice all the time."
Team 4	4	SW1:"The same for me, I have added those [systemic] ideas to the child-protection plan and have gone through them on a general level… [−−] but I haven't changed my actual practice at all."
Site 2		
Team	4	"SW1: I think perhaps we have adopted systemic thinking somehow. But maybe we could use those techniques more. I: What techniques have you used so far? SW2: Well I don't know if I have used any of those techniques we went through in the training. SW1: Probably mostly the same [laughs] and familiar ones, cards and games."
Site 3		
Team 1	6	SW2: "This implementation has activated me to try some techniques, for example, I had never used that reflective discussion in a family meeting before."
Team 2	3	"I: [D]o you feel that you work differently than before [the implementation]? All SWs: No."
Team 3	3	SW1: "It is very difficult to identify any changes in my own practice. I just noticed that I don't know these [systemic] techniques."
Team 4	9	SW2: "We used [the genogram], and met both parents separately many times." SW1: "Since the implementation we have started to think more not only about the families' systems, but also that we are part of those systems."

six days, and to have continued for longer than a few months. Consistently, our quantitative findings reveal that, on the national level, 40 per cent of social workers were satisfied with their training, whereas 43 per cent were not. Likewise, only a quarter of the survey respondents perceived the training as concrete enough, whereas a third said it had given them a clear idea of how to follow systemic practice in their work. Some of the interviewees suggested that the lack of clarity might be connected to the poor quality of the ToT and the trainers' inexperience in systemic social work. Indeed, 29 per cent of the trainers felt that they were not properly equipped to instruct the teams in the use of genograms or the art of hypothesising (Aaltio & Isokuortti, 2019).

On the general level, training-related challenges could explain the lack of clarity concerning the content of systemic practice described in Section 3.2.1. The results of the Spearman correlation indicate a significant positive association between overall satisfaction with training and overall SPM clarity ($rs(48) = 0.36$, $p = .010$). The quantitative findings also reveal that, in the view of 46 per cent of the respondents, there was not enough supporting material to enable them to incorporate systemic practice into their routines.

The coaching provided by the trainers was not considered particularly useful in many of the interviewed teams. However, the view in four of them was that the clinician or other colleague trained in family therapy not only helped the social workers to engage in systemic practice (see Section 3.2.3), but also supplemented the formal training and coaching with their hands-on teaching. As one of them concluded: "And I feel that [our clinician] has actually taught us to use those circular questions and to externalise [problems], and the dialogical [approach] … all those things, really hands-on teaching, and it has been, at least for me, extremely important".

3.2.3. *Participant responsiveness*

3.2.3.1. Social workers. Participant responsiveness was categorised in two themes: positive and negative experiences. Overall, the social workers were very satisfied with the SPM, which evidently facilitated its implementation. Of all the citations related to participant responses ($N = 377$), 71 per cent were positive.

Two elements in particular concerned positive experiences: the perceptions that SPM improved work-related wellbeing and practice, and that it provided a safe learning environment. For the most part, the interviewees' perceptions of the model's usefulness were connected to the clinician's involvement in the team meetings. In applying systemic thinking and techniques during the meetings the clinician helped the social workers to consider the families' situations from multiple perspectives and prevented them from making hasty decisions on the cases. The mutual reflection also eased the individual burden of responsibility and made it easier to plan interventions with the family. Some of the interviewees acknowledged that the model had not only enhanced their work-related well-being and sense of meaningfulness, but had also induced positive change among families, such as with case closure. Applying the techniques enabled them to evoke change in family members and help them "to take responsibility and come up with solutions" themselves, instead of the professionals making the decisions. The social workers also appreciated the coordinator's help in taking minutes in the meetings, scheduling appointments and contacting other service providers, all of which helped them to improve their practice.

According to the interviews, the practitioners in two teams had formed a safe learning environment enhancing the uptake of systemic practice. All these team members shared high learning motivation and had clear roles. One of the teams had a well-established relationship with their clinician. The other team saw the clinician only in biweekly meetings but received practical support both from a practitioner trained in family therapy and from their skilful consultant social worker. These positive interrelationships and hands-on coaching in systemic practice (see Section 3.2.2) gave the social workers the confidence to discuss their feelings in the team meetings and to try out the techniques in practice. Their positive experiences of practice enhanced their feelings of competence, which in turn strengthened their commitment to systemic practice.

The negative responses were associated with two elements: the low fidelity of the model and the impracticality of team meetings. Although the social workers were able to apply the systemic approach with a few families (see Section 3.1.3), they were still responsible for their whole caseload. The interviewees felt obliged to "prioritise" families involved in the intervention, which created a split sense of the reality: service as usual with some families and systemic practice with others. The "cheap

version" of the model was a major concern. One team suggested that their motivation for SPM would be optimal if they implemented "the ideal model". Despite the challenges, however, the social workers seemed to appreciate the model in itself. Moreover, although most of them were satisfied with the team meetings, a few thought they were too long and unhelpful. This finding could be attributed to the large team size (as many as six SWs), which was detrimental to reflective discussion and lengthened the meetings. In sum, the social workers' perceptions on negative experiences and their linkage to motivation indicate that low fidelity might decrease participant responsiveness. The discrepancy concerning the team meetings could also reflect inadequate integration between the therapeutic approach and statutory social work.

A comparison of the qualitative results with our survey results revealed that 79 per cent of the survey respondents ($n = 56$) wished to continue using the model, and 76 per cent would recommend it to colleagues. As indicated in the interviews, the willingness to continue the implementation among all survey respondents was associated with their positive experiences of the weekly team meetings ($rs(51) = 0.78$, $p < .001$), and support from the clinician ($rs(51) = 0.63$, $p < .001$): yet, 44 per cent of them experienced the implementation as a burden. Nevertheless, half of these respondents wanted to continue using the model, which illustrates its high level of acceptability.

3.2.3.2. Families. Despite the limited evidence of systemic practice on the family level, most interviewees were satisfied with the help received from the child protective services. In a similar vein, a forthcoming pilot study reveals high service-related satisfaction among service-user respondents. However, there were no differences between the service-as-usual and the SPM groups, implying that the parents were equally satisfied in both. According to the interviewees, service users whose social worker applied systemic techniques with them or had attended the reflective team meeting themselves felt that it gave them new perspectives on their situation. One father said that discussing family questions with several professionals in the meeting was, at its best, "empowering". However, five of the six users who knew that their case had been discussed in the team meeting were somewhat disappointed that the discussion had little impact on their family's service, indicating relatively weak linkage between the meetings and the practice. As one mother pointed out, "of course, the social workers can think about good practices or means to help a family, but I think that those means should be brought concretely to the family level".

3.2.4. Organisational factors

The following three themes were connected to organisational factors: caseloads, the leaders' commitment to change and staff turnover. Of all the citations related to organisational factors ($N = 368$), 91 per cent were coded as barriers. The social workers were able to discuss only a few of their cases in the team meetings, given their high caseloads, and consequently engaged in systemic practice mainly with these specific families (see Section 3.2.3). One interviewee who was dealing with 47 child protection cases said that her team frequently had to cancel meetings due to time pressure. There was too little time to meet the families and implement new techniques in any case, which in turn weakened the effect of the training: as one interviewee remarked, "I haven't had time even to try another kind of approach in my work".

According to our survey results, and contrary to the nationally recommended 20 cases per social worker, the average load increased from 32 to 35 cases in the course of the implementation. Surprisingly, the ANOVA results indicated that there were no significant differences among the fidelity groups regarding the number of cases discussed in a team meeting ($F(2, 53) = 0.61$, $p = .548$), the use of key techniques ($F(2, 53) = 0.46$, $p = .635$), or intensive casework ($F(2, 53) = 1.57$, $p = .218$). The Spearman correlation also gave interesting results indicating the lack of an association between caseloads and the willingness to continue with the SPM ($rs(51) = 0.25$, $p = .071$), and between caseloads and experienced burden ($rs(54) = -0.26$, $p = .058$). However, as in the interviews, the survey data revealed a significant association between caseloads and the willingness to recommend SPM to colleagues ($rs(52) = 0.34$, $p = .012$). As anticipated, respondents with higher caseloads were less willing to recommend the model.

The interviewees were divided in their opinions about the commitment of agency leaders to the SPM. Those in site two in particular described their senior managers as supportive of their work in that the agency had lessened their caseloads prior to the intervention. Interestingly, this team scored only four on a scale from zero to ten indicating that organisational support was not connected with fidelity to systemic practice in this context (see Table 7). Several social workers at other sites were disappointed in their agency leaders' lack of commitment, such as not arranging reasonable caseloads and poorly communicating team responsibilities. Consequently, some teams wondered whether the "disorganised" nature of the project was attributable to the lack of vision or implementation strategy. Although most agencies carried out structural changes, in some teams the contributions of the clinician and the consultant social worker were cut during the implementation, which frustrated the social workers. Another stress-inducing factor was that some agencies had initiated simultaneous non-intervention-related reforms such as moving into new offices. The imbalance between the SPM objectives and the circumstances caused further frustration. In comparison, only one in ten survey respondents thought that systemic thinking was embedded in their whole organisation.

Finally, although some teams experienced little or no staff turnover, others went through major changes during the implementation. At its most severe, teams were left with only a few members who had participated in the training. Given the high staff turnover and frequent sickness absences, a number of interviewees expressed the view that the original idea of small teams with a maximum of three social workers was not feasible. On the other hand, the low staff turnover in other teams helped to establish stable interrelationships and fostered mutual learning (see Section 3.2.3.1).

4. Discussion

The purpose of this study was to assess the fidelity of the SPM and to formulate hypotheses concerning the potential influencing factors based on our mixed-methods data. We observed considerable variability in the areas of fidelity measurement. Unexpectedly, there was also high variation between individual teams within the same sites. Whereas some adhered to systemic practice, others scarcely used such techniques, and even discontinued their implementation during the course of the study. We identified several implementation barriers: (1) a lack of clarity regarding systemic practice (*intervention complexity*), (2) insufficient training (*facilitation strategies*), and (3) high caseloads and staff turnover, and a lack of leader commitment to change (*organisational factors*). Regardless of these challenges, the social workers positively engaged with the SPM (*participant responsiveness*). The teams that were more successful in adopting systemic practice enjoyed a positive learning environment (*participant responsiveness)* with little staff turnover (*organisational factors*) and received hands-on coaching from their team member who was specialised in systemic family therapy (*facilitation strategies*). It is interesting that although most teams reported several of the barriers listed above, they differed in their ability to engage in systemic practice. This finding underscores the significance of the facilitators.

Our results further support the association between different influencing factors (Carroll et al., 2007). For example, one hypothesis is that high motivation and a good team atmosphere could strengthen employee commitment to coaching, which in turn could be attributed to their skills and engagement with systemic practice. According to another hypothesis, against expectations, the level of fidelity might influence certain factors: low coverage could weaken participant motivation, for example, which in turn could hinder learning and systemic practice.

These hypotheses should be refined in future research, and tested in different contexts.

We identified the following similarities between our findings and the results of previous research on the original UK model. First, Bostock et al. (2017) and Laird et al. (2018) report variation in delivering intended structural changes and conducting systemic practice, having also reported similar organisational barriers such as high caseloads. Moreover, Berrick et al. (2016) found that the caseloads of English social workers were half the size of those of their Finnish colleagues (i.e., 19–21 children per worker in England contrasted with 46–48 in Finland). To increase the coverage of systemic practice, in other words to allow time to reflect on cases and to interact with families, it is crucial to decrease the workload of those concerned in all kinds of settings. Second, we found that the clinician's role in maintaining systemic practice was crucial (Bostock et al., 2017, 2019).

We also found certain differences with regard to previous research. First, our findings were contradictory in relation to UK evaluations implying high levels of satisfaction with training (Bostock et al. 2017; Dugmore, Partridge, Sethi, & Krupa-Flasinska, 2018; Laird et al., 2018). The inconsistency could be attributable to the high number of local trainers who received the same ToT but differed in terms of experience and possibly also in training and practice skills. The UK agencies received training and coaching from a social enterprise, whose founders led the systemic change in Hackney (Bostock et al., 2017). This resource is not easily transferrable to other countries, and we discuss the implications in Section 4.1. The vague intervention description known to impede implementation (Hasson, 2010) may also explain the lack of clarity concerning systemic practice in this study. Second, in contrast to findings in England (Bostock et al., 2017; Laird et al., 2018), the social workers in this study seldom discussed potential conflicts between the systemic approach and child risk management. This discrepancy could be attributed to the different orientations in the child protection systems in England and Finland: Finland has traditionally been characterised as family-service-oriented in contrast to the risk-oriented English system (Gilbert, 1997; Gilbert et al., 2011). Thus, it may be that a therapeutic orientation per se is more suited to the Finnish than to the English context. In fact, the foreign origin of the model came up in the discussions with social workers only a few times, which supports its transportability.

Finally, our results were mixed on the question of engaging the whole organisation, including senior leaders, in systemic change. On the one hand, the differing implementation outcomes, both highly positive and highly negative, within one site indicate that a reform of the whole system might not be a prerequisite for the implementation of the SPM, as implied in previous evaluations (Bostock et al., 2017; Laird et al., 2018; see also Sheehan et al., 2018). On the other hand, our findings indicate that leader support is essential in arranging the intended structural changes, providing the facilitation strategies and preventing staff turnover.

4.1. Implications for policy and practice

According to our findings, the SPM was widely accepted among social workers, but there were problems related to staff training and the transference of these ideas to social work practice, and it was difficult to create an organisational culture that would support systemic practice. Given that the model comprises multiple overlapping components and is demanding in terms of practical skills (see also Craig et al., 2013), to support its future implementation it would be useful to develop a manual based on the programme theory. This kind of guidance would also enhance the model's transportability. In particular, further attention should to be given to the roles of the consultant social worker and the coordinator, which have been found crucial for the maintenance of systemic practice in previous studies (Bostock et al., 2017; Forrester et al., 2013). Given that the training should give a clear and coherent picture of systemic practice that its users will understand, and equip them with necessary skills, future implementers should assess whether the ToT is the most effective solution. Furthermore, it is vital to maintain high-quality supervision to support systemic practice (Bostock et al., 2017, 2019). The need for ongoing technical assistance has also been noted in other studies (Meyers et al., 2012; Sanclimenti et al., 2017).

The results of this study, although preliminary, have implications concerning implementing practice models across countries. We have shown that without adequate implementation support, practice models presumably fail to achieve the anticipated outcomes, or succeed to a limited extent. Given the complex nature of children's services as a change environment (Mildon et al., 2013), implementers should assess their readiness for change and, when necessary, enhance the organisational capacity (e.g., resources, infrastructure) before introducing an innovation. To ensure that this happens, those in charge of developing and disseminating practice models should list the resources that are generally required for their implementation. It is also necessary to formulate a comprehensive implementation strategy, including long-term maintenance support (Mildon & Shlonsky, 2011). These measures will also help to avoid unintentional harmful implementation effects such as practitioner frustration and opportunity costs. Finally, inherent in practice models is the potential for improving social work practice with a distinct and consistent approach, but endless adaptation and 'hybrid models' create challenges in terms of both implementation and evaluation (Baginsky et al., 2020). Various frameworks could be used to facilitate decision-making in adapting the models to any given setting (Stirman et al., 2013; Miller et al., 2020).

4.2. Limitations and future research

Two major limitations of this study are the use of self-reported fidelity data and the lack of a validated fidelity-measurement tool for the SPM. However, the findings give useful insights that could enhance the development of such a tool in the future. Although the two researchers had independent responsibilities in terms of gathering and analysing specific datasets, we worked to minimise a potential source of bias by regularly discussing data collection and interpretation throughout the study process. The small sample size limits the representativeness, however. In addition, the unequal distribution of respondent social workers in the fidelity categories made further statistical analysis, especially analyses of variance, difficult to interpret. Nevertheless, the mixed-methods design allowed us to conduct a comprehensive analysis of fidelity, the influencing factors and their interrelationship (see also Toomey et al., 2020). Specifically, the qualitative assessment of systemic practice produced a detailed picture of the team-level differences. We have also demonstrated the need for future measurements of fidelity to focus not only on the agency level but also on the team and the individual levels. We acknowledge that observations of direct practice and team meetings would have enriched the interview and survey data in terms of assessing the fidelity of systemic practice. Finally, the study excluded the perspectives of leaders and trainers. We recommend further research with a stronger focus on organisational factors (including sufficient resourcing and the organisational culture) and the role of leaders, and on support for leaders in implementation efforts (see also Baginsky et al., 2020).

5. Conclusions

Changing practice in social care for children is challenging. Adding international transportability and the further adaptation of practice models to the implementation challenge considerably increases the complexity. Fidelity measurement helps to determine implementation outcomes, which are vital not only to leaders and practitioners but also to outcome evaluators. Furthermore, a better understanding of the barriers and facilitators will help policy makers and professionals in future implementation efforts in this field. In sum, we suggest that the RSW model could be adopted in new settings with adequate

implementation support, but we also demonstrate the complexity of children's social care as a change environment.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Nanne Isokuortti: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Project administration. **Elina Aaltio:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Mirja Satka, Nelli Hankonen and Mengyao Lu for their insightful comments on the manuscript. We also thank the Master's students who contributed to the data collection. We are also grateful to Joan Nordlund for help with language revision. Finally, we give special thanks to the National Institute of Health and Welfare and all the professionals and families who participated in this research project.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.childyouth.2020.105647.

References

Aaltio, E. & Isokuortti, N. (2019). Systeemisen lastensuojelun toimintamallin pilotointi. Valtakunnallinen arviointi [The initial implementation of the Systemic Practice Model for child protection. National evaluation]. Report 3/2019. Helsinki: National Institute for Health and Welfare.

Antle, B. F., Christensen, D. N., van Zyl, M. A., & Barbee, A. P. (2012). The impact of the Solution Based Casework (SBC) practice model on federal outcomes in public child welfare. *Child Abuse & Neglect, 36*(4), 342–353. https://doi.org/10.1016/j.chiabu.2011.10.009

Baginsky, M., Ixer, G., & Manthorpe, J. (2020). Practice frameworks in children's services in England: An attempt to steer social work back on course? *Practice (Birmingham, England)*, 1–17. https://doi.org/10.1080/09503153.2019.1709634

Barbee, A., Christensen, D., Antle, B., Wandersman, A., & Cahn, K. (2011). Successful adoption and implementation of a comprehensive casework practice model in a public child welfare agency: Application of the Getting to Outcomes (GTO) model. *Children and Youth Services Review, 33*(5), 622–633. https://doi.org/10.1016/j.childyouth.2010.11.008

Bellg, A., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D., Ory, M., … Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology, 23*(5), 443–451. https://doi.org/10.1037/0278-6133.23.5.443

Berrick, J., Dickens, J., Pösö, T., & Skivenes, M. (2016). Time, institutional support, and quality of decision making in child protection: A cross-country analysis. *Human Service Organizations: Management, Leadership & Governance, 40*(5), 451–468. https://doi.org/10.1080/23303131.2016.1159637

Bostock, L., Forrester, D., Patrizo, L., Godfrey, T., Zonouzi, M., Bird, H., … Tinarwo, M. (2017). *Scaling and deepening Reclaiming Social Work model: Evaluation report.* London: Department for Education.

Bostock, L., Patrizo, L., Godfrey, T., & Forrester, D. (2019). What is the impact of supervision on direct practice with families? *Children and Youth Services Review, 105*. https://doi.org/10.1016/j.childyouth.2019.104428

Braun, V., & Clarke, V. (2013). *Successful qualitative research. A practical guide for beginners*. London: SAGE Publications.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. (Debate). *Implementation Science, 2*(40), 40.

Crea, T. M., Usher, C. L., & Wildfire, J. B. (2009). Implementation fidelity of Team Decisionmaking. *Children and Youth Services Review, 31*(1), 119–124. https://doi.org/10.1016/j.childyouth.2008.06.005

Creswell, J. W., & Plano Clark, V. L. (2008). *The mixed methods reader*. Thousand Oaks, Calif: SAGE Publications.

Cross, S., Hubbard, A., & Munro, E. (2010). Reclaiming Social Work London Borough of Hackney Children and Young People's Services: Independent evaluation.

Dugmore, P., Partridge, K., Sethi, I., & Krupa-Flasinska, M. (2018). Systemic supervision in statutory social work in the UK: systemic rucksacks and bells that ring. *European Journal of Social Work, 21*(3), 400–414. https://doi.org/10.1080/13691457.2018.1446914

Durlak, J. A., & Dupre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Forrester, D., Westlake, D., McCann, M., Thurnham, A., Shefer, G., Glynn, G., & Killian, M. (2013). *Reclaiming social work? An evaluation of systemic units as an approach to delivering children's services*. Luton: University of Bedfordshire.

Gilbert, N. (1997). *Combatting child abuse: International perspectives and trends*. New York: Oxford University Press.

Gilbert, N., Parton, N., & Skivenes, M. (2011). *Child protection systems: International trends and orientations*. New York (N.Y.): Oxford University Press.

Gillingham, P. (2018). Evaluation of practice frameworks for social work with children and families: Exploring the challenges. *Journal of Public Child Welfare, 12*(2), 190–203. https://doi.org/10.1080/15548732.2017.1392391

Goodman, S., & Trowler, I. (2012). *Social work reclaimed: Innovative frameworks for child and family social work practice*. London: Jessica Kingsley Publishers.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and evaluating complex interventions: The new Medical Research Council guidance. *International Journal of Nursing Studies, 50*(5), 587–592. https://doi.org/10.1016/j.ijnurstu.2012.09.010

Greenhalgh, T., Pawson, R., Wong, G., Westhorp, G., Greenhalgh, J., Manzano, A. & Jagosh, J. (2017). What realists mean by context; or, why nothing works everywhere or for everyone. The RAMESES II Project.

Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation Science: IS, 5*(1), 67. https://doi.org/10.1186/1748-5908-5-67

Isokuortti, N., Aaltio, E., Laajasalo, T., & Barlow, J. (2020). Effectiveness of child protection practice models: A systematic review. *Child Abuse & Neglect, 108*, 104632. https://doi.org/10.1016/j.chiabu.2020.104632

Laird, S. E., Morris, K., Archard, P., & Clawson, R. (2017). Working with the whole family: What case files tell us about social work practices. *Child & Family Social Work, 22*(3), 1322. https://doi.org/10.1111/cfs.12349

Laird, S. E., Morris, K., Archard, P., & Clawson, R. (2018). Changing practice: The possibilities and limits for reshaping social work practice. *Qualitative Social Work, 17*(4), 577–593. https://doi.org/10.1177/1473325016688371

Lahtinen, P., Männistö, L. & Raivio, M. (2017). Kohti suomalaista systeemistä lastensuojelun toimintamallia: Keskeisiä periaatteita ja reunaehtoja [Toward the Finnish Systemic Practice Model for child protection. Core principles and preconditions]. Discussion paper 7/2017. Helsinki: National Institute for Health and Welfare.

Lambert, D., Richards, T., & Merrill, T. (2016). Keys to implementation of child welfare systems change initiatives. *Journal of Public Child Welfare, 10*(2), 132–151. https://doi.org/10.1080/15548732.2015.1113226

Meyers, D., Durlak, J., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*(3–4), 462–480. https://doi.org/10.1007/s10464-012-9522-x

Mildon, R., & Shlonsky, A. (2011). Bridge over troubled water: Using implementation science to facilitate effective services in child welfare. *Child Abuse & Neglect, 35*(9), 753–756. https://doi.org/10.1016/j.chiabu.2011.07.001

Mildon, R., Dickinson, N., & Shlonsky, A. (2013). Using implementation science to improve service and practice in child welfare. Actions and Essential Elements. In A. Shlonsky, & R. Benbenishty (Eds.), *From evidence to outcomes in child welfare: An international reader* (pp. 83–101). Oxford: Oxford University Press.

Miller, C., Wiltsey-Stirman, S., & Baumannn, A. (2020). Iterative Decision-making for Evaluation of Adaptations (IDEA): A decision tree for balancing adaptation, fidelity, and intervention impact. *Journal of Community Psychology, 48*(4), 1163–1177. https://doi.org/10.1002/jcop.22279

Morris, K., Archard, P. J., Laird, S. E., & Clawson, R. (2018). Family experiences of children's social care involvement following a social work change programme. *Journal of Social Work Practice, 32*(3), 237–250. https://doi.org/10.1080/02650533.2017.1326473

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: SAGE.

Pipkin, S., Sterrett, E. M., Antle, B., & Christensen, D. N. (2013). Washington State's adoption of a child welfare practice model: An illustration of the Getting To Outcomes implementation framework. *Children and Youth Services Review, 35*(12), 1923–1932. https://doi.org/10.1016/j.childyouth.2013.09.017

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., … Hensley, M. (2010). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(2), 65–76. https://doi.org/10.1007/s10488-010-0319-7

Roberts, Y. H., Caslor, M., Turnell, A., Pearson, K., & Pecora, P. J. (2019). An international effort to develop a fidelity measure for signs of safety®. *Research on Social Work Practice, 29*(5), 562–571. https://doi.org/10.1177/1049731518754724

Sanclimenti, J. G., Caceda-Castro, L. E., & Desantis, J. P. (2017). Child welfare practice model implementation projects: Lessons learned. *Journal of Public Child Welfare, 11*(3), 279–298. https://doi.org/10.1080/15548732.2016.1275920

Sheehan, L., O'Donnell, C., Brand, S. L., Forrester, D., Addis, S., El-Banna, A., … Nurmatov, U. (2018). *Signs of Safety: Findings from a mixed-methods systematic review focussed on reducing the need for children to be in care.* London: What Works Centre for Children's Social Care.

Stirman, S., Miller, C., Toder, K., & Calloway, A. (2013). Development of a framework and coding system for modifications and adaptations of evidence-based interventions. *Implementation Science: IS, 8*(1), 65. https://doi.org/10.1186/1748-5908-8-65

Sundell, K., Ferrer-Wreder, L., & Fraser, M. W. (2014). Going global: A model for evaluating empirically supported family-based interventions in new contexts. *Evaluation & the Health Professions, 37*(2), 203–230. https://doi.org/10.1177/0163278712469813

Toomey, E., Hardeman, W., Hankonen, N., Byrne, M., Mcsharry, J., Matvienko-Sikar, K., & Lorencatto, F. (2020). Focusing on fidelity: Narrative review and recommendations for improving intervention fidelity within trials of health behaviour change interventions. *Health Psychology and Behavioral Medicine, 8*(1), 132–151. https://doi.org/10.1080/21642850.2020.1738935

von Thiele Schwarz, U., Aarons, G., & Hasson, H. (2019). The Value Equation: Three complementary propositions for reconciling fidelity and adaptation in evidence-based practice implementation. *BMC Health Services Research, 19*(1), 868–910. https://doi.org/10.1186/s12913-019-4668-y

IV

DEVELOPING A PROGRAMME THEORY FOR THE SYSTEMIC PRACTICE MODEL IN CHILDREN'S SOCIAL CARE: KEY INFORMANTS' PERSPECTIVES

by

Aaltio, E. & Isokuortti, N. 2021

Child and Family Social Work, 27(3), 444-453.

https://doi.org/10.1111/cfs.12896

Reproduced with kind permission by Wiley.

Developing a programme theory for the Systemic Practice Model in children's social care: key informants' perspective

Elina Aaltio[1] and Nanne Isokuortti[2]

[1] Doctoral candidate, Department of Social Sciences and Philosophy (Social Work), University of Jyväskylä, Finland & Visiting researcher, Department of Children, Young People and Families, Finnish Institute for Health and Welfare. Email: elina.m.aaltio@jyu.fi.

[2] Doctoral candidate, Faculty of Social Sciences (Social Work), University of Helsinki, Finland. Email: nanne.isokuortti@helsinki.fi.

Correspondence concerning this article should be addressed to Elina Aaltio, University of Jyväskylä, Department of Social Sciences and Philosophy (Social Work), Keskussairaalantie 2, PO Box 35, 40014 Jyväskylä, Finland. Email: elina.m.aaltio@jyu.fi

Acknowledgments

We would like to thank Johanna Kiili, Mikko Mäntysaari, Donald Forrester and Mirja Satka for their supportive and helpful comments on the article draft. We also thank Michael Freeman for help with language editing. We are grateful to the National Institute of Health and Welfare for their support and funding of the research project and to all the participants who contributed to the project.

Developing a programme theory for the Systemic Practice Model in children's social care: key informants' perspectives

Abstract

Vague programme descriptions are known to impede implementation and evaluation. Yet social work change programmes often fail to describe in detail how they aim to provide better outcomes for children and families. This study focused on the Systemic Practice Model (SPM), which is a Finnish adaptation of Reclaiming Social Work, a practice model developed in England. The SPM aims to deliver systemic practice in children's social care and has recently been widely disseminated across Finland. However, research has found both considerable variation in its delivery and a lack of clarity about what it is. This study, applying realist evaluation, aimed to formulate a programme theory based on the perceptions of key informants (n=12) involved in the national development and dissemination of the SPM during the period 2017–2019. The analysis yielded three core components (a systemic team, systemic weekly meetings, and systemic practice) and two context-mechanism-outcome (CMO) configurations, which represent causal chains of the SPM. The findings should enable its quality implementation and meaningful outcome and process evaluation. The findings will also aid organisations and practitioners both to identify changes required in their service provision and prevent adverse effects. This programme theory should be tested and refined with empirical data.

Key words: child and family social work, child protection, practice model, programme theory, realist evaluation, systemic social work

Introduction

Attempts to improve the quality and effectiveness of child and family social work have motivated decision-makers and practitioners to search for novel approaches to guide their service delivery. Along with distinct evidence-based practices, service organisations have increasingly adopted practice models (Baginsky et al., 2020; Gillingham, 2018). These models have a clear theory and value base, and they define the practices and skills that social work practitioners should use at all stages of their casework in order to optimise the safety and well-being of children receiving these services (Barbee et al., 2011). In Finland, a government-led initiative funded the nation-wide adaptation and dissemination of one such practice model, the Systemic Practice Model (SPM), based on the Reclaiming Social Work (RSW) model, during 2017–2019. The RSW introduces the ideas and methods of systemic family therapy in the context of statutory child and family social work (Goodman and Trowler, 2012).

Evidence on the utility of practice models is limited (Isokuortti et al., 2020), hence more high-quality studies are needed to evaluate whether these models improve outcomes for children and families. One of the challenges in measuring the effectiveness of social work interventions has been the lack of a logic model or a programme theory that would stipulate an intervention's change mechanisms and guide evaluation and interpretation (Pecora et al., 2006). The lack of a detailed intervention description may also impair the processes of implementation and evaluation (Hasson, 2010). If intervention users are unable to put the intervention into practice in a meaningful way, it is only possible to evaluate its implementation and not its outcomes. Although the need for programme theories has been recognized in previous research (Crampton, 2006; Gillingham, 2018), only a few published examples of these theories exist in the context of children's social services (see Turnell and Murphy, 2017; Sheehan et al., 2018).

In Finland, the SPM was developed and disseminated in a creative and collaborative manner. No manual or uniform training materials were offered on the local level. Instead, the content of the SPM was developed during the Training of Trainers (ToT) and by local teams implementing the SPM. In the nationally coordinated ToT, a group of trainees were taught SPM related skills and knowledge to be passed on to the implementation teams. While the practitioners positively engaged with the systemic approach, in particular the involvement of a clinician in the weekly team meetings, to many it remained unclear precisely how the SPM differed from service as usual (Isokuortti and Aaltio, 2020). This study sought to clarify these issues in collaboration with the key informants involved in the national development and dissemination of the SPM during 2017–2019. Based on their perceptions, this study aimed to formulate a programme theory for the SPM. A programme theory comprises the set of assumptions established by a programme's designers about how and why they expect the programme to attain its outcomes and in what context (Marchal et al., 2018). The core components of the model are its most essential and indispensable components (Fixsen et al., 2005). As the study applied a realist evaluation (Pawson and Tilley, 1997), we were especially interested in the model's mechanisms, i.e., the resources it provides and how these should work to change human reasoning and behaviour. Since the model seeks to bring about change on multiple levels, we explored the causal chains that connect the changes on the first level to the outcomes on the next (Shaw et al., 2018). The resulting theory may serve as a starting point for further outcome and process evaluations of the SPM. Our research questions were:

1) What, according to key informants, is the programme theory of the SPM?
2) More specifically, what are the core components of the SPM and how do these differ from service as usual?
3) What are the mechanisms and causal chains of the SPM that provide more effective service compared to service as usual?

The Finnish context for children's social care and Reclaiming Social Work

A comparative study of child protection systems (Gilbert et al., 2011) categorised those of the Anglo-American countries as risk-oriented 'child protection orientation' and those of the Nordic and continental countries as more therapeutic 'family service orientation', while acknowledging that both categories had elements in common and an emerging child-focused orientation. The system in England, from where the RSW has been imported, is a hybrid that supports vulnerable families but is also oriented towards child protection, whereas the Finnish system is primarily a family-service oriented system which has incorporated some elements of a child-focused orientation. In general, the English system involves more national regulation, whereas the Finnish system is a deregulated system that allows more professional discretion in decision making (Berrick et al., 2015). However, compared to their English counterparts, Finnish social workers' perceived work pressure is higher, and they have over double the caseload (Berrick et al., 2016). In Finland, moral distress, i.e., work-related malaise due to inability to practice in a morally appropriate way, has impaired social workers' well-being and motivation to remain in the sector (Mänttäri-van der Kuip, 2016).

Following the high-profile death of an 8-year-old girl under a child protection plan in 2012, the Finnish government commissioned a report on the state of Finland's child protection services (Kananoja et al., 2013). The problems identified included high staff turnover, poor management, high caseloads, incomprehensive understanding of families' needs and lack of service-user participation. The Finnish government responded with a programme to reform child and family services that was to be implemented during 2016–2019. Key stakeholders in the central government perceived RSW as a promising solution in reforming children's social care. Consequently, the SPM, an adaptation of RSW, had initially been implemented in 31 municipal children's services sites across Finland by the summer of 2018. The present

researchers had no part in the implementation decision. The model selection process will be analysed in a forthcoming study.

The RSW model has not been manualised, thereby allowing room for free adjustment. The original model was developed in Hackney Children's Services by Trowler and Goodman (2012), whose idea was to create a model that would promote good practice and overcome problems within the child protection system, such as skills deficits, a risk-averse mentality and reliance on procedural approaches at the expense of creative and independent thinking. In the original model (Goodman and Trowler, 2012), professionals work in small multi-disciplinary systemic units. Each unit has a qualified social worker, a child practitioner, a unit coordinator, and a systemic family therapist as a clinician, and is led by a consultant social worker, that is, a senior practitioner who has more competence in systemic practice. The units hold weekly meetings which serve as case supervision sessions. In other UK agencies, these units have been replaced by larger systemic teams consisting of more than one case-holding social worker and occasionally without the clinician (Bostock et al., 2017). According to Bostock et al. (2019a), the following features characterise a systemic case discussion: 1) patterns of family relationships and narratives are actively explored; 2) family relationships are set in the wider social context; 3) a child and family focus is present during the conversation; 4) there is clarity about the potential risks to the child(ren); 5) discussion is curious and reflective; 6) different hypotheses are generated and/or evidence presented that challenge established theories about the family; 7) hypotheses are developed into clear and actionable conversations with the families; and 8) discussion is collaborative and involves all group members.

The most important similarities between RSW and the SPM include practitioner training in the systemic approach, i.e., seeing the family as a system and carrying out interventions aimed at changing the system's functioning (Forrester et al., 2013), the inclusion of a clinician in the team, and the use of systemic tools and methods with families and in weekly meetings. During

its initial implementation, new elements were incorporated into the SPM, such as inviting the family to participate in the team meeting. At the same time, some important elements were excluded, such as the on-site coaching of managers in how to create systemic change in the organisation. In Finland, for legislative reasons, social workers hold case responsibility instead of the consultant social worker.

Previous evaluations suggest that the RSW model has improved the quality of local child protection practice (Bostock et al., 2017, 2019b; Forrester et al., 2013). Forrester et al. (2013) outlined six features that distinguished the original Hackney units from conventional teams: shared work, in-depth case discussion, a shared systemic approach, skills development, special roles, and low caseloads. In addition, the authors outlined a model describing how the core elements link to the target outcomes. The evaluations conducted in Hackney (Cross et al., 2010; Forrester et al., 2013) and in new sites (Bostock et al., 2017, Bostock and Newlands, 2020, Laird et al. 2017, 2018; Morris et al., 2017) have yielded important knowledge on how systemic practice works and which contextual factors are essential for implementing the model successfully. Nevertheless, to date, no evidence has been presented on the effectiveness of the RSW model. Additionally, the follow-up study (Bostock and Newlands 2020) has identified further changes in the delivery of the model in the UK, most notably reductions in training, clinician input and group supervision and the implementation of a second approach, i.e., Signs of Safety (SoS), to guide practice. According to Baginsky et al. (2020), eight local authorities in the UK report using both SoS and a systemic approach while 134 use SoS only. To our best knowledge, the precise number of sites using RSW alone or some other systemic approach has not been published.

Our evaluation of the implementation of the SPM in Finland (Isokuortti and Aaltio, 2020) revealed wide variation in fidelity to the model across 23 implementation sites and the teams within them. Several practitioners reported that a lack of clarity concerning systemic practice,

insufficient training, staff turnover, and a lack of leader commitment to change impeded implementation of the model. High caseloads had a negative impact on implementation, as this meant, among other things, that social workers had insufficient time for family meetings and learning new techniques. We also found that coaching and positive experiences of the model facilitated implementation. Our initial evaluation (Aaltio and Isokuortti, 2019) concluded that all stakeholders would benefit from a clearer description of the core components of the SPM and of the mechanisms that were expected to improve the quality and outcomes of child and family social work, i.e., a programme theory. Thus, we (the researchers) suggested organising a workshop process for the key individuals in the central government agency responsible for the model's development and dissemination to refine the description of the components and create a programme theory for the SPM in a collective process.

Data and methods

Methodological framework

This study applied a realist evaluation framework (Pawson & Tilley, 1997), which is a form of theory-driven evaluation research based on a realist philosophy of science. A realist programme theory "will explain how and why different outcomes are generated in different contexts" (Greenhalg et al., 2017a). Programmes seek to change people's behaviour by changing their *reasoning* through the offer of *resources* (Pawson, 2013) that may, for example, be material, emotional or social (Greenhalg et al., 2017b). The combination of resources and reasoning forms a *mechanism* (M), which generates the programme's *outcomes* (O). However, these mechanisms will only fire in the right *context* (C). Here, context refers not only to institutional settings or the wider social, economic, and cultural infrastructure, but also to the individual characteristics and capacities of stakeholders and their interpersonal relations (Pawson, 2013).

To explain why a programme has or has not worked, a realist evaluation aims to analyse the relationships between the mechanism and the context with context-mechanism-outcome (CMO) configurations. Dalkin et al. (2015) highlight the temporal order of the process. Resources are introduced in a context and together these two affect the reasoning and behaviour of the programme participants. Ultimately, the changes in reasoning and behaviour lead to specific outcomes.

Furthermore, programmes often involve long causal chains (Shaw et al., 2018). This means that the programme may aim to change the reasoning and behaviour of multiple stakeholders on different levels such as leaders, practitioners, and service users. A mechanism on the first level may become the context of the next mechanism in this chain. Whether something is categorized as a mechanism or a context depends on the level of analysis in focus at any given time.

Finally, social programmes are generally expressed as a set of ideas about how to change something for the better. The evaluator's task is to articulate these as theories and refine them in a continuous process (Pawson, 2013). Greenhalg et al. (2017a) posit that an overall programme theory comprises several subtheories that can be tested in future evaluations. The findings are then used to further refine the programme theory. Hence, the programme theory is both the starting point and the unit of analysis (Pawson, 2013).

Participants, procedure and analytical framework

Our programme theory for the SPM was created in collaboration with a group of key informants (n=12) in five workshops between January and August 2019. In addition, the researchers and most of the workshop participants participated in a 5-day systemic training course between the workshops run by two British RSW experts. The training provided additional information about the original model and its functions and was used as a reference point in the workshop

discussions. We also conducted two interviews, the first with the two RSW experts and the second with one Finnish workshop participant. The workshops lasted between 140 and 200 minutes and the interviews between 85 and 110 minutes. The first three workshops and the interviews were audio-recorded (12.25 hours in total) and transcribed verbatim. As supplementary data, we drew on the research literature (e.g., systemic family therapy literature, RSW evaluations), notes from the last two workshops, including written comments received from the workshop participants, notes made during the training course, and personal communications (notes from two meetings and emails) with four informants: three workshop participants in charge of SPM dissemination and one key informant outside the workshop process.

The semi-structured interview protocol (SSIP) proposed by Gugiu and Rodriguez (2007) for constructing logic models guided our data collection and analysis. A logic model is a visual map or a narrative description of the programme's components and desired results. The SSIP protocol aims to aid evaluators and programme managers to create a description of the programme and to build a programme theory defining the connections among the logic model elements. The seven phases of the SSIP are: (1) identify key informants and basic background and contextual information; (2) generate logic model elements, i.e., outcomes (e.g., changes in behaviour or in the organisation), activities (e.g., specific actions and processes used to produce outputs and outcomes), outputs (e.g., services, techniques) and inputs (e.g., resources invested and used by the programme); (3) model these elements with key informants; (4) build a rational theory; (5) develop a programme theory; (6) prioritise logic model elements; and (7) build a graphical or tabular logic model. We adapted this protocol to a realist evaluation by first outlining a researcher's theory (Pawson and Tilley 1997, p. 159-161), which was then refined in the workshops. In addition, we elaborated the core components and logic model elements from the realist perspective by formulating hypotheses on the mechanisms of the SPM in

different contexts. Finally, instead of building a logic model, we formulated CMO configurations.

Table 1 provides a summary of the collaborative process for generating the SPM programme theory. In phase one, we used purposeful sampling to identify the workshop participants and the interviewees. The workshop participants (n=12) were key individuals in the development of the model, the training programme and national dissemination. Of these twelve, eight were practitioners in children's social care and six worked in an SPM team. The interviewees were specialised in RSW, systemic family therapy, and supervision. The researchers jointly identified the objectives and content for each workshop, designed the interview protocols and mutually conducted these sessions. An informed consent was requested from all research participants prior to the data collection.

[Table 1 here]

The first three workshops were used to refine the logic model elements and explore mechanisms through which these generate the programme outcomes in a given context (phases two and three). Our researcher's theory of the core components, key concepts and logic model elements was used as a starting point for the discussion. We had based this proposed theory on our initial findings on the implementation of the SPM (Aaltio and Isokuortti, 2019) and other evaluators' previous findings on the RSW. We asked the workshop participants to discuss and refine these components, and then formulate hypotheses on the causal relationships between the SPM components in different contexts.

To build a rational theory (phase four), the researchers negotiated the proposed programme theory elements with the participants in the workshops. After the first three workshops, we produced a first draft of the programme theory based on the transcripts and notes from these workshops and two key informant interviews (phase five). The transcripts were analysed with

ATLAS.ti. Data were first coded by using key components and other key concepts relating to the SPM (e.g., genograms, hypothesising) as coding labels. During the data analysis, we first produced a detailed description of each core component based on the data. Next, we identified contradictions and alternative ways of understanding the components and their functions and highlighted these in the draft. The supplementary data were used to add details and references to theories and methods. The draft was introduced to the research participants for collective discussion and refinement and to prioritise its elements (phase six). Finally, the coded data were further categorised into the following themes: context, mechanisms, and outcomes. We then constructed hypotheses on the causal chains generated by the SPM. Narrative and tabular versions of the two CMO configurations were created for this paper (phase seven).

Results

The analysis resulted in three core components, the ideal context for these components, and two CMO configurations representing the causal chains of the SPM.

Core component 1: Systemic team

The workshop participants explained that a systemic team is formed from a permanent group of social work practitioners with a shared approach and orientation to systemic practice. An ideal systemic team was seen as comprising a consultant social worker, a systemic family therapist (a clinician), a coordinator and two to three social workers. Additionally, a team could include one or more family practitioners. According to the participants, the team shares "the responsibility for interpreting the family situation and making decisions regarding the steps to be taken." The minimum number of social workers in a team was considered to be two so that they can fill in for and provide each other with peer support. The maximum number was three in order to ensure time and space for systemic team discussion. The consultant social worker should be a senior practitioner who functions as a "mentor," "leads the team," and supervises the social workers in the weekly meetings. Alongside the clinician, the consultant social worker

ensures that the team maintains a systemic orientation. The participants identified the clinician's task as presenting new insights and generating reflexive curiosity while supporting the practitioners in adopting systemic thinking, methods, and tools. Although the clinician was expected to bring "a new perspective from the outside," "the clinician's position is not that of an outsider consultant, but a team member." The coordinator, in turn, provides administrative support for the social workers and contacts the families

Defining a systemic team was not a straightforward process. During the workshops, the participants had divergent views on the composition of a team. For some, a systemic team could mean a "network," i.e., a group of professionals from different organisations brought together to support a family systemically while others favoured the original idea of having permanent teams that would serve as "a support structure" for frontline social workers. The logic behind the first option was that in smaller organisations it had been possible to form ad hoc systemic teams to support one social worker and each case. Since such tailoring could not work in larger organisations, the systemic team would have to comprise permanent members only. However, the participants supported the idea of ad hoc teams in the form of a multiagency practice that was being developed in a parallel process.

In comparison to service as usual, a systemic team would include a family therapist as a permanent team member. Furthermore, a systemic team would engage more in mutual reflection from multiple perspectives. Practitioners would also share responsibility and knowledge of family cases, which would make for a "consistently high quality" of work.

Core component 2: Systemic weekly meetings

The workshop participants concluded that systemic case discussions should take place in systemic weekly meetings, and all team members should participate actively in the discussion. By reflecting on a case, the team can help the case-holding social worker to generate hypotheses

and plan the next steps. As one participant formulated: “Practitioners will gain multiple perspectives and find multiple directions in which to proceed.” The workshop participants agreed that the nature of these discussions should be “curious” and “respectful” towards the other team members and client families, “as if the client was present.” Tools and methods specific to systemic practice (see core component 3) would be used to support the systemic nature of the discussion. Team meetings should help the practitioners to “slow down” and ponder “what this case is about” instead of just thinking about what to do. To classify the different purposes of case discussions, the participants referred to the three Domains of Action proposed by Lang et al. (1990). In the ‘domain of aesthetics’, practitioners consider the ethical aspects of their work. In the ‘domain of explanation’, they explore various ideas and perspectives with curiosity. In the ‘domain of production’, they decide how to act based upon the case discussion. With the help of the team, practitioners should also reflect on their role as part of the system, as their behaviour could, in some cases, maintain the problems of the family: “So you might consider how you could change your own behaviour to promote change.”

During the workshop process, the participants’ views on the two meeting-related themes differed. First, some practitioners felt that the meeting should focus on reflecting one or two specific cases, whereas others felt that the meeting should serve as the main forum for case supervision and include discussion of most of the current cases. The participants eventually decided that each social worker should have an opportunity to introduce one or two cases for reflection by their team each week.

Second, the participants suggested three functions for the weekly team meetings: as a forum for 1) collegial reflection and learning, 2) practice, and 3) multiagency practice. With respect to the first function, the meetings would serve as a space for reflection and learning exclusive to the team members. With the help of the clinician and the consultant social worker, the meeting would form a safe place for social workers to reflect on possible prejudices and

emotions towards family, both of which might unconsciously affect their decision-making. With respect to the second function, the families would be invited to the weekly meetings. In this way, the expertise and multiple perspectives of the team would be at the family's disposal. With respect to the third function, both the families and all relevant professionals working with a family would be invited to the meeting. Some participants argued that practitioners "need an opportunity to reflect on their practice and how to process their emotions with the service user later on without the presence of the family." Others maintained that in principle all the discussions between practitioners should be open to families, whether present or not: "[…] you first need to clear the air between you and your client so that you can achieve change." After careful consideration, the participants decided that the weekly meeting should be exclusive to the systemic team members. The reason for this decision was more practical than ideological: having families and other professionals present would take up more time than if cases were discussed by the practitioners alone.

In contrast to service as usual, where case supervision is typically conducted once a month and purchased from the private sector, the systemic team meeting should serve as a case supervision session to be guided by the consultant social worker and the clinician.

Core component 3: Systemic practice

The participants defined systemic practice as purposeful, relationship-based, and systemic-oriented work with families, in which the focus is on the relationships between child and family, and the physical and immaterial living-environment of both, including the child protection system and practitioners. The key methods and techniques of systemic practice would include the use of genograms along with formulating hypotheses, as proposed by Cecchin (1987), and circular and other types of questions, as proposed by Tomm (1988). The participants welcomed additional tools, such as a timeline, but the use of these tools was not deliberated in the workshops. According to the participants, the value of the proposed key

methods is that they provide insights and new perspectives on family dynamics, history, patterns, and narratives: “Both the service user and the practitioner […] can see the service user’s situation and experiences from a new angle.” Such an approach aims to increase practitioners’ understanding and family members’ compassion, self-understanding, and empowerment. One participant described it as follows: “It can also decrease parent’s guilt to see the whole picture and realise that it’s not just me, a bad parent, but it’s me as part of a family system.” In addition, this approach can improve “parents’ ability to mentalise,” i.e., to understand their child’s mental state. The overall purpose is to strengthen the interaction between family members and help them to generate solutions to their own problems. Moreover, the practitioner needs to have “trust in the family’s strengths.” However, to successfully broach difficult and sensitive issues, the social worker first needs to create a trusting and affirmative relationship with the child and parents.

Given that the SPM aims at child-centred practice, the participants emphasized the importance of direct contact with the child. However, to create this contact and ensure the child’s safety and well-being, the social worker needs to build trust with the parents “since the children are bound to the adults and the family.” The participants concluded that practitioners need to be flexible when considering the order in which they approach family members in a given case. Participants referred to the idea of ‘seeking permission’ (Aggett et al., 2015), i.e., requesting clients’ permission to proceed at all steps in the process of engagement as a way of forging a good relationship. Similarly, practitioners should be aware of the limits of their knowledge (Anderson, 1997) and be respectful and curious (Cecchin, 1987). Thus, the aim would be to maximise the sharing of power between the social worker and the family while keeping children safe, “[moving] away from the idea of the social worker knowing it better.” To benefit from the full potential of systemic methods and techniques, practitioners should meet their families frequently.

In service as usual, social workers often focus more on other tasks (e.g., managing services, administration, court preparations) than direct practice. The participants stated that in systemic practice the social worker forms a relationship with the children and parents instead of delegating this task to other practitioners. Relationship-based and curious practice were seen as the key elements of systemic practice in contrast to the more managerial and instructive approach to practice with families.

An ideal context for implementing the core components

The participants acknowledged that in contrast to the RSW programme, which aims to reform all levels of the organisation from management to face-to-face practice, the focus of the SPM, and hence its training, had been on the team level. Although a detailed analysis of how the SPM would promote systemic change on the organisational level was beyond the scope of this study, the participants were asked to outline an ideal context for a systemic team and practice. According to the participants, leaders and managers need to analyse "the whole chain" of services "systemically […] beyond organisational boundaries," and provide sufficient resources for frontline practice. To do this, the managers need "coaching" and "peer support." Managers should also participate in practitioners' training, so that they can better understand what systemic practice requires: "[T]hey will then realise that this takes time. This is what I want my employees to do. Hang on, we want more resources.") All in all, there should be "a shared understanding" between the organisation's leaders, managers and practitioners.

Causal chains of the Systemic Practice Model

Based on the researchers' synthesis of all the results presented above, we formulated two CMO configurations illustrating how use of the resources provided by the SPM can, in the right context, result in reasoning that generates the target outcomes (see Tables 2 and 3). Since the SPM entails long causal chains (Shaw et al., 2018), the process is disaggregated into two phases in which the intermediate outcomes of the previous phase, in particular, serve as a context or

mechanism for the next phase. We first outline the interrelationships of the context, mechanisms and outcomes on the level of a systemic team, and then move on to the level of the family.

The ideal context for a systemic team is the presence of a systemic organisational culture and management that support the team's frontline practitioners and daily routines (Table 2). An ideal team provides a safe learning environment for practitioners. By providing systemic training and coaching, the SPM equips practitioners with new knowledge and understanding not only of family dynamics and the factors underlying child abuse and neglect but also of family functioning. The social workers in the team learn new communication skills and strategies that can help in addressing sensitive questions with family members. Second, team discussions lead to shared responsibility, which in turn helps social workers make better decisions for families. Systemic training and case supervision change the reasoning of practitioners, improve their skills, and change service-as-usual practice into systemic practice. The new skills learnt include permission-seeking and building a relationship with families that is both respectful and curious. When social workers feel they have time and permission to use their skills and expertise, share responsibility and receive collegial support from the team in dealing with difficult cases, their work-related wellbeing improves.

[Table 2 here]

The intermediate outcomes on the practitioner level now form the context and mechanisms for the family level (Table 3). The systemic practice and practitioners serve as a resource for family members seeking to change their behaviour or situation. The ultimate aim in children's social care is to reach a point where it is safe to close the case. With the help of a systemically trained social worker, family members can more easily identify problematic interaction and communication. Thus, the practitioner's aim is to improve the family's communication and

dynamics. A skilful social worker is able to evoke the motivation to change in family members themselves instead of telling them what to do. This can provide family members with new insights, in turn increasing the likelihood of longer-lasting behaviour change. The final outcomes of systemic practice include the increased safety and wellbeing of parent/s and child/ren, both of which have the potential to last longer than in service as usual.

[Table 3 here]

4 Discussion

The purpose of this study was to formulate a programme theory for the SPM including a coherent description of its core components and the mechanisms that create the intended outcomes in a given context. Building a programme theory of this kind with stakeholders is crucial for several reasons. First, we argue that a shared understanding of the key concepts and their concrete content is a prerequisite for further outcome or process evaluations. Indeed, a clear theory helps researchers to identify whether the expected change has occurred. Second, an ideal description, or a *normative theory* (Chen 1990, p. 43*)*, aids organisations and practitioners to identify the changes needed in their service provision. Third, as pointed out by Urk et al. (2016), theory-based interventions may be more effective than non-theory-based interventions since they may result in a better understanding of human behaviour. Fourth, as also noted by Urk et al. (2016), the underlying theory can reveal weak links and potential conflicts in the hypothesised change process. Fifth, given that programmes may also have unintended and even harmful outcomes, programme theories can prevent potential adverse effects (Bonell et al., 2015).

Based on our analysis of the key informants' perspectives along with supplementary data, we were able to 1) define the composition of systemic team, 2) clarify the main function for weekly meetings, and 3) produce a coherent description of the methods and objectives of systemic

practice in the context of child and family social work. Furthermore, we identified the expected differences between the SPM and service as usual. Finally, our analysis of two CMO configurations illustrates how the core components of the SPM would produce the intended outcome in an ideal context.

However, creating and applying a programme theory for a new, non-manualised programme is a challenging task. Despite the collaborative nature of the workshop process, the consensus reached between the participants may only be temporary. The interests of those developing the services and those evaluating them might not converge. For instance, disseminators and implementers might feel the need to combine aspects of various practice models while evaluators might prefer a stepwise implementation process controlling for only a few variables (Baginsky et al., 2020). If the model is being developed in an on-going process of co-production, it might be almost impossible to keep up with the inputs and intended outputs let alone succeed in the data collection.

In addition, the mechanisms of the SPM only fire in the right context. However, creating an ideal context is demanding, and the model itself is complex (see also Isokuortti and Aaltio, 2020). Thus, we need to recognise that the programme described in this paper may not ultimately be the one that is applied in practice. Nevertheless, it is essential to have a commonly accepted description of the core components and their relation to the context and outcomes against which subsequently amended versions can be compared.

Based on our findings, we also propose that when decision-makers select interventions for dissemination, they should pay more attention to the intervention's formulation and justification. If an intervention lacks a coherent and concrete programme theory, this should be formulated before any large-scale dissemination. This helps to prevent adverse outcomes and guide the practitioners delivering the intervention.

Although the original RSW model was not implemented as such in the Finnish context, it guided the present initiative. For instance, the key features of the original RSW model identified by Forrester et al. (2013) are in line with the core components of the SPM. Furthermore, previous research and learning experiences aided this project. For example, Bostock et al. (2017) highlight the decisive roles of the consultant social worker and the coordinator in ensuring good practice within the team. In Finland, in contrast, creating change relied heavily on the clinicians and the trainers during the initial implementation of the SPM (Authors' own, 2019). During the workshop, we were able to reconsider the role of each team member. This and other findings can be used to improve the focus of the SPM training modules, and the implementation strategy and everyday delivery of the SPM in systemic teams.

Limitations and future directions

This study has its limitations. First, as its purpose was to form a national consensus on the SPM in collaboration with the central government agency, participation was limited to key individuals responsible for the development and dissemination of the SPM. Nevertheless, we consider frontline practitioners' views vital and have previously explored these in depth in a parallel study (Isokuortti and Aaltio, 2020). We discovered that many social workers found the application of systemic practice in real-world settings confusing. Consequently, its implementation resulted in only limited changes, as also confirmed by service users. Hence, we concluded that the involvement of families and frontline practitioners at this point would have been premature. However, the input of frontline practitioners and families from agencies that have moved closer to full implementation will be essential in refining this programme theory in future research.

Second, the study focused on the team and family levels. Although the participants recognised the need for organisational change, its operationalisation was just beginning in a separate process that did not involve us. Owing to the different time frames and nature of these

processes, we decided not to change the scope of our study. Nevertheless, both we the researchers and our participants acknowledged the need to expand the programme theory to include the organisational level in the future.

Third, the study relied primarily on workshop data. However, grounding the study in a realist methodology enabled in-depth reflections on the model's context, mechanisms, and outcomes. Furthermore, we found that the protocol proposed by Gugiu and Rodriguez (2007) had added value in the process. Defining the programme's inputs, activities and outputs helped us to describe the core components of the SPM before discussing how and why these might provide better outcomes. Fourth, given that the lead author was responsible for coding the data, we aimed to diminish potential bias by continuous reflection as well as mutually collecting and interpreting the data.

Fifth, our findings represent an ideal description of the SPM. In real-world settings, the hypotheses outlined here might not be supported or, although they may seem reasonable, the programme does not have sufficient leverage to bring about change. Future evaluations should focus on testing these hypotheses in real-life conditions.

Finally, based on this data and practitioners' experiences (Isokuortti and Aaltio, 2020), the SPM appears to have the potential, if fully implemented, to improve practice by furnishing frontline practitioners with new skills and knowledge and enabling more case supervision.

Conclusions

A detailed description of an intervention is a prerequisite for meaningful evaluation. In a realist evaluation, both the researchers and the participants play an active role in producing this description, i.e., the programme theory. The findings of this study are intended to guide the testing and refinement of the theory in the future as well as assist other researchers and

developers in undertaking similar collaborative efforts. In addition, a programme theory is equally important for quality implementation.

References

Aaltio, E. & Isokuortti, N. (2019). Systeemisen lastensuojelun toimintamallin pilotointi. Valtakunnallinen arviointi [The initial implementation of the Systemic Practice Model for child protection. National evaluation]. Report 3/2019. Helsinki: National Institute for Health and Welfare.

Aggett, P., Swainson, M., & Tapsell, D. (2015). 'Seeking permission': an interviewing stance for finding connection with hard to reach families. *Journal of Family Therapy*, 37(2), 190–209. https://doi.org/10.1111/j.1467-6427.2011.00558.x

Anderson, H. (1997). *Conversation, Language, and Possibilities*. New York: Basic Books.

Baginsky, M., Ixer, G., & Manthorpe, J. (2020). Practice frameworks in children's services in England: An attempt to steer social work back on course? *Practice*, 33(1), 3–19. https://doi.org/10.1080/09503153.2019.1709634

Barbee, A. P., Christensen, D., Antle, B., Wandersman, A., & Cahn, K. (2011). Successful adoption and implementation of a comprehensive casework practice model in a public child welfare agency: Application of the Getting to Outcomes (GTO) model. *Children and Youth Services Review*, 33(5), 622–633. https://doi.org/10.1016/j.childyouth.2010.11.008

Berrick, J. D., Dickens, J., Pösö, T., & Skivenes, M. (2015). Children's involvement in care order decision-making: A cross-country analysis. *Child Abuse & Neglect*, 49, 128–141. https://doi.org/10.1016/j.chiabu.2015.07.001

Berrick, J. D., Dickens, J., Pösö, T. & Skivenes, M. (2016). Time, Institutional Support, and Quality of Decision Making in Child Protection: A Cross-Country Analysis. *Human Service Organizations: Management, Leadership & Governance*, 40(5), 451–468. https://doi.org/10.1080/23303131.2016.1159637

Bonell, C., Jamal, F., Melendez-Torres, G.J., & Cummins, S. (2015). 'Dark logic': Theorizing the harmful consequences of public health interventions. *Epidemiology & Community Health,* 69(1), 95–98. http://dx.doi.org/10.1136/jech-2014-204671

Bostock, L., Forrester, D., Patrizo, L., Godfrey, T., Zonouzi, M., Antonopoulou, V., Bird, H., & Tinarwo, M. (2017). Scaling and deepening the Reclaiming Social Work model. Evaluation Report. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/625227/Scaling_and_deepening_the_Reclaiming_Social_Work_model.pdf

Bostock, L., Patrizo, L., Godfrey, T., Munro, E., & Forrester, D. (2019a). How do we assess the quality of group supervision? Developing a coding framework. *Children and Youth Services Review,* 100, 515–524. https://doi.org/10.1016/j.childyouth.2019.03.027

Bostock, L., Patrizo, L., Godfrey, T., & Forrester, D. (2019b) What is the impact of supervision on direct practice with families? *Children and Youth Services Review,* 105, Article 104428. https://doi.org/10.1016/j.childyouth.2019.104428

Bostock, L., & Newlands, F. (2020). Scaling and deepening the Reclaiming Social Work model: longitudinal follow up. Evaluation report. Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/933120/Reclaiming_Social_Work_-_Bedfordshire.pdf

Gugiu, P. C., & Rodriguez-Campos, L. (2007). Semi-structured interview protocol for constructing logic models. *Evaluation and Program Planning,* 30(4), 339–350. https://doi.org/10.1016/j.evalprogplan.2007.08.004

Cecchin, G. (1987). Hypothesizing, circularity, and neutrality revisited: An invitation to curiosity. *Family Process* 26(4), 405–413. https://doi.org/10.1111/j.1545-5300.1987.00405.x

Chen, H.-T. (1990). *Theory-driven evaluations*. Newbury Park: Sage Publications.

Crampton, D. (2007). Research Review: Family group decision-making: a promising practice in need of more programme theory and research. *Child and Family Social Work*, 12(2), 202–209. https://doi.org/10.1111/j.1365-2206.2006.00442.x

Cross, S., Hubbard, A., & Munro, E. (2010). *Reclaiming Social Work. London Borough of Hackney. Children and Young People's Services. Part 1: Independent Evaluation*. https://core.ac.uk/download/pdf/30322069.pdf

Dalkin, S. M., Greenhalg, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, 10, Article 49 (2015). https://doi.org/10.1186/s13012-015-0237-x

Fixsen, D., Naoom, S., Blase, K., Friedman, R., & Wallace, F. (2005). Implementation Research: A Synthesis of the Literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Forrester, D., Westlake, D., McCann, M., Thurnham, A., Shefer, G., Glynn, G., & Killian, M. (2013). *Reclaiming Social Work? An Evaluation of Systemic Units as an Approach to Delivering Children's Services. Final report of a comparative study of practice and the factors shaping it in three local authorities*. Bedford: University of Bedfordshire.

Gilbert, N., Parton, N., & Skivenes, M. (Eds.). (2011). *Child protection systems: International trends and orientations.* New York: Oxford University Press.

Gillingham, P. (2018). Evaluation of Practice Frameworks for Social Work with Children and Families: Exploring the Challenges. *Journal of Public Child Welfare*, 12(2), 190–203. https://doi.org/10.1080/15548732.2017.1392391

Goodman, S., & Trowler, I. (Eds.). (2012). *Social Work Reclaimed. Innovative Frameworks for Child and Family Social Work Practice*. London and Philadelphia: Jessica Kingley Publishers.

Greenhalg, T., Greenhalg, J., Pawson, R., Manzano, A., Wong, G., Jagosh, J. & Westhorp, G. (2017a). "Theory" in realist evaluation. https://www.ramesesproject.org/media/RAMESES_II_Theory_in_realist_evaluation.pdf

Greenhalg, T., Greenhalg, J., Pawson, R., Manzano, A., Wong, G., Jagosh, J. & Westhorp, G. (2017b). What is a mechanism? https://www.ramesesproject.org/media/RAMESES_II_What_is_a_mechanism.pdf

Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation Science,* 5, Article 67 (2010). https://doi.org/10.1186/1748-5908-5-67

Isokuortti, N. & Aaltio, E. (2020). Fidelity and influencing factors in the Systemic Practice Model of children's social care in Finland. *Children and Youth Services Review*, 119, Article 105647. https://doi.org/10.1016/j.childyouth.2020.105647

Isokuortti, N., Aaltio, E., Laajasalo, T. & Barlow, J. (2020). Effectiveness of child protection practice models: a systematic review. *Child Abuse & Neglect,* 108, Article 104632. https://doi.org/10.1016/j.chiabu.2020.104632

Kananoja, A., Lavikainen, M., & Oranen, M. (2013). *Toimiva lastensuojelu. Selvitysryhmän loppuraportti [Functioning Child Protection. Final report of the clearing group]*. Ministry of Social Affairs and Health publications 2013:19. Helsinki: Sosiaali- ja terveysministeriö [Ministry of Social Affairs and Health].

Laird, S., Morris, K., Archad, P., & Clawson, R. (2017). Working with the Whole Family. What case files tell us about Social Work Practices. *Child & Family Social Work*, 22(3), 1322–1329. https://doi.org/10.1111/cfs.12349

Laird, S., Morris, K., Archad, P., & Clawson, R. (2018). Changing practice: The possibilities and limits for reshaping social work practice. *Qualitative Social Work*, 17(4), 577–593. https://doi.org/10.1177/1473325016688371

Lang, P., Little, M., & Cronen, V. (1990). The systemic professional. Domains of action and the question of neutrality. *Human systems*, 1, 39–55.

Marchal, B., Kegels, G., & Van Belle, S. (2018). Theory and realist methods. In N, Emmel, J, Greenhalg, A, Manzano, M, Monaghan, & S, Dalkin (Eds.), *Doing realist research* (pp.79–90). London: SAGE Publications Ltd.

Morris, K., Archard, P. J., Laird, S., & Clawson, R. (2017). Family experiences of children's social care involvement following a social work change programme. *Journal of Social Work Practice*, 32(3), 237–250. https://doi.org/10.1080/02650533.2017.1326473

Mänttäri-van der Kuip, M. (2016). Moral distress among social workers: The role of insufficient resources. *International Journal of Social Welfare*, 25(1), 86–97. https://doi.org/10.1111/ijsw.12163

Pawson, R. (2013). *The science of evaluation. A Realist Manifesto*. London: Sage.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.

Pecora, P. J., McAuley, C., & Rose, W. (2006). Effectiveness of Child Welfare Interventions: Issues and Challenges. In C. McAuley, P. J. Pecora, & W. Rose (Eds.), *Enhancing Well-Being of Children and Families through Effective Interventions* (pp. 14–20). Philadelphia: Jessica Kingsley Publishers.

Shaw, J., Gray, S. C., Baker, G. R., Denis, J.-L., Breton, M., Gutberg, J., Embuldeniya, G., Carswell, P., Dunham, A., McKillop, A., Kenealy, T., Sheridan, N., & Wodchis, W. (2018). Mechanisms, contexts and points of contention: operationalizing realist-informed research for complex health interventions. *BMC Medical Research Methodology*, 18, Article 178 (2018). https://doi.org/10.1186/s12874-018-0641-4

Sheehan, L., O'Donnell, C., Brand, S. L., Forrester, D., Addis, S., El-Banna, A., Kemp, A., & Nurmatov, U. (2018). *Signs of Safety: Findings from a mixed-methods systematic review focussed on reducing the need for children to be in care.* London: What Works Centre for Children's Social Care. https://whatworks-csc.org.uk/wp-content/uploads/Signs_of_Safety_a_mixed_methods_systematic_review.pdf

Tomm, K. (1988). Interventive Interviewing: Part III. Intending to Ask Lineal, Circular, Strategic, or Reflexive Questions? *Family Process*, 27(1), 1–15. https://doi.org/10.1111/j.1545-5300.1988.00001.x

Trowler, I. & Goodman, S. (2012). Conclusion and Reflections on Our Journey. In S. Goodman, & I. Trowler (Eds.), *Social Work Reclaimed. Innovative Frameworks for Child and Family Social Work Practice* (pp.160–169). London and Philadelphia: Jessica Kingley Publishers.

Turnell, A. & Murphy, T. (2017) *Signs of Safety. Comprehensive Briefing Paper*. 4th edition. East Perth: Elia International Ltd. https://knowledgebank.signsofsafety.net/resources/introduction-to-signs-of-safety/signs-of-safety-comprehensive-briefing-paper/signs-of-safety-comprehensive-briefing-paper-en/signs-of-safety-comprehensive-briefing-paper?

Urk, F., Grant, S., & Bonell, C. (2016). Involving stakeholders in programme theory specification: discussion of a systematic, consensus-based approach. *Evidence & Policy,* 12 (4), 541–557. https://doi.org/10.1332/174426415X14474260456850

Table 1. Summary of the process for generating the SPM programme theory in accordance with the SSIP (Gugiu & Rodriquez, 2007).

Phase	Research and workshop activities	Data used in this phase
1. Identifying key informants and basic information	Purposeful sampling Formulating a researcher's theory	Previous findings on SPM and RSW evaluation
2. Generating logic model elements	Workshop 1: Introducing the researchers' theory, asking participants to refine definitions and explicate justifications for including or excluding components, inputs, activities, outputs and outcomes	
	Workshop 2: Discussing the researchers' proposition of key approaches, theories and methods underpinning the SPM. Defining the concrete resources (inputs, activities, outputs) offered by the SPM and the ideal context	Notes from interview 1, RSW training
3. Modelling logic model elements with key informants	Workshop 3: Creating hypotheses of the mechanisms and causal chains	
4. Building a rational theory	Workshops 1–3 and the first stakeholder meeting: Negotiating with the key informants	
5. Developing a programme theory	Analysing the transcripts Formulating the first draft of the programme theory Workshop 4: Discussing the draft	Transcripts of workshops 1–3, interview 2
6. Prioritising logic model elements	Revision of the draft Workshop 5, second stakeholder meeting, and personal communication: Refining the revised draft	Notes from workshop 4, personal communication
7. Formulating CMO configurations	Formulating tabular and narrative CMO configurations	All data

Table 2. The intended causal chains at the practitioner level.

Practitioner level			
Resources	**Context**	**Reasoning**	**Intermediate outcomes**
Systemic team + Systemic training and coaching + Systemic weekly meetings providing systemic case supervision	Organisation and leaders who implement daily practices and resources to support the proper functioning of the systemic teams, e.g., ensuring reasonable caseloads and recruiting the requisite team members + A team that forms a safe space for learning	Practitioners expand their understanding of family dynamics and begin to approach problems systemically from multiple perspectives + Mutual learning and reflection + Systemic team shares responsibility which provides emotional and practical support for practitioners	Intensive systemic practice, including respectful, power-sharing and curious relationships with families + Improved work-related wellbeing + Decreased staff turnover

Table 3. Intended causal chains at the family level.

Family level			
Resource	**Context**	**Reasoning**	**Intermediate and long-term outcomes**
Skilful, and permanent practitioners who experience wellbeing + Systemic social work practice	The family meeting context is respectful, power-sharing, and curious + Regular and frequent meetings with the same practitioner/s	Identifying multiple perspectives + Identifying problematic interaction and communication, e.g., family patterns or scripts + Family's own motivation for change aroused	New insights and change in beliefs + Improved family dynamics + Increased safety + Improved parent/ child wellbeing