Author(s): Isokuortti, Nanne; Aaltio, Elina; Laajasalo, Taina; Barlow, Jane

Title: Effectiveness of child protection practice models : a systematic review

Please cite the original version:

Effectiveness of child protection practice models: a systematic review

Nanne Isokuortti, University of Helsinki, Faculty of Social Sciences (Social Work), Unioninkatu 37, P.O. Box 54, 00014 University of Helsinki, Finland. E-mail address: nanne.isokuortti@helsinki.fi.

Elina Aaltio, Department of Social Sciences and Philosophy, University of Jyväskylä, Finland. Keskussairaalantie 2, P.O. Box 35, 40014 University of Jyväskylä, Finland. E-mail address: elina.m.aaltio@jyu.fi.

Taina Laajasalo, Department of Psychology and Logopedics, University of Helsinki, Finland. Address: Haartmaninkatu 8, P.O. Box 63, 00014 University of Helsinki, Finland. E-mail address: taina.laajasalo@helsinki.fi.

Jane Barlow, Department of Social Policy and Intervention, University of Oxford, UK, 32 Wellington Square Oxford, OX1 2ER, UK. E-mail address: jane.barlow@spi.ox.ac.uk.

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 address: nanne.isokuortti@helsinki.fi. Tel. +358407510954 (Nanne Isokuortti).

Abstract

Background: Attempts to improve child protection outcomes by implementing social work practice models embedded in a particular theory and practice approach, have increased internationally over the past decade.

Objective: To assess the evidence of the effectiveness of child protection practice models in improving outcomes for children and families.

Participants and setting: Children < 18 years and their families involved in child protection services.

Methods: A systematic review was conducted to synthesize evidence regarding the effectiveness of child protection practice models. Systematic searches across 10 electronic databases and grey literature were conducted to identify quasi-experimental studies minimally. Included studies were critically appraised and the findings summarized narratively.

Results: Five papers, representing six studies, focusing on three practice models (Solution-Based Casework; Signs of Safety; and Reclaiming Social Work) met the inclusion criteria. All studies applied a quasi-experimental design. Overall, the quality of the evidence was rated as being poor, with studies suffering from a risk of selection bias, small sample sizes and short-term follow up.

Conclusions: Despite the popularity of practice models, the evidence base for their effectiveness is still limited. The results suggest that high-quality studies are urgently needed to evaluate the impact of practice models in improving the outcomes of child-protection-involved families. The findings also illustrate the difficulties of conducting high-quality outcome evaluations in children's social care, and these challenges and future directions for research, are discussed.

1. Background

Every year 1.5-5% of children in the UK, USA, Australia, and Canada are reported to child protection agencies for all types of child maltreatment (Gilbert et al., 2009). Child protection services have a vital role in protecting children from serious harm. Nevertheless, many countries have encountered problems within child protection services such as the demands of bureaucracy reducing social workers' capacity to work directly with children and families, an increased workload and a high degree of work pressure (e.g., Berrick, Dickens, Pösö, & Skivenes, 2016; Holmes, & McDermid, 2013; Munro, 2011; STM, 2019). At its most severe, many countries have witnessed high-profile deaths of children involved in child protection services, some of which have drawn attention to the ability of children's social care to keep children safe (Holmes, & McDermid, 2013).

During the past two decades, child protection practice models (also known as practice frameworks) that are embedded in a particular theory and practice approach, have become increasingly popular in multiple countries, e.g., the United States, Australia, and the United Kingdom and other European countries (e.g., Baginsky, Moriarty, & Manthorpe, 2019; Gillingham, 2018; Laird, Morris, Archard, & Clawson, 2018). Barbee, Christensen, Antle, Wandersman, and Cahn (2011) 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 model 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 the child welfare system. (p. 623)

The overall aim of these models is to improve the quality of child protection services and

outcomes for children and families, by adopting a clear theoretical and practical approach to

social work practice (Gillingham, 2018). Existing reviews of child protection practice models

include three reviews of Signs of Safety (Baginsky, Moriarty, & Manthorpe, 2019; Bunn, 2013;

Sheehan et al., 2018) and an overview of the Solution-Based Casework model (Gillingham,

2018). Despite the growing body of research, there are no existing systematic reviews that focus

explicitly on assessing the effectiveness of all practice models in improving outcomes for

children involved in child protection services. Specifically, we are interested in assessing to

which extent the models provide intended effects in real-world settings. The aim of the current

review was therefore minimally to synthesise data from all quasi-experimental studies (i.e., pre-

post comparison group design studies) evaluating the effectiveness of child protection practice

models compared to regular child protection practice in improving outcomes for children and

families.

2. Method

2.1. Research question and eligibility criteria

The review protocol (CRD42018111918) was registered to PROSPERO (International

Prospective Register of Systematic Reviews). In contrast to the protocol, after preliminary

searches we added one database (i.e., Scopus) to our database search. In addition, we merged two

eligibility criteria (a description of practitioner skills and specified set of tools) into one

description of practitioner skills and/or tools.

The review question was "how effective are child protection practice models in improving outcomes for children aged 0-17 years and their parents involved in child protection services." For the purpose of this study, we defined a practice model as follows: the model had to be designed to improve child protection outcomes, and the model's aims and methods of achieving these should be clearly defined; it should involve all the following elements: i) a clear theoretical basis, ii) a framework for client practice, and iii) description of practitioner skills and/or tools. The model may also include a definition of values and reforms to workforce and structure. The models did not need to be licensed but did need to focus on statutory child protection casework practice provided by public authorities. Additionally, the model had to be intended for use in all stages of the child protection process, and not for example, only in the assessment stage. Although assessment is an integral part of a child protection process, our focus was on practice models as a whole. For this reason, differential response options (Child Welfare Information Gateway, & Children's Bureau, 2014) are also excluded. These models aim to change all child protection practice, and not only one part of it. This approach also builds on a systematic review focusing on assessment models (Barlow, Fisher, & Jonas, 2012). The Family Group Conference (FGC) and its adaptions were as such excluded from this review because they do not meet all the above practice model eligibility criteria. Specifically, FGC provides a framework for decision-making but does not shape all practice as explicated above. We also excluded locally developed innovations that have not been disseminated to other agencies as we wanted to identify evidence regarding practice models with at least limited evidence of transferability and scalability.

The main outcomes of interest involved all child-related outcomes (using parent- child- social worker or teacher-reports, client record data; or objective measures of outcome) relating to social, emotional or behavioural functioning, school-related outcomes, etc. We also extracted

data for all parent-related outcomes such as parental mental health, attitudes and behavior, as well as family level outcomes such as family functioning. These outcomes were selected because there is currently no consensus on the most important outcomes of children's social services (Forrester, 2017), and as such we treated a range of improvements in child and family well-being as important outcomes for child protection services. Table 1 summarizes our inclusion and exclusion criteria.

[Insert Table 1 here]

2.2. Systematic searches, data extraction and synthesis

The search procedure was as follows. First, ten electronic databases were searched between February and March 2019: Applied Social Sciences Index & Abstracts (ASSIA); Web of Science (Social Sciences Citation Index, Conference Proceedings Citation Index- Social Science & Humanities, Emerging Sources Citation Index); Cumulative Index to Nursing and Allied Health Literature (CINAHL); EBSCOhost Ebook Collection; Medical Literature Analysis and Retrieval System Online (MEDLINE); OATD - Open Access Theses and Dissertations; PsychINFO; Scopus; Social Services Abstracts; Sociological Abstracts. The following search strings were used: ("practice model*" OR "practice framework*" OR "practice approach*" OR "practice program*") AND ("child* protect*" OR "child* welfare" OR "child* safeguard*") AND (effect* OR outcome*).

Second, reference lists of full texts were screened and assessed for eligibility. The full texts were also assessed as to whether they focused on an eligible intervention, i.e., child protection practice model. If the full paper did not describe the elements of the model in sufficient detail, we searched for a model handbook or other descriptive material that was publicly and freely

available in order to assess whether the model met the inclusion criteria. Based on our screening of the full texts, three eligible models were identified (i.e., Solution-Based Casework, Reclaiming Social Work and Signs of Safety), which we then hand searched in additional key databases (ASSIA and Social Services Abstracts) and grey literature (Google and Google Scholar) for eligible studies between June and August 2019. Third, additional searches were conducted in Cochrane and Campbell Collaboration libraries. Fourth, reference lists of included studies were screened.

One reviewer screened the titles and abstracts of all electronic database references identified by the search strategy. Three reviewers searched the grey literature. Clearly irrelevant references were excluded. In order to be selected, the abstracts had to clearly identify the population and model described above. RefWorks was used to manage references and remove duplicates. An eligibility form developed from the inclusion criteria was used for screening abstracts and full texts. Three reviewers assessed independently full text of studies that were likely to meet inclusion criteria. When the reviewers' conclusions differed, the study was reviewed jointly or resolved by a fourth reviewer.

The following data was extracted from included studies: authors, publication date and type, setting, study design and methods, name of the model, brief description of the model, participants/sample, comparison, outcomes, and effect sizes. We used the Quality Assessment Tool for Quantitative Studies (Thomas, Ciliska, Dobbins, & Micucci, 2004) for each of the papers that were reviewed. We added a question "Was the study conducted by researchers independent of the developer?"

Meta-analysis was not performed because of the small number of included studies and high level of heterogeneity in terms of the included models and outcomes.

3. Results

3.1. Description of the studies

In total, 1360 possibly eligible citations were identified from all searches. After screening the titles and abstracts, 77 full-text articles were screened for inclusion. Our final sample consisted of five papers (representing six studies) focusing on three models. Figure 1 displays the PRISMA flow diagram of our search and selection process.

[Insert Figure 1 here]

Three papers (representing four studies) focused on Solution-Based Casework (SBC), one on the Reclaiming Social Work (RSW) model, and one on Signs of Safety (SoS) were included (see Table 2 for a description of the models). SBC and SoS are rooted in a solution-based approach, whereas RSW involves a systemic approach, and all of them emphasize the relational aspect of social work practice. All of these models are applied in public child protection service settings, and all evaluations were conducted in a context and involved child-protection-involved families as study participants.

[Insert Table 2 here]

All included studies applied a quasi-experimental design. Four papers were peer-reviewed articles (Antle, Barbee, Christensen, & Martin, 2008; Antle, Barbee, Christensen, & Sullivan, 2009; Antle, Christensen, van Zyl, & Barbee, 2012; Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018), while one was a study report (Bostock et al., 2017). All studies were conducted in high-income countries, three in the USA (Antle, Antle, Barbee, Christensen, & Martin, 2008; Barbee, Christensen, & Sullivan, 2009; Antle, Christensen, van Zyl, & Barbee, 2012), one in the

UK (Bostock et al., 2017), and one in the Netherlands (Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018).

The control groups had either received less training in the model of interest (Antle, Barbee, Christensen, & Martin, 2008; Bostock et al., 2017) or were using a different approach to the model of interest, i.e., Intensive Family Case Management, representing "the standard approach at the involved child welfare agency" (Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018, p. 180). Antle, Christensen, van Zyl, and Barbee (2012) compared a high adherence-SBC implementation group and a low adherence-SBC implementation group. The implementation level was evaluated with the public child welfare system's Continuous Quality Improvement tool representing core elements of the SBC model. Only Antle, Barbee, Christensen, and Sullivan (2009) reported that the control group did not implement the SBC model at all. Nevertheless, given that the authors of the study also reported that the workers were referred to groups according to the degree of implementation of the SBC model, it is assumed that all workers were somewhat familiar with the model, especially since statewide implementation efforts had already taken place. The follow-up periods were: 3 months (Bostock et al., 2017; Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018) and 6 months (Antle, Barbee, Christensen, & Sullivan, 2009). No follow-up period was reported for Antle, Barbee, Christensen, and Martin (2008) or Antle, Christensen, van Zyl, and Barbee (2012).

Outcomes of interest were child maltreatment based on state-level maltreatment recidivism referrals (Antle, Barbee, Christensen, & Sullivan, 2009), federal outcomes of safety, i.e., the protection of children from abuse and neglect, the maintenance of children in their own homes and services to prevent removal and risk of harm (Antle, Christensen, van Zyl, & Barbee, 2012), and self-report instruments, i.e., the Actuarial Risk Assessment Instrument Youth Protection and

the Child Abuse Potential Inventory (Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018), assessment of family and service system empowerment, i.e. using the Family Empowerment Scale (Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018), federal outcomes of well-being, i.e., involvement of the family in case planning, meeting educational needs, children receiving services to meet their physical and mental health need (Antle, Christensen, van Zyl, & Barbee, 2012) as well as achievement of case goals and objectives (Antle, Barbee, Christensen, & Sullivan, 2009; Antle, Barbee, Christensen, & Martin 2008), federal outcomes of permanency, i.e., elements of foster care, reunification, permanency goals, and adoption of children as well as preservation of family relationships and connections (Antle, Christensen, van Zyl, & Barbee, 2012), entry to care (Bostock et al., 2017) and other legal actions (Antle, Barbee, Christensen, & Martin, 2008). Antle, Christensen, van Zyl, and Barbee (2012) involved limited descriptions of the definitions and content of the federal measures of safety, permanency and well-being.

Included studies had an unclear or high risk of bias in several domains of the quality assessment tool (see Appendix for study and participant characteristics). The sample sizes were small as four studies involved 100 (Antle, Barbee, Christensen, & Martin, 2008, study two) or less (Antle, Barbee, Christensen, & Martin, 2008, study one; Bostock et al., 2017; Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018) cases. Two studies had a large sample size, 4,559 cases in total (Antle, Christensen, van Zyl, & Barbee, 2012) and 760 cases from 77 practitioners (Antle, Barbee, Christensen, & Sullivan, 2009).

Antle, Christensen, van Zyl, and Barbee (2012) specified that all cases that were selected for the target state's Continuous Quality Improvement (CQI) process, were also selected for the study. These CQI cases were randomly selected from all 9 service regions of the state on a monthly basis. In other studies, it was unclear how the participants were selected to the intervention

group. Antle, Barbee, Christensen, and Sullivan (2009) reported that they had selected all open cases for the two study groups from both SBC workers and control group workers, the latter having been assigned to the two groups based on degree of implementation of the model, not further defined. Likewise, Antle, Barbee, Christensen, and Martin (2008) used degree of training (study one) and implementation scores (study two) to distinguish the study groups, but these processes were not defined. Reekers, Dijkstra, Stams, Asschera, and Creemers (2018) reported that SoS group was selected from the agency where the qualitative study was undertaken, but it was unclear how the team of seven workers was originally selected to implement SoS in that agency. Finally, Bostock et al., (2017) reported that due to recruitment challenges, most of the sample consisted of teams and workers who volunteered to ask families whether they would participate in the study. The authors' initial plan involved recruiting a random sample of families allocated to specific teams over the study period. None of the studies applied randomization or blinding both of which are difficult to conduct for child protection interventions.

In terms of the comparability of the groups, Reekers, Dijkstra, Stams, Asschera, and Creemers (2018) used propensity score matching to match families. Bostock et al., (2017) compared family welfare scores, parent identified concerns and social workers' ratings of concerns at baseline between the two study groups but not the baseline demographic characteristics of the groups. In both study one and two presented in Antle, Barbee, Christensen, and Martin (2008), the authors compared for differences in baseline characteristics, such as type of maltreatment involved for the intervention and control groups, in the analysis. Although Antle, Barbee, Christensen, and Sullivan (2009) had made efforts to compare worker differences "by matching the sample along a number of dimensions known to affect child welfare outcomes" (p. 1350), the authors did not describe what this involved. However, they specified that selecting all cases from participating

workers would provide "a balance of cases by type of maltreatment, severity of maltreatment, comorbid factors, prior involvement with the child welfare agency, and demographic characteristics of the families" (p. 1349). Antle, Christensen, van Zyl, and Barbee (2012) did not comment on baseline characteristics between low and high SBC adherence groups.

3.2. Main findings

Antle, Christensen, van Zyl, and Barbee (2012) found that there was a significant difference between high adherence and low adherence SBC groups favouring the former, for all federal outcomes (permanency, well-being and safety, all p-values < .0001). Further analyses demonstrated that a high degree of implementing the SBC key skills predicted overall safety, permanency, and well-being significantly (p = .001). Antle, Barbee, Christensen, and Sullivan (2009) found that cases in the SBC model group experienced significantly fewer recidivism referrals (351) compared to the control group with 358 referrals (t (73) = - 4.52, p < .0001). An earlier evaluation by Antle, Barbee, Christensen, and Martin (2008) found that the SBC group was more likely to reach the case goals and objectives than the control group. In the first study reported in Antle, Barbee, Christensen, and Martin (2008), the mean number of goals was 6.00 in the intervention group (SD = 2.62) compared to 1.09 (SD = 2.21) in the control group (p < .0001), and the overall effect size (standardized mean difference) was 2.21. There were also significantly less legal actions (e.g., child removals) in the SBC group compared to the control group (2.46 vs. 4.5; p < .001). In the second study reported in Antle, Barbee, Christensen, and Martin (2008), individual- and family level objectives were met in 16.3% of SBC families, while none of objectives were reached in the control group the difference being statistically significant for both family level objectives (x2(2) = 8.25, p < .05) and individual level objectives (p < .05). Effect size (absolute risk reduction) was 16.3.

Bostock et al., (2017) found no significant difference in the number of children entering care in the RSW group compared with the control group. Reekers, Dijkstra, Stams, Asschera, and Creemers (2018) found that SoS and control group were equally effective in reducing the risk of child maltreatment (no effect for time*group, $p = 0.17$, $\eta p2 = 0.05$). Likewise, no significant main effect for time*group was found either for the family empowerment score or for the service system empowerment score.

4. Discussion

The purpose of this review was to synthesize evidence addressing the effectiveness of child protection practice models compared to regular child protection practice. Although the implementation of these models represent a potential improvement on standard practice, and thereby the possibility of improving outcomes for children and families, we identified few controlled studies assessing the effectiveness of the models in terms of key child- and family-level outcomes. Further, based on the quality assessment, the identified studies were weak methodologically in terms of the risk of selection bias, small sample sizes leading to the studies being underpowered with limited statistical analyses, short-term follow up, and reliance on single-source data. The studies were also poorly reported making further assessment of bias difficult. While a number of studies have also found positive practitioner experiences regarding the use of other models such as SoS and RSW (e.g., Bostock et al., 2017; Sheehan et al., 2018) and a statistically significant relationship between systemic supervision quality and overall quality of direct child protection practice (Bostock, Patrizo, Godfrey, & Forrester, 2019), the current findings suggest that there is still a lack of rigorous evidence demonstrating that these models lead to better outcomes for children and families. Furthermore, as a result of the focus on

models as a whole, the findings of this review do not enable us to assess to what extent the presence or absence of different components are influencing the results, something that requires a much larger body of evidence to be able to assess.

The conduct of high-quality outcome evaluations in child protection settings is an extremely challenging task, and as such the included studies represent an important attempt to address the above evidence gap. Indeed, there are several inherent challenges in studying the effects of the practice models that may explain the paucity of high-quality quantitative studies. Examples include variety of work in children's social care, issues in the operationalization and measurement of outcome variables, cultural and organizational resistance and lack of research infrastructure in social services, assessment of fidelity, and difficulties in the recruitment and retention of participants when practitioners have limited resources and child protection-involved families have highly complex live situations (e.g., Forrester, 2017; Gillingham, 2018; Mezey et al., 2015). Furthermore, implementing practice models in different contexts involves several implementation barriers that also create challenges to an outcome evaluation (e.g., Bostock et al., 2017; Roberts, Caslor, Turnell, Pearson, & Pecora, 2019). For example, previous studies have identified that e.g., leadership and organisational climate, training and coaching, alignment with other organisational systems and initiatives, time and resources and staff permanency influence implementation of practice models (Antle et al., 2012; Lambert, Richards, & Merrill, 2016; Sanclimenti, Caceda-Castro, & DeSantis, 2017; Sheehan et al., 2018). Despite this, rigorous evaluations within child protections have been successfully conducted (e.g., Chaffin et al., 2004) and these new models of practice are deserving of the same level of rigour in terms of evaluation.

4.1. Strengths and limitations

There are a number of limitations to this systematic review. The small number of studies meeting our inclusion criteria is likely related to our stringent inclusion criteria, which were targeted at identifying studies of practice models with child- and family-level outcome measures in public child protection services. Although there is a growing body of literature regarding practice models, the terms used to describe these vary significantly (e.g., "practice frameworks", "change programmes", and "intervention models" - Gillingham, 2018; Lwin, Versanov, Cheung, Goodman, & Andrews, 2014; Laird, Morris, Archard, & Clawson, 2018); and as such, while we developed a comprehensive set of search terms with the aim of increasing the sensitivity of the search, we may not have identified all existing models. Furthermore, the focus on studies written in English means that we may well have failed to identify evaluations of other practice models published in non-English language journals (e.g., Holmgård Sørensen, 2009; Vink, de Wolff, van Dommelen, Bartelink, & van der Veen 2017). Psychosocial treatments and family preservation services (as opposed to practice models) that have proven effective in relevant child related outcomes, such as preventing child maltreatment or out-of-home placements in CPS population, were also excluded (e.g., Bezeczky et al., 2020; Chaffin et al., 2004).

4.2. Implications for future research

This review identifies several avenues for future research. Recent years have seen an intense debate regarding whether randomised controlled trials (RCT's) are applicable to the field of social sciences. RCT's are seen as lacking nuance and disregarding the social realities, context and complexities of the situations that child protective services face (De Jong, Schout, & Abma, 2015). Furthermore, it has been argued that while RCT's provide "an unbiased estimate", their results are not generalizable since these estimates apply only to the sample selected for the trial (Deaton, &

Cartwright, 2018). Defendants have argued that despite its flaws, RCT's are still the best choice in terms of being able to make causal inferences (Creemers et al., 2017), but that they need to be combined with other methods and theorisation of "why things work" and how this will vary across different contexts (Deaton, & Cartwright, 2018). There is also a need to think about the outcomes in children's social care in innovative ways (Forrester, 2017).

Despite these challenges, if we are to assess the effectiveness of practice models, the field should aim for rigorous mixed-method studies such as realist randomised trial designs (e.g., Bonell, Warren, Fletcher, & Viner 2016) or other kinds of high-quality study designs. For example, where randomization is not possible, quasi-experimental methods should be used such as difference in difference designs with propensity score matching (see, for example, Austin, 2011). These experimental and quasi-experimental designs allow for appraisal of the effectiveness of the models, whilst also addressing implementation in differing contexts in terms of the participating services. They can also be used to detect unexpected effects as well as the subjective experiences of the participating professionals and families. When the aim is to measure child- and family-level outcomes, multiple data sources and informants should be included, such as data gathered directly from children and youth (Sweeting, 2001). It should also be noted that effective study designs require sufficient time and funding (Baginsky, Moriarty, & Manthorpe, 2019). Specifically, funding one large-scale high-quality evaluation might create more robust knowledge in terms of service improvement compared to several smaller initiatives.

Finally, lack of high-quality evidence does not mean that child protection practice models do not work, nor do we suggest that agencies should forgo applying them as part of their service provision. However, it is important that leaders and practitioners in children's services acknowledge that the absence of effectiveness research is problematic, and work alongside researchers to secure the

necessary funding to undertake such evaluation prior to any large-scale implementation. Therefore, despite the high level of interest with regard to the use of practice models, the current findings suggest that service providers should proceed with caution, in terms of their implementation.

5. Conclusions

Child protection practice models have been widely adopted in a number of countries. Despite the growing body of research on these models, evidence of their effectiveness in terms of child- and parent-related outcomes, is still limited. Although the conduct of outcome evaluations in children's social care involves multiple challenges, the results of this review suggest that more high-quality studies are urgently needed to evaluate which, if any, of these practice models improve outcomes for child-protection-involved families.

Funding

References

Antle, B. F., Barbee, A. P., Christensen, D. N. & Martin, M. H. (2008). Solution-Based Casework in Child Welfare: Preliminary Evaluation Research. *Journal of Public Child Welfare, 2*(2), 197-227. doi: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. doi: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 Solution Based Casework (SBC) practice model on federal outcomes in public child welfare. *Child Abuse & Neglect, 36*(4), 342-353. doi:10.1016/j.chiabu.2011.10.009.

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research: Propensity Score Analysis, 46*(3), pp. 399-424. doi:10.1080/00273171.2011.568786.

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 182 child welfare agency: Application of the Getting to Outcomes (GTO) model. *Children and Youth Services Review, 33*(5), 622–633. doi:10.1016/j.childyouth.2010.11.008.

Baginsky, M., Moriarty, J. & Manthorpe, J. (2019). Signs of Safety: Lessons learnt from evaluations. *Journal of Children's Services, 14*(2), 107-123. doi:10.1108/JCS-11-2018-0028.

Barlow, J., Fisher, J. & Jonas, D. (2012). Systematic review of models of analysing significant harm. Oxford University: Department for Education.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/183949/DFE-RR199.pdf

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), pp. 451-468. doi:10.1080/23303131.2016.1159637.

Bezeczky, Z., El-Banna, A., Petrou, S., Kemp, A., Scourfield, J., Forrester, D. & Nurmatov, U. B. (2020). Intensive Family Preservation Services to prevent out-of-home placement of children: A systematic review and meta-analysis. *Child Abuse & Neglect, 102 (2020)*104394. doi:10.1016/j.chiabu.2020.104394.

Bonell, C., Warren, E., Fletcher, A., & Viner, R. (2016). Realist trials and the testing of context-mechanism-outcome configurations: a response to Van Belle et al. *Trials*, 17(1), 478. doi:10.1186/s13063-016-1613-9.

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*. doi:10.1016/j.childyouth.2019.104428.

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. Children's Social Care Innovation Programme Evaluation Report 45. Tilda Goldberg Centre for Social Work and Social Care, University of Bedfordshire & CASCADE: Children's Social Research and Development Centre, University of Cardiff.

https://www.gov.uk/government/publications/scaling-and-deepening-the-reclaiming-social-work-model

Bunn, A. (2013). Signs of safety® in England: An NSPCC commissioned report on the Signs of Safety model in child protection. London, UK: NSPCC. https://www.cornwall.gov.uk/media/10965889/signs-of-safety-model-nspcc-report.pdf

Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., . . . Bonner, B. L. (2004). Parent–Child Interaction Therapy With Physically Abusive Parents: Efficacy for Reducing Future Abuse Reports. *Journal of Consulting and Clinical Psychology, 72*(3), 500-510. doi:10.1037/0022-006X.72.3.500.

Child Welfare Information Gateway & Children's Bureau. (2011). *Differential Response to Reports of Child Abuse and Neglect.* https://www.childwelfare.gov/pubPDFs/differential_response.pdf

Creemers, H. E., Sundell, K., Deković, M., Dijkstra, S., Stams, G. J. J. M. & Asscher, J. J. (2017). When the 'Golden' Standard Should Be the General Standard: Response to a Commentary on the Use of Randomised Controlled Trials to Examine the Effectiveness of Family Group Conferencing. *British Journal of Social Work, 47*(4), 1262-1267. doi:10.1093/bjsw/bcw060.

Deaton, A. & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine, 210*, 2-21. doi:10.1016/j.socscimed.2017.12.005.

De Jong, G., Schout, G. & Abma, T. (2015). Examining the Effects of Family Group Conferencing with Randomised Controlled Trials: The Golden Standard? *The British Journal of Social Work, 45*(5), 1623-1629. doi:10.1093/bjsw/bcv027.

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

Gilbert, R., Kemp, A., Thoburn, J., Sidebotham, P., Radford, L., Glaser, D. & Macmillan, H. L. (2009). Recognising and responding to child maltreatment. *The Lancet, 373*(9658), 167-180. doi:10.1016/S0140-6736(08)61707-9.

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. doi:10.1080/15548732.2017.1392391.

Holmes, L. & Mcdermid, S. (2013). How social workers spend their time in frontline children's social care in England. *Journal of Children's Services, 8*(2), 123-133. doi:10.1108/JCS-03-2013-0005.

Holmgard Sorensen, T. (2009). Med Familien I Centrum - en historie om implementering af anerkendende og løsningsfokuserede metoder i familiebehandling og sagsbehandling. Evalueringsrapport fra projekt Familien i Centrum i Københavns Kommune. [Family in Focus - a story of implementing acknowledged and solution-focused methods in family and case management. Evaluation report from the project Family in Focus in the City of Copenhagen]. Servicestyrelsen. https://www.kk.dk/sites/default/files/migrated/sc/Med-familien-i-Centrum---en-historie-om-implementering-af-anerkendende-og-loesningsfokuseret.pdf?token=iMqxdZcI

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. doi:10.1177/1473325016688371.

Lambert, D., Richards, T. & Merrill, T. (2016). Keys to Implementation of Child Welfare Systems Change Initiatives. Journal of Public Child Welfare, 10(2), pp. 132-151. doi:10.1080/15548732.2015.1113226.

Lwin, K., Versanov, A., Cheung, C., Goodman, D. & Andrews, N. (2014). The Use of Mapping in Child Welfare Investigations: A Strength-based Hybrid Intervention. *Child Care in Practice, 20*(1), 81-97. doi:10.1080/13575279.2013.847055.

Mezey, G., Robinson, F., Campbell, R., Gillard, S., Macdonald, G., Meyer, D., … White, S. (2015). Challenges to undertaking randomised trials with looked after children in social care settings. *Trials*, *16*, 206. doi:10.1186/s13063-015-0708-z.

Munro, E. (2011). The Munro review of child protection. Final report. A child centred system. London: Department for Education.

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. doi:10.1016/j.childyouth.2018.06.012.

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. doi: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), pp. 279-298. doi:10.1080/15548732.2016.1275920.

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 focused 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

STM. (2019). Selvityshenkilön arvio lastensuojelun henkilöstön riittävyydestä ja ratkaisuehdotukset kuormituksen vähentämiseksi: loppuraportti. [Investigation person's assessment of child protection staff's sufficiency and proposed solutions to decrease burden: final report]. Ministry of Social Affairs and Health reports and notes 2019:4. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161379/Selvityshenkilon_edellytykset _lastensuojelun.pdf?sequence=4&isAllowed=y

Sweeting, H. (2001). Our family, whose perspective? An investigation of children's family life and health. *Journal of Adolescence, 24*(2), 229-250. doi:10.1006/jado.2001.0376.

Thomas, B. H., Ciliska, D., Dobbins, M. & Micucci, S. (2004). A Process for Systematically Reviewing the Literature: Providing the Research Evidence for Public Health Nursing Interventions. *Worldviews on Evidence-Based Nursing, 1*(3), 176-184. doi:10.1111/j.1524-475X.2004.04006.x.

Vink, R., de Wolff, M., van Dommelen, P., Bartelink, C., and van der Veen, S. (2017). Empowered door Signs of Safety? Onderzoek naar de werkzaamheid van Signs of Safety in de Jeugdzorg. [Empowered by Signs of Safety? Effectiveness study of Signs of Safety in Youth Care]. Leiden, TNO. https://www.nji.nl/nl/Download-NJi/Publicatie-NJi/Empowered-door-Signs-of-Safety-Onderzoek.pdf

Table 1.

Inclusion and Exclusion Criteria

<table>
<tr><td>
Inclusion criteria:
1. Population: Children aged 0-17 years and parents involved in child protection services
2. Intervention of interest: Child protection practice models (licensed and non-licensed)
3. Outcomes: All child-related outcomes (using parent- child- social worker or teacher-reports, client record data; or objective measures of outcome) relating to social, emotional or behavioral functioning; school-related outcomes etc. Additional outcomes were all parent related outcomes (as above) such as parental mental health; attitudes and behavior; etc or family outcomes such as family functioning.
4. Comparison group: Child protection social work practice that used no specific model or other well-matched control groups.
5. Study setting: Statutory child protection social work practice provided by public authorities.
6. Study type: Quantitative studies that are minimally controlled before-after studies.
7. Publication type: Any type.
8. Languages: English only.
9. Data range: From 1990 until March 2019 (when main search was executed).
</td></tr>
<tr><td>
Exclusion criteria:
1. Ineligible population. Models targeted to specific populations or conditions (e.g., children with disabilities). Studies that were not conducted in child protection settings and models that were not provided by public authorities.
2. Ineligible intervention. All models that focused only on assessment or residential treatment. Locally developed innovations that had not been disseminated to other agencies. The Family Group Conference and its adaptions.
</td></tr>
</table>

Table 2.

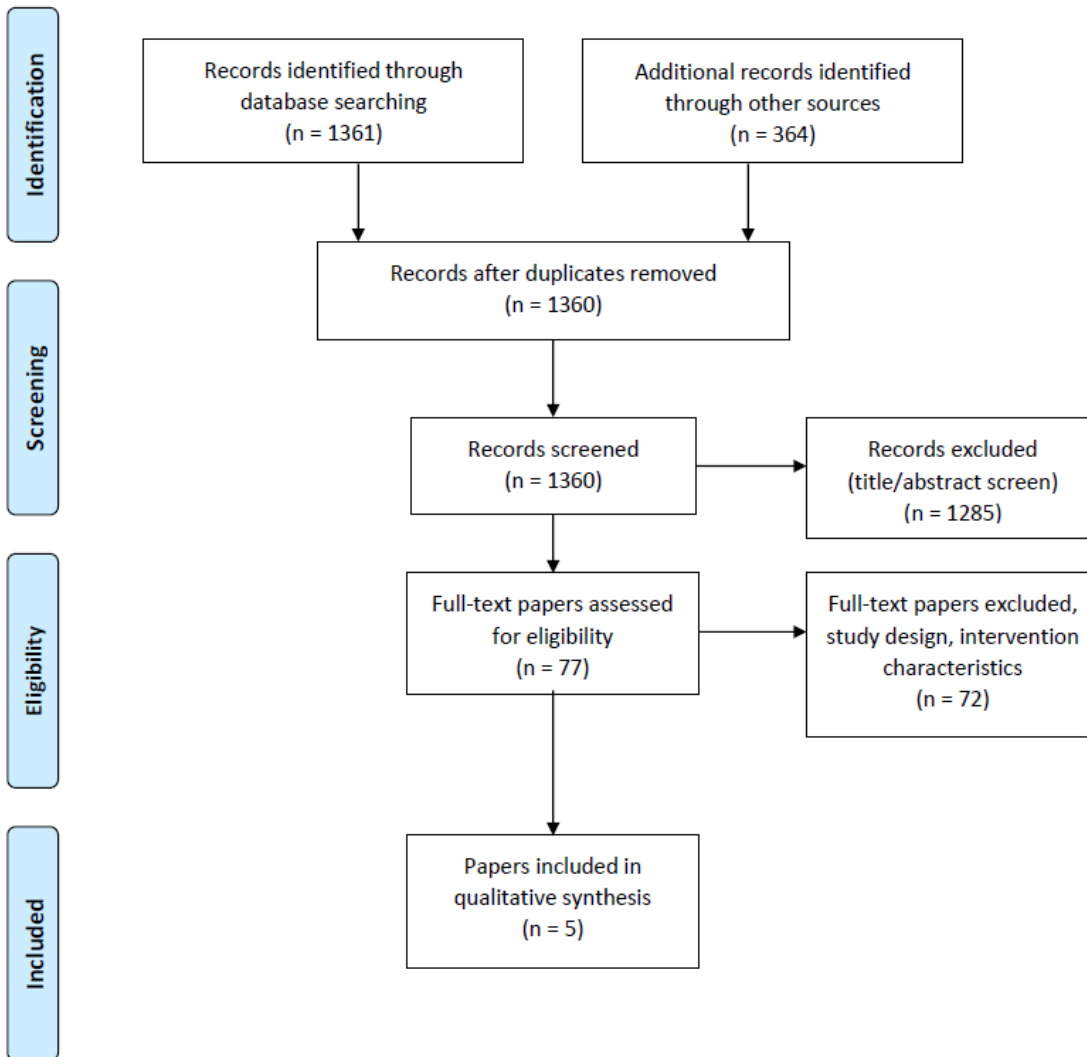*Model characteristics of included studies*

Model description	Included studies	Theoretical basis	Key skills and tools
Solution-Based Casework (SBC) promotes strengths-based practice where full partnership with the family is central. Developed in the USA in the 1990's, the model presumes families already possess skills that can be used to prevent child maltreatment. Further, it presumes that families progress through developmental stages, and many of the problems they encounter can be described as non-pathological, situational, universal and related to developmental tasks. Finally, in order to prevent relapses to high-risk behavior, the parents are assisted in identifying the situations and behavior patterns associated with child maltreatment. Case plans and objectives are formulated for family-level as well as individual level.	(Antle, Barbee, Christensen, & Martin, 2008; Antle, Barbee, Christensen, & Sullivan, 2009; Antle, Christensen, van Zyl, & Barbee, 2012)	Solution-focused family therapy, family life cycle theory, and relapse prevention strategies drawing from cognitive behavior therapy.	Solution-focused interviewing techniques encouraging the parents to identify strengths and exceptions to problematic situations. Parents are helped to develop strategies to avoid destructive behavior patterns and situations.
Reclaiming Social Work (RSW) model is a systemic approach to child protection rooted in systemic family therapy. The model was developed in London Borough of Hackney	(Bostock et al., 2017).	Milan School of social constructivist family therapy.	Systemic family therapy techniques such as hypothesizing, using genograms to understand family patterns, reflexivity and curiosity, use of reflexive questions,

children's social care in 2000's. In systemic practice, families are viewed as systems instead of individuals, and multiple perspectives and solutions to problems are reflected. Change is facilitated by encouraging reflexivity and new insights on how beliefs and circular patterns of behaviour affect others. In order to learn and maintain systemic practice, social workers and child practitioners work in small systemic units, which are led by a consultant social worker, who has the ultimate responsibility for case decision-making. Each unit also has a qualified systemic family therapist as a clinician, and a unit coordinator providing administrative support. Systemic units hold weekly unit meetings, which are the main forum for shared decision making and case supervision.			approaching families with curiosity rather than making assumptions.
Signs of Safety (SoS) is a strengths-based and safety-focused approach to child protection practice. Developed through the 1990's in Western Australia by Andrew Turnell and Steve Edwards in collaboration with child protection practitioners, the model draws upon techniques from	(Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018)	Strengths-based and safety-focused, particularly Solution Focused Brief therapy.	Techniques from Solution Focused Brief therapy, such as working with family strengths and resources, finding exceptions, goal setting and scaling, using a transparent approach.

Solution Focused Brief therapy and has two core principles: establishing a working relationship between professionals and parents and supporting parental empowerment. Ultimately, the aim is to involve children and families in effective safety planning to improve the child safety.			

Figure 1.

PRISMA flow diagram

Appendix.

Study and participant characteristics.

Study	Model	Sample	Comparison	Outcome	Effect	EPHPP rating	Comment
(Antle, Barbee, Christensen, & Martin, 2008)							
Study one	Solution-Based Casework (SBC)	Total of 48 cases SBC-group: 27 cases Control group: 21 cases	Cases of practitioners who had not received the SBC training (or had received a lower degree of training, i.e., a team in which only the supervisor received 1 day of training on SBC but not others)	Case outcomes: achievement of case goals and objectives	Mean number of goals 6.00 in the intervention group (SD = 2.62) vs, 1.09 (*SD* = 2.21) in the control group, (F(1) = 30.53, *p* < .0001). Effect size (standardized mean difference) 2.21. There were significantly less legal actions (including removal of children from home) in SBC group compared to the comparison group (2.46 vs. 4.5; t (45 = 3.65, *p* < .001). Further, in multiple regression analysis the level of SBC implementation predicted the number of goals achieved (*p* < .05).	Weak	Selection bias Small sample Intervention group consisted of a team with a high degree of training in SBC, while the control group of a team with a lower degree of training. However, the team selection process was vaguely described.
Study two	Solution-Based Casework (SBC)	Total of 100 cases SBC-group: 50 cases Control group: 50 cases	Cases of service providers who implemented the SBC weakly (below the implementation level median score of the whole sample)	Case outcomes: achievement of case goals and objectives	Individual- and family level objectives were met in 16.3% families in the intervention group and for 0% in the control group. The difference was statistically significant for both family level objectives (x2(2) = 8.25, *p* < .05) and individual level objectives (x2(2) = 8.25, *p* < .05). Effect size (absolute risk reduction) was 16.3.	Weak	Selection bias Small sample Cases were selected based upon degree of implementation, but the case selection process was vaguely described.

(Antle, Barbee, Christensen, & Sullivan, 2009)	Solution-Based Casework (SBC)	Total of 77 practitioners and 760 cases SBC-group: 39 practitioners and 339 cases Control group: 38 practitioners and 421 cases	Cases of workers who did not implement the SBC model. Workers were assigned to groups based upon degree of implementation of the model.	6-month standardized state-level abuse recidivism data	Number of recidivism referrals 350.69 for the SBC group, whereas 538.00 for the control group in the 6-month follow-up. The difference was statistically significant, t (73) = - 4.52, $p < .0001$.		The numbers are reported as averages instead of actual numbers of recidivism referrals per group No description of the case characteristics Limited data on how the workers were selected to the study Limited description of study groups' level of implementation
(Antle, Christensen, van Zyl, & Barbee, 2012)	Solution-Based Casework (SBC)	4,559 CPS cases from Kentucky. All cases that were selected for the state's Continuous Quality Improveme nt (CQI) process during a 4-year time period (2004–2008) were included in the study. The CQI	Level of implementation of SBC model, comparison of high and low SBC adherence cases. This assessment was based on the Continuous Quality Improvement (CQI) - measure.	Outcomes were federal definitions of -safety: (1) protection of children from abuse and neglect (2) maintenance of children in their own homes and services to prevent removal and risk of harm) -permanency (1) elements of foster care, reunification, permanency goals, and	According to t-tests, there was a significant difference between high adherence and low adherence SBC groups for all federal outcomes (permanency, well-being and safety). In regression analyses, and the use of SBC predicted overall safety, permanency, and well-being significantly ($p = .001$). Different factors of SBC contributed differently to the outcomes, SBC-intake and investigation being the most important factor in predicting overall safety (Beta 0.592, 95% CI 0.482-0.528), SBC-case management in predicting overall permanency (Beta 0.418 95% CI 0.302-0.399), while for overall well-being SBC-ongoing, SBC-case management and SBC- case	Weak	Limited data on the case characteristics Broad outcome variables with limited descriptions of the definitions and content

		cases were randomly selected from all nine service regions of the state on a monthly basis.		adoption of children (2) preservation of family relationships and connections) - well being (1) involvement of the family in case planning, meeting educational needs (2) children receiving services to meet their physical and mental health need).	planning, all made substantial contributions to predicting overall well-being scores (Betas ranging from 0.299 to 0.323). The mean percentage scores for all the outcome variables were better for the high adherence SBC cases compared to low adherence cases, High adherence groups were able to meet and exceed the federal standards on all the safety, permanency and well-being variables, whereas for low adherence this was only true for permanency related outcomes that were related to children's living situations.		
(Bostock et al., 2017)	Reclaiming Social Work (RSW) model	Total: 86 families RSW-group: 34 families Control group: 52 families	Service as usual	Entry to care	Number of children entering care at T2: n = 0 in RSW group, n = 2 in comparison group, the difference was statistically non-significant (p-value not reported).	Weak	Selection bias Small sample Control group involved a range of different types of team setup and training, including previous systemic training While child protection specific baseline characteristics were compared between study groups, demographic characteristics were not

							Attrition ruled out using intended indicators in the analysis other than entry to care Short duration of the study (3 months)
(Reekers, Dijkstra, Stams, Asschera, & Creemers, 2018)	Signs of Safety (SoS)	Total of 37 families SoS-group: Parent report: 18 families Social worker report: 17 families Control group: Parent report: 20 families Social worker report: 20 families	Care as usual	Child maltreatment measured with ARIJ The risk of child maltreatment measured with the ARIJ (social worker) and CAPI (parents)	Due to the small number of reported child maltreatment cases in T2 (one in SoS-group and one in the control group), the logistic regression was not performed. Mean score for the risk of child maltreatment in the social worker report at T1 was 0.52 (SD = 0,31) and at T2 0.28 (SD = 0.29) in the SoS group vs. 0.38 (SD = 0.22) and 0.28 (SD = 0.26) in the control group. In the parent report, the mean score was at T1 0.16 (SD = 0.18) and at T2 0.13 (SD = 0.14) in the SoS-group, and 0.17 (SD = 0.23) and 0.16 (SD = 0.15) in the control group. SoS and control group were equally effective in reducing the risk of child maltreatment (no effect for time*group, Wilks'Λ = 0.95, F(1, 35) = 1.99, p = 0.17, $\eta p2$ = 0.05).		Selection bias Small sample Missing data was imputed. Short duration of the study (3 months) Control group used Intensive Family Case Management that was described as a standard approach in the agency Propensity score matching was used to match families
				Family and service system empowerment measured with the FES	Mean score for family empowerment was at T1 4.28 (SD = 0.59) and at T2 4.36 (SD = 0.39) in the SoS-group, vs. at T1 4.19 (SD = 0.51) and at T2 4.38 (SD = 0.43) in the control group.		

| | | | | | In terms of the service system empowerment the mean score was at T1 3.95 ($SD = 0.53$) and T2 4.05 ($SD = 0.54$) for the SoS-group, and at T1 4.08 ($SD = 0.44$) and T2 4.25 ($SD = 0.53$) for the control group. No significant main effect time*group was found either for the family empowerment score or for the service system empowerment score. | | |