



# NO PLACE LIKE HOME

Final Progress Report

*December 2015*

*The Kempe Center for the Prevention and Treatment of Child Abuse & Neglect  
and  
Casey Family Programs*

**Authors:**

Heather Allan, MSW  
Tyler Corwin, MA  
John Fluke, PhD  
Nicole Harlaar, PhD  
Dana Hollinshead, PhD  
Anita Horner, BA  
Erin Maher, PhD  
Lisa Merkel-Holguin, MSW

**Acknowledgment**

*This project was funded through the Children's Bureau, U.S. Department of Health and Human Services, Administration of Children and Families, Children's Bureau, grant # 90F0051.  
This product expresses the views of the authors, not the views of the Children's Bureau.*

## Table of Contents

Executive Summary .....	1
I. Overview of the Community, Population and Needs .....	5
A. Grantee Information.....	5
B. Community Setting .....	6
C. Project Need.....	8
D. Population .....	9
II. Overview of the Program Model .....	11
A. Project Goals and Objectives .....	11
B. NPLH Service Model.....	13
C. Project Activities.....	14
III. Collaboration.....	19
A. Inner-Agency Collaboration .....	19
B. Inter-Agency Collaboration .....	22
IV. Sustainability.....	23
V. Evaluation Planning .....	25
A. Process Evaluation Methodology .....	25
B. Outcome Evaluation Methodology .....	31
C. Evaluation Plan Revisions .....	38
D. Evaluation Trainings and Engagement .....	40
VI. Process Evaluation Results .....	41
A. General Child Welfare Staff Survey .....	41
B. Fidelity Results - Larimer County .....	43
C. Fidelity Results - Texas DFPS Region 3 .....	55
VII. Outcome Evaluation Results .....	69
A. Outcome Analysis - Larimer County .....	69
B. Outcome Analysis - Texas Region 3 .....	80
C. Cost Analysis .....	91
D. Additional Analyses.....	94
VIII. Evaluation Discussion.....	98
A. Challenges.....	98
B. Limitations .....	100
IX. Conclusions.....	102

A. Major Takeaways.....	102
B. Contributions.....	104
C. Next Steps .....	105
X. Recommendations.....	106
Appendix A. NPLH Logic Model.....	112
Appendix B. NPLH Training and Technical Assistance .....	113
Appendix C. NPLH Instrument List with References .....	116
Appendix D. General Staff Survey Cross-Site Results.....	120
Appendix E. Larimer County Propensity Score Matching (PSM) .....	135
Appendix F. Larimer County Child Characteristics Prior to PSM.....	140
Appendix G. The Implementation of Family Group Decision Making in South Dakota during the No Place Like Home project.....	145
Appendix H. Results of Bayesian Model Averaging.....	168

## Executive Summary

The No Place Like Home (NPLH) evaluation project was a 3-year multi-agency collaboration funded by a Family Connections grant from the U. S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. With the receipt of one year no-cost extension, the grant operated between October 2011 and September 2015. This project strategically combined the assets of the Kempe Center for the Prevention and Treatment of Child Abuse & Neglect, Casey Family Programs and three child welfare agencies with mature family group decision making (FGDM) programs to test the effectiveness of FGDM in safely preventing children receiving in-home child welfare services from entering or re-entering foster care. The project sites—Larimer County Department of Human Services, Colorado; South Dakota Department of Social Services, Rapid City; and Texas Department of Family and Protective Services, Dallas and Tarrant Counties —have over two decades of collective experience in delivering FGDM to children and families who come to the attention of child welfare.

The NPLH project had robust process, outcomes and cost components geared toward answering the following questions:

1. Are children in families in the focus population who experience FGDM interventions less likely to experience placement compared to children in the control group?
2. If children are placed out of home, are they more likely to be placed with relatives compared to the control group?
3. Are families in the population who experience FGDM interventions as likely as families in the control group to experience child maltreatment re-reports or re-reports with substantiation?
4. Are families who experience FGDM processes more satisfied with their experiences with child welfare compared to children in the control group?
5. For all of the outcomes identified above (placement, relative placement, reporting, and satisfaction) are families less likely to have disparate experiences based on race or ethnicity compared to families in the control group?

The process evaluation design included a global or agency-level assessment of how FGDM was designed and implemented by the agency. Key to the process evaluation was determining how the agency- and case-level views are actually integrated, the degree to which policy is consistent with practice, and how agencies identify and resolve implementation problems. The process evaluation involved a combination of qualitative (focus groups and interviews) and quantitative (staff and caregiver surveys; tracking sheet; fidelity surveys) data.

The outcome evaluation consisted of an experimental or quasi-experimental longitudinal design, specific to the practice site, and covering a minimum of 18 months of potential follow-up. The data collection period began in October 2012 and referrals to the project ceased on April 30, 2014 in Larimer County, CO and June 30, 2014 in Texas. A no-cost extension was sought and approved to allow for the maximum data collection period as well as to allow for data analysis. The final set of administrative data (e.g. SACWIS data) was pulled on these two groups to track services and outcomes, such as maltreatment recurrence and placements, in June 2015. During the ~24 month data collection period, interim results were used for formative work with the sites and to inform technical assistance and training efforts. The final six months of the no cost extension project period were used to consolidate data, complete the analysis, and prepare the evaluation report.

The major findings from this study are as follows:

- Low frequency events, such as child removals, can pose a challenge to detecting significant effects.
- Fidelity index scores indicate overall favorable responses from both family and professionals, but professionals and facilitators generally had slightly higher fidelity ratings.
- Orientation toward child safety versus family preservation depended on job type, years of experience, and shared vision.
- Higher ratings of perceived FGC effectiveness depended on worker type, perceptions of local services, and belief in families' ability to construct plans to address issues.
- No statistically significant differences between treatment and comparison/control groups in screened-in re-referrals were found in Larimer or Texas.

- In the case of Larimer, FSRTs may produce some of the same results as FUMs and FGCs for this stage of service (as indicated by lack of significant findings between treatment and comparison groups in Larimer).
- No statistically significant differences between treatment and comparison/control groups in removals were found in Larimer or Texas. Impacts from intermediate outcomes such as more comprehensive, family-centric plans and increases in social support may take longer to be reflected in long-term outcomes.
- Other unmeasured benefits may be present, such as child well-being, positive impacts on agency culture, consistency with agency values, and family engagement over time.
- Propensity score matching is a very complex methodology with important limitations. For this study important questions emerged regarding the difficulty of obtaining appropriate control samples and matching variables.

In summary, this large-scale, multi-site, multi-method evaluation showcased some relevant and unexamined findings related to family meetings. A new fidelity index was created and tested for this project. The FGCs in TX and CO showed favorable fidelity to the model from both family members and the child welfare staff who participated. And, the psychometric analyses performed on the instrument in this project have informed further development of this tool, which will be a useful asset to the field in both practice and research.

Staff attitudes and buy-in are such a critical component of effective practice, as has been demonstrated from studies of implementation science and organizational culture. The worker survey findings shed light on staff attitudes as pertaining to FGCs. Higher ratings of FGC effectiveness depend on whether workers carry a caseload, perceptions of local services, and belief in family abilities to construct plans to address issues. Similarly, worker orientation to child safety versus family preservation also depends on whether or not a worker carries a caseload, years of experience, and perceptions of a shared vision within the organization.

In terms of outcomes, in TX, it was found that families receiving an FGC were perceived by their caseworkers as having a greater amount of social support at the close of their case than families who did not receive an FGC. The importance of social support in the context of child maltreatment has been well established (Thompson, 2015). This finding is an important one. However, significant effects were not found on the likelihood of a re-report or an out-of-home

placement for the full samples in TX. Two racial and ethnic differences in these outcomes were detected, but, as stated in the report, they need additional exploration to rule out spurious results.

# I. Overview of the Community, Population and Needs

## A. Grantee Information

The No Place Like Home (NPLH) evaluation project was a 3-year multi-agency collaboration funded by a Family Connections grant from the U. S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. With the receipt of a no-cost extension of one year, the grant operated between October 2011 and September 2015. This project strategically combined the assets of the Kempe Center for the Prevention and Treatment of Child Abuse & Neglect, Casey Family Programs and three child welfare agencies with mature family group decision making (FGDM) programs to test the effectiveness of FGDM in safely preventing children receiving in-home child welfare services from entering or re-entering foster care. The project sites—Larimer County Department of Human Services, Colorado; South Dakota Department of Social Services, Rapid City; and Texas Department of Family and Protective Services, Dallas and Tarrant Counties —have over two decades of collective experience in delivering FGDM to children and families who come to the attention of child welfare.

The American Humane Association was awarded the grant in 2011. However, when project staff migrated to the Kempe Center for the Prevention and Treatment of Child Abuse & Neglect in 2012, an organization that was ineligible to serve as the lead grantee agency, Larimer County Department of Human Services (LCDHS) assumed that role. LCDHS is an organization that, over the past two decades, has sought to evolve from an 'old fashioned' child welfare system into a modern, family-centered organization that asks families and agency professionals to make decisions together. As a result, Larimer County has reduced the number of children in higher



levels of care, increased services to children in their homes, and improved their agencies' federal outcomes significantly over the last several years.

## B. Community Setting

### Larimer County, Colorado

Larimer County has a population of approximately 325,000 people and is the 6<sup>th</sup> most populated county in Colorado. Geographically Larimer is the 9<sup>th</sup> largest Colorado county, covering 2,596 square miles. Fort Collins is the county seat and home to Colorado State University and the largest municipality within the county serving as home to about half of the counties' people. The community is predominately non-Hispanic white (approximately 83%) with a 10% Hispanic population (U.S Census, 2015). Larimer County is a mix of suburban, urban, and rural tracts. Larimer County Department of Human Services (LCDHS) has been implementing Family Group Conferences since 1998 and has since expanded their family meeting array to include Family Unity Meetings (FUMs) and Family and Safety Resource Team Meetings (FSRTs). In fiscal year 2013 LCDHS received 6,181 child protection referrals and conducted 2,656 assessments. Of the children served by LCDHS in that year, 95.2% were able to remain in-home (e.g. did not experience an out-of-home placement) and 2,052 children (duplicated) participated in a family meeting with 207 participating specifically in an FGC (Allan et al., 2014).

### Texas Region 3

Dallas and Tarrant Counties comprise Region 3 in Texas and are the 3<sup>rd</sup> and 5<sup>th</sup> most populous counties in Texas, respectively, with a combined population of approximately 4,400,000 people. Both are largely urban counties housing the cities of Dallas and Ft. Worth. Dallas County's population is 31% white (non-Hispanic), 23% African American, and 39%

Hispanic. Tarrant County's population is approximately 50% white (non-Hispanic), 15% African American, and 27% Hispanic (U.S. Census, 2015).

Texas Department of Family and Protective Services (DFPS) began implementing Family Group Decision Making using the Family Group Conferencing model in 2002 following the federal Child and Family Services Review (CFSR) which found that Texas CPS was having difficulty engaging families in the case planning and decision-making processes. In addition to FGC, Texas utilizes Family Team Meetings (FTMs) as well as Circles of Support meetings for its youth aging out of the foster care system. In Texas, FGCs are held both during the Family Based Safety Services (FBSS) and Conservatorship stages of service. In fiscal year 2013 Dallas and Tarrant Counties received a combined 37,881 child protection referrals of which 33,457 were accepted for an investigation of which 22,286 were not opened for services following the investigation. During that same fiscal year, approximately 4,000 families and 11,000 children received in-home services (FBSS) comprising the target population for the NPLH project (DFPS 2013 Data Book, 2013). The intent of FBSS is to put in place a range of prevention and support services to decrease the likelihood that children are removed from their homes.

### Rapid City, South Dakota

Rapid City is the county seat of Pennington County, South Dakota which has a population of approximately 110,000 people, 80% of whom are non-Hispanic white and approximately 10% of whom are Native American (U.S. Census, 2015). While Rapid City is the second largest city in South Dakota, the county is largely rural and borders the Pine Ridge Indian Reservation to the South.

South Dakota CPS began implementing FGCs in the Rapid City region in 2005, making it the only non-tribal child protection authority utilizing the practice in the state at that time. In 2006,

the agency expanded its family engagement work by implementing Team Decision Making models, and later added Concurrent Planning Meetings and Placement Team meetings resulting in a continuum of FGDM processes throughout the child welfare system.

### C. Project Need

While the FGDM research base continues to grow—nationally and internationally—there are clearly critical gaps in knowledge that this project sought to help answer. The California Evidence Based Clearinghouse for Child Welfare, using a Scientific Rating Scale based on the published, peer-reviewed research available, rates FGDM as a practice with “promising research evidence” in achieving safety, permanency and well-being for children. This is because the majority of FGDM evaluation to date lacks sufficient rigor, meaning not enough studies have deployed experimental or quasi-experimental designs. Table 1 details some of the foci of existing research and some key research gaps, with an important notation that the items listed in Column A still need additional rigorous study. This project sought to address many of these gaps by implementing a rigorous evaluation design to examine topics other research has tried to address.

**Table 1. Foci of Existing FGDM Research and Current Research Gaps**

<b>A. Some Foci of Existing Research</b>	<b>B. Current Research Gaps</b>
Planning and execution of family meetings	Long-term implementation and follow-through of the resulting plans by families and service providers
Participant satisfaction, including families and child welfare agency staff	The level, quality, and accessibility of available resources to meet family needs identified in plans
Impact of FGDM on placement with relatives and keeping siblings together	The implementation and impact of follow-up meetings on child safety, permanency and well-being outcomes
Impact of FGDM on disproportionality and disparate outcomes in child welfare for certain ethnic groups	How FGDM models prevent the need for out-of-home placement and serve the in-home services population

Given the research to date and the existing research gaps, this project was strategically structured to build knowledge on the effectiveness of FGDM models in maintaining child safety and preventing foster care placement for children receiving in-home services. Currently, there is insufficient research that gives child welfare systems enough evidence to support the full-scale implementation of FGDM models for children receiving in-home services, resulting in a small percentage of these children and families partnering with child welfare agencies in service planning and decision making. In rigorously evaluating practice in three jurisdictions that cover rural, suburban, and urban areas with diverse populations, it is hoped that the findings from the NPLH project enhances the knowledge base in the field.

#### D. Population

The No Place Like Home project's focus population was children and families receiving in-home child welfare services, with children at risk of entering/re-entering foster care in three child welfare agencies with mature FGDM programs: Colorado (Larimer County), South Dakota (Rapid City) and Texas (Region 3-Dallas and Tarrant Counties). Through the NPLH project, these communities sought to expand, enhance, and further saturate their FGDM models with this focus population. Collectively, these targeted geographic areas provided significant urban, rural, and suburban diversity. Individually, each geographic area experiences disparate child welfare outcomes for its respective predominate under-represented racial and ethnic group—Latinos in Larimer County; American Indians in SD; and African Americans in TX—and thus it was a goal of this project to better understand FGDM through these various cultural lenses. As described above and in the project logic model, the proposed focus population is children and families receiving in-home services whereas most FGDM evaluations have focused on children placed out of their home. The intention of focusing on the in-home services population was to help to

fill an important gap in the knowledge of FGDM effectiveness. By more narrowly focusing the scope of the evaluation on this population rather than the range of children and families who might benefit, the evaluators were striving to maintain consistency and comparability across the three sites. From the perspective of the goals of the project to advance knowledge of effective FGDM implementation and practice, this allowed for the development of an evaluation design that concentrated on processes and outcomes that pertained to this population around prevention (such as placement prevention and the prevention of maltreatment recurrence). In addition, the focus population was meant to permit the integration of intervention fidelity within sites, which was hoped to further strengthen the evaluation.

For the most part, the project was implemented as intended. The target populations remained the same in both Texas and Larimer County; however, due to lower than expected response rates in South Dakota, the target population was expanded to include the out-of-home service population and, in addition, the randomized assignment process was suspended in that site (for more information see V.C. Evaluation Plan Revisions).

## II. Overview of the Program Model

### A. Project Goals and Objectives

The primary goal of NPLH was to address existing knowledge gaps in FGDM research around children and families receiving in-home services through a comprehensive evaluation design of FGDM implementation, processes, and outcomes (all tied to costs). Key features of the outcome evaluation were a randomized control trial (RCT) in Texas, an intent-to-treat design in South Dakota, and a quasi-experimental propensity score matching design in Larimer County, Colorado. In other words, all three sites were designed to feature rigorous longitudinal and experimental or quasi-experimental designs addressing FGDM process, outcomes and cost effectiveness. In addition, given the over-representation of Latinos, African Americans and American Indians that exist in this project's targeted geographic areas, this project sought to provide guidance and evidence about the effectiveness of FGDM in supporting culturally diverse populations.

Many states have identified FGDM models as a strategy to implement in order to sufficiently conform to many CFSR items, and, more importantly, to improve their practice so that outcomes for children and families are optimized. While this project's state partners have been implementing FGDM models for many years, and local (Larimer County) and statewide (Texas) research has demonstrated positive outcomes, there was still significant room for improvement in implementation and evaluation. Table 2 shows the Areas Needing Improvement (ANI) by project state as determined by the CFSR reviewers. Given these state's maturity, tenure and investment in FGDM implementation, they were well positioned to make measureable improvements through participation in the NPLH project.

Additional goals of the No Place Like Home project included individualized, advanced training, technical assistance and coaching of the sites' staff and community; peer networking; and a highly tactical and extensive dissemination component designed to reach thousands of stakeholders so that the knowledge be gleaned from this project could be shared toward the ultimate goal of improving child welfare practices and policies at local, state, and national levels. Please see Appendix A for a copy of the project logic model which ties all project activities to goals and objectives.

**Table 2. CFSR Areas Needing Improvement within Project States**

CFSR Item	Description	State		
		CO (11/2009)	SD (4/2009)	TX (3/2009)
<b>Item 2:</b> Repeat Maltreatment	Reviewers were to determine if there had been a substantiated or indicated maltreatment report on the family during the period under review	Strength	ANI	ANI
<b>Item 3:</b> Services to family to protect child(ren) in the home and prevent removal or reentry into foster care	Reviewers assessed whether, in responding to a substantiated maltreatment report or risk of harm, the agency made diligent efforts to provide services to families to prevent placement of children in foster care while ensuring their safety	ANI	ANI	ANI
<b>Item 14:</b> Preserving Connections	Reviewers were to determine whether the agency had made, or was making, diligent efforts to preserve the child's connections to neighborhood, community, heritage, family, faith, and friends while the child was in foster care	ANI	Strength	ANI
<b>Item 15:</b> Relative Placement	Reviewers were to determine whether the agency made diligent efforts to locate and assess both maternal and paternal relatives as potential placement resources for children in foster care	ANI	ANI	ANI
<b>Item 17:</b> Needs and Services of Child, Parents and Foster Parents	Reviewers were to determine whether the agency had adequately assessed the needs of children, parents, and foster parents and provided the services necessary to meet those needs	ANI	ANI	ANI

<b>Item 18.</b> Child and Family Involvement in Case Planning	Reviewers were to determine whether parents and children (if age-appropriate) had been involved in the case planning process, and if not, whether their involvement was contrary to the child's best interest	ANI	ANI	ANI
<b>Item 25.</b> The state provides a process that ensures that each child has a written case plan to be developed jointly with the child's parent(s) that includes the required provisions		ANI	ANI	ANI

## B. NPLH Service Model

The three public child welfare agencies (Larimer County, Colorado; South Dakota Department of Social Services; and Texas CPS) that comprise this project team were intentionally selected to partner with the evaluator organizations in part due to their extensive history implementing FGDM, and more specifically FGC, over many years (13 years in Larimer County; 8 years in Texas, and 6 years in Rapid City, South Dakota) and further, had embedded FGDM models into the fabric of everyday child welfare practice. It was assumed that these agencies' staff expertise, commitment to, and competency in FGDM would result in seasoned veterans who could rapidly deploy the FGDM models, make mid-course corrections, and help analyze the results. Additionally, each of these agencies were vested in evaluating the impacts of FGDM, and were well-positioned to translate the research into meaningful practice and policy changes that will positively impact children, families and the system professionals, both locally and nationally. And finally, it was assured that FGDM would remain a core and sustained activity post-grant, and that the evaluation findings would be embedded into future plans for growth and sustainability. For these reasons, the project team was confident that the information gained from project implementation, process and outcome evaluations would make a significant contribution to the field.



Due to the variety of meeting models being implemented in each site as well as the established model fidelity associated with the Family Group Conference (FGC) model, the FGC was the primary meeting model of focus across sites for this project. Rated as a promising “3” on a 1-5 scale of level of research evidence by the California Evidence-Based Clearinghouse (2011), which indicates the need for additional research to establish this as an evidence-based practice, there was an established knowledge base to use to focus on the FGC. However, due to the low numbers of FGCs conducted with the target population in Larimer County (~20 over the life of the project) and the county’s interest in ascertaining what, if any, differences in impact the two meeting types may have on families, Family Unity Meetings (FUMs) were assessed as well in that site.

### C. Project Activities

#### Training and Technical Assistance

The first six months of the project were dedicated to working with each site’s practice staff and leadership—including administrators, program managers, supervisors, community partners, and others—to help them install and/or enhance FGDM implementation with this target population. The Kempe Center provided a significant amount of training and technical assistance in helping each site configure their case flow process between the various FGDM models that they implement. In addition, the Kempe Center conducted a training needs assessment with each public child welfare agency which resulted in a 21-month training plan for each site which was implemented in the first 2 years of the project. Depending on the needs of the site, introductory through advanced trainings and coaching sessions were provided to a variety of stakeholders, though were mostly focused on the FGDM facilitator/coordinators and referring worker staff of the public child welfare agencies. These trainings sought to effectively build skill and

competency for all professionals who have a role in the implementation of FGDM. A range of topics were covered, including domestic violence, substance abuse, mental health, managing emotions, facilitating teams, family finding, involving children and youth, managing conflict, and disproportionality and disparities. All of these trainings combined didactic, small-group, large group, case example, video, simulation, coaching and other effective learning strategies. Please see Appendix B. NPLH Training and Technical Assistance for a comprehensive list of trainings provided in each site over the life of the project.

### Peer Networking

Peer networks can improve staff retention by supporting professional development and growth, facilitating mutual learning, empowering staff, and building a sense of community among child welfare staff. Peer networks can provide opportunities for professionals to learn from one another as they build continuous quality improvement processes into their program improvement efforts. Peer networks offer an efficient and effective way to reach middle managers, supervisors, and front line staff who play a critical role in implementing systems reform. In addition, peer networks also provide opportunities for staff to develop and practice leadership skills such as mentoring and coaching. Given the likely benefits of peer networking, this was an important project activity and had multiple components. As part of the Children's Bureau T/TA network and the Family Connections Grant cluster, peer networking among this project's child welfare agencies began with bi-monthly conference calls during the early months of the grant and among all cluster grantees at the kickoff meeting held in Washington, DC in November 2011 and at the other annual grantees meeting in 2013.

Occurring quarterly throughout the grant, teleconferences were held with leadership from each site to discuss project progress, milestones, issues related to implementation and the

evaluation. With the support of the Project Coordinator, the peer network determined their quarterly agenda to meet their needs, and between teleconference calls, had the opportunity to further their relationship through emails and phone calls. In addition, in each year of the grant, each of the sites was slated to host a 2-day in-person meeting of a total of 8-10 representatives from the other sites. While Larimer County hosted the first visit in Year 1, Dallas and Tarrant Counties hosted in Year 2, South Dakota opted to hold their visit in Olmsted County, Minnesota (a 2012 Family Connection grantee) in Year 3, and all sites sent representatives to Minnesota in that final year. This format allowed for a more intensive learning exchange between the core members of the peer network, in addition to the opportunity to observe practice in each other's sites. These in-person gatherings further provided the evaluation team an opportunity to observe and partner with programmatic staff around evaluation lessons and possible interpretations of the data.

### Evaluation

As described in much greater detail in Sections V-0, the bulk of project resources were dedicated to a rigorous evaluation of FGDM practice in the three project sites at both process and outcome levels.

### Dissemination

*Webinars.* Webinar technology provides an easy-to-use tool to rapidly and broadly disseminate cutting-edge project information. It provides those engaged or interested in FGDM to learn real-time lessons and strategies around implementation and evaluation. Kempe developed and provided 5 webinars during each project year on topics related to FGDM practice and evaluation and led by well-respected experts in the field from around the country, many of whom were current or former Family Connection grantees. Each webinar was free of charge and

could accommodate the first 200 registrants. Each webinar was recorded and posted on Kempe's FGDM website ([www.FGDM.org](http://www.FGDM.org)), available to the public, and downloadable for free.

*Conference and other presentations.* Between Kempe, CFP and the three public child welfare agencies, four conference presentations were conducted over the life of the project, the bulk of which were held in year 3 and the 4<sup>th</sup> no-cost extension year, when there were an increased amount of information and results to be shared. Conferences presented at included the 19th National Conference on Child Abuse and Neglect and Kempe's annual 2012, 2014, and 2015 Conferences on FGDM and Family Engagement. Other mid-project presentations were made in person or via webinar to staff in Larimer, Dallas and Tarrant counties and Rapid City and at the headquarters of Casey Family Programs.

*E-blasts and website.* Throughout the life of the grant, Kempe sent approximately 24 e-blasts to approximately 5,000 individuals each time who have indicated an interest in FGDM. CFP also has a semi-annual research news e-blast that reaches thousands. These vehicles for dissemination were utilized throughout the life of the grant to disseminate project-related information as well as to advertise for the webinar and conference opportunities. In addition, Kempe's FGDM website houses all project reports, webinar recordings, PowerPoint presentations, journal citations (and links, if appropriate) and evaluation findings. This website receives approximately 3,000 hits per month.

*Texas DFPS Newsletters.* The Texas Project Coordinator sent out a monthly newsletter to DFPS staff, beginning April 2013 and ending in April 2014 with the conclusion of the referral period. The purpose of the newsletter was to provide immediate feedback to referring staff, provide consistent information about the project, generate referrals, and showcase practice tips and other information on FGDM implementation in Dallas and Tarrant Counties.

*Journal articles.* Thus far, project members have had four scholarly articles accepted for publication in professional journals. These include:

Merkel-Holguin, L. (2015, Winter). Family Group Decision Making as a Culturally-Responsive Child Welfare Practice. In LaLiberte, T., Crudo, T., Ombisa Skallet, H., & Day, P. (Eds.). *CW360°: Culturally responsive child welfare practice*. St. Paul, MN: Center for Advanced Studies in Child Welfare, University of Minnesota.

Williams, J., Merkel-Holguin, L., Allan, H., Maher, E., Fluke, J., and Hollinshead, D., (September 2015), Factors associated with staff perceptions of the effectiveness of family group conferences. *Journal for the Society of Social Work and Research*, 6(3), 343-366.

Merkel-Holguin, L., and Marcynyszyn, L. (2014). The Complexity of Fidelity in Measuring System Change: The Case of Family Group Decision Making. *British Journal of Social Work*, 1-13. doi: 10.1093/bjsw/bcu092.

Allan, H. & Maher, E. (2014). All in the family: Variations in the use of family meetings in child welfare. *Child Welfare*, 92(6), 97-110.

Three other manuscripts are currently under development on such topics as children's participation in FGDM processes, FGDM and social supports, and staff perceptions of the tension between child protection and family preservation.

### III. Collaboration

The three sites of the NPLH project have been implementing various types of family meetings for more than a decade.<sup>1</sup> These public child welfare agencies are grounded by family meeting policies and procedures and have an infrastructure to support family meeting practices, including dedicated staff (supervisors and coordinators), and policies that, in theory, drive referrals. Thus, the presence of collaboration at various levels reflects, in most instances, the efforts of FGDM-dedicated staff and leadership who have worked to install these practices within these child welfare systems, some of which predated the NPLH grant award.

#### A. Inner-Agency Collaboration

With regard to FGDM services in Larimer, Tarrant and Dallas Counties, there are four types of child welfare professionals who typically play a critical role in the implementation of family meetings: leadership, supervisors, child welfare caseworkers, and FGDM Coordinators.

##### Leadership

Even with policies and procedures that require FGDM to occur at specific times in a case trajectory, such is the case in these three counties, the support of leadership and administrators is important to embedding FGDM into the fabric of everyday child welfare decision-making processes. According to the Larimer County FGDM Supervisor, “Family meetings are an expectation,” and this is supported by agency leadership and management, resulting in infrastructure and resources being allocated to facilitate family meetings, and staff standardly referring families to participate in these decision-making processes. Historically, in Dallas and Tarrant Counties, when leadership created additional checks and balances around FGDM

---

<sup>1</sup> While Rapid City, South Dakota was one of the NPLH sites, this main report does not include their experiences with implementation. Rather, this can be found in Appendix G: *The Implementation of Family Group Decision Making in South Dakota during the No Place Like Home Project—Final Report*.

implementation, such as reviewing referral numbers at the unit, supervisor, or caseworker levels, the number of referrals increased. The lesson here is that policy alone is often insufficient to embed change in complex environments like child welfare agencies.

### Supervisors

Caseworker supervisors' support of FGDM processes may impact the likelihood in which their caseworkers refer to FGDM, even in instances where policy dictates a referral. In Dallas and Tarrant Counties, during the project, there were FBSS units that did not refer cases for an FGC and one barrier presented was the perception that supervisors must attend all meetings scheduled for their caseworkers, which was perceived as being an insurmountable time burden on those staff. However, this was a misconception that was dispelled with the aid of messaging from leadership and training/coaching to FBSS staff. In addition, Texas focus group respondents noted that a positive relationship between the FBSS supervisors and FGDM supervisors was important to support programmatic implementation. These supervisor teams, along with administrators, have stated that they are once again using data to isolate referral patterns among caseworkers toward the goal of increasing referrals.

### Child Welfare Caseworkers

These three counties reflect the struggles experienced nationally related to child welfare workforce recruitment and retention, resulting in levels of turnover that make it essential for FGDM staff and others to continually educate, train and seek buy-in from caseworkers to refer and perform their responsibilities in the various family meetings. Caseworkers are pivotal to the implementation of family meetings, in that they are a main referral source, and at the meeting they provide critical information that is used for decision-making purposes. In Larimer County, caseworker referrals to FSRTs and FUMs were commonplace, reflecting an expectation. On the

other hand, this was not the case for FGCs. Families offered FGCs likely have participated in other types of family meetings, and have either a caseworker or Coordinator who believes that a more resource-intensive meeting to find and prepare members of the family group will result in decision making to improve or expedite the improvement of outcomes.

During the NPLH project period in Dallas and Tarrant Counties, FGDM staff needed to market the FGDM service to FBSS caseworkers because, though policy dictates that referrals are mandatory in some instances, they are not automatic, unlike in the Conservatorship stage of service referrals. At the end of the NPLH project, focus group respondents in Dallas and Tarrant Counties noted there was a greater level of collaboration between FBSS caseworkers and FGDM staff. This is likely because of the education and training efforts that occurred throughout the NPLH project, FGDM staff's efforts to promote the FGDM service, and FBSS caseworkers' positive experiences with family meetings.

#### FGDM Coordinators

As ambassadors and implementers of the FGDM service, Coordinators are critical in delivering and sustaining FGDM practice. In all of the sites, Coordinators received training, coaching and support—either through their agency or from the Kempe Center—to continue the installation of the program as intended and with fidelity. In Dallas and Tarrant Counties, they were responsible for program marketing to their caseworker counterparts to increase referrals with the FBSS target population. In all three sites, the FGDM Supervisor coached and supported the FGDM Coordinators in implementing quality practice, with a continued emphasis on model fidelity.



## B. Inter-Agency Collaboration

There were some unique collaborative relationships between the child welfare agencies and community organizations around the implementation of FGDM in some of the sites as well. For example, in Tarrant County, New Day Services spearheads the county's Fatherhood Initiative and plays an active role in family finding, with an emphasis on fathers and paternal relatives. Through their contract with DFPS, New Day Services assists the FGDM Coordinator with widening the circle of participants in FGCs.

Larimer County has contracted with a community-based agency that has hired family advocates to attend FSRT meetings to help elevate parents' voices and to facilitate their understanding of these meetings. These family advocates attend all family meetings (including FSRTs) as they are available and as the family requests their participation. In their role, the family advocates are also positioned to challenge the agency, if necessary, and to create support teams for the family. In addition, another community agency, Turning Point, is contracted to partner with an internal DHS staff person to supplement the FGDM Coordinators' family finding activities.

## IV. Sustainability

The NPLH project provided the three project counties with the opportunity to objectively review data related to FGDM implementation. While all counties intend to sustain FGDM, it is hoped that the evaluative results will help inform and nuance continued implementation efforts.

In addition to the internal infrastructure that Larimer County has built since 1997 to install family engagement practices in its agency, there are external supports that solidify this direction. Three examples are illustrative. First, family engagement is now embedded in Colorado's Children's Code (Volume 7) which is the Colorado Department of Human Services Rules and Regulations. Second family engagement is a core part of the Title IV-E Waiver project whereby states enter into agreements with the federal government to reinvest unspent or saved foster care dollars to prevention and in-home services efforts. Lastly, family engagement is realized through the Senate Bill 1451 (Colorado Collaborative Management Program where the county departments of human/social services and other mandatory agencies (judicial, public health, probation; schools) develop multi-agency services to provide to children and families. In addition, Larimer County will learn from installing family advocates into their planning activities, and from the continued research conducted by the Social Work Research Center at Colorado State University.

In Dallas County, in addition to policy guiding and sustaining implementation, it appears that when program administrators encourage FGDM processes and require FTMs before they will approve an emergency removal, family meeting referrals across all stages of service will increase. In addition, presentations at unit meetings, reviewing administrative data to promote data-driven decision making, and simplifying the referral process at the FBSS stage of service are some of the other sustainability strategies that have been implemented.

In Tarrant County, as in Dallas County where policy guides referrals, sustainability efforts are supported by the FGDM State Lead at the Texas Department of Family and Protective Services. The FGDM Supervisor will continue to support growth and quality of practice amongst the FGDM Coordinators. In addition, as part of Foster Care Redesign in Texas, DFPS is contracting with All Church Homes, Child and Family Services (ACH) to provide a continuum of foster care placement and services to children and families from 7 counties in the southwest portion of DFPS Region 3 (known as Region 3b and includes Tarrant County). Beginning in September 2014, ACH began using FGCs as part of their service planning requirements. Region 3b's (including Tarrant County) collaboration with ACH is supporting the growth of FGC practice in the community agency.

In general, across Texas, additional FTEs in the FGDM service area, dedicated to engaging young people 14 and older in transition planning through Circles of Support meeting processes, will also result in further cementing of the service in the DFPS fabric of services.

## V. Evaluation Planning

The evaluation for the No Place Like Home project was a major aspect of the project and a large part of the reason practice sites with well-established practice were selected as it was hoped that such practice strength would facilitate the implementation of rigorous evaluation designs in each site, each appropriate to their particular context. Experimental or quasi-experimental designs were planned for the outcome evaluation in each of the three sites: a randomized control (RCT) trial in Texas; an intent to treat design with randomly assigned controls in South Dakota; and a propensity score matching design for Larimer County, Colorado. In addition, a process and cost evaluation was planned for each of the three sites with cross-site analysis intended, as appropriate, to discern commonalities in practice and outcomes across the three practice sites.

### A. Process Evaluation Methodology

Process evaluations focus on *how* something happened rather than the results obtained. These evaluations are particularly useful for dissemination and replication of interventions because practice factors may help explain how and why an intervention is effective, and for whom, in ways that a purely outcome-focused evaluation cannot. The purpose of this process evaluation was to shed light on the mechanisms underlying any hypothesized effects as well as to provide helpful information to the three NPLH practice sites and other jurisdictions about implementation of FGDM. Ultimately, based on data sources, the following process-oriented questions (as well as others) were addressed through this evaluation: the types of family meetings public child welfare agencies use and the general flow of those meetings; the characteristics of families who participate in family meetings, including risk and protective factors; the barriers to implementation and uptake of FGDM exist in the agency or community and how are those addressed in developing and maintaining the FGDM models; the relationship

between worker characteristics (demographics, background, years of experiences, attitudes) and their perceptions on the effectiveness of FGCS; the length of time between referral and the actual meeting; a description of who is invited and who attends family meetings; whether the type of family meeting results in a different number and composition of meeting participants; the extent of fidelity to the core components of FGDM based on type of respondent; and worker perception of social support for families who receive family meetings versus those who do not.

### Process Evaluation Design

The process evaluation design included two features, which were closely linked to the research questions. First, a global or agency-level assessment of how FGDM was designed and implemented by the agency. This assessment included the barriers and strategies for FGDM uptake, and how these have changed over time as the project progressed, with a strong focus on fidelity. Second, a more detailed view of processes at the case-level in conjunction with the outcome evaluation were planned to understand how other factors contribute to the success of FGDM. Key to the process evaluation was determining how the agency- and case-level views are actually integrated, the degree to which policy is consistent with practice, and how agencies identify and resolve implementation problems. The process evaluation involved a combination of qualitative (focus groups and interviews) and quantitative (staff and caregiver surveys; tracking sheet) data, which are described in more detail later.

*FGDM Fidelity Analysis Methodology.* Fidelity data were collected using a survey administered at the time of the family meeting, with a section of pre- and post-meeting questions. One version was completed by the meeting coordinator, while another version was disseminated to all participants, whether they identified as family, like-family or service providers. Thus, the fidelity analysis is based on a series of questions, many which are common to both surveys that

employed a six point Likert scale where the answer options were: Strongly Disagree (1); Disagree (2); Slightly Disagree (3); Slightly Agree (4); Agree (5) or Strongly Agree (6). All questions were written so that agreement indicated that the meeting had greater alignment with fidelity principles.

Three methods were used to examine the results: computation of average Likert Scores and the calculation of Fidelity Index (FI) and fidelity domain scores (described further below) for all of the eligible family meetings part of this project. For both types of scores (Likert and FI), prior to computation, some questions were eliminated from the calculations because they were too procedurally based (e.g., the facilitator described his or her role; the roles of the family members and service providers were clearly described; the process for calling another family meeting was described). For some questions on the Participant Fidelity Survey, only the responses from the “family/like-family” respondents were used in the FI calculations. Table 4, below, reflects the final set of questions used to assess fidelity. Likert, FI, and domain scores were computed for three groups of participant types: family and like-family, professionals/service providers (see Table 3), and FGDM facilitator/coordinators, in order to examine whether or not respondents scored meeting fidelity similarly. Asterisks indicate that only the family/like-family respondents’ answers to these questions were used in score calculations.

**Table 3: Categorical Breakdown of Family and Like-Family Respondents**

<b>Family</b>	<b>Like-Family</b>
Child	Mom’s significant other
Mom	Dad’s significant other
Dad	Neighbors
Stepmom	Family friends
Stepdad	Clergy
Sibling	Godmother/Godfather
Maternal aunt/uncle/cousin	
Maternal grandparent	
Paternal aunt/uncle/cousin	
Paternal grandparent	

First, average Likert scores were computed for each meeting using data from participant and facilitator surveys. This approach enabled consideration of the average level of agreement between respondents regarding whether the meeting incorporated fidelity-related elements. For example, using the values of the response options (described above) a score of 5 or above indicates that, on average, respondents generally agreed that the elements inquired about were present in the meeting they attended.

Second, to reflect overall meeting fidelity an FI score was calculated using all of the questions reflected in Table 4 below. For all of the indices, the Likert scale responses were recoded to form a dichotomous variable indicating the presence or absence of fidelity. Thus, answers of “Agree” or “Strongly Agree,” indicate the respondent endorsed fidelity (and that question gets a score of 1) while “Strongly Disagree,” “Disagree,” “Slightly Disagree,” and “Slightly Agree” indicate lack of fidelity (and that question gets a score of 0). As the FI score reflects the average degree of the fidelity, or lack thereof, across all questions, it *represents the percentage of questions for which respondents indicated fidelity was present for each meeting*. For example, an FI score of 0.50 indicates a respondent Agreed or Strongly Agreed with 50% of the available questions, which would indicate relatively weak fidelity. An FI score of 0.75 or greater indicates a moderate adherence to fidelity for a given meeting where scores closer to or equaling 1.00 indicate strong or complete adherence.

Finally, in addition to the overall FI score, four domains of fidelity were identified: *Preparedness, Inclusion and Respect, Family Leadership, and Transparent Planning*. These four domains mirror the core principles and practices of family group decision making as established by the National Center on Family Group Decision Making (2012). The questions from the participant and coordinator/facilitator fidelity surveys were fit into these domains and are

provided below in Table 4. The relationship between the FI score and the fidelity domain scores is that the FI score utilizes answers to all of the questions underlying the four domains combined, to reflect an overall indicator of model fidelity.

**Table 4. FGDM Fidelity Domains and Corresponding Survey Questions**

<b>Domain</b>	<b>Participant Survey Questions</b>	<b>Facilitator Survey Questions</b>
<b>Family Leadership</b>	<ul style="list-style-type: none"> <li>• The right people were at the family meeting</li> <li>• The facilitator was flexible in meeting the needs of the participants</li> <li>• There was the chance to ask questions about the information presented by child welfare and other professionals</li> <li>• Paid professionals did not tell the family how to solve CPS’s concerns.</li> <li>• Others listened to my opinions about what was best for the child</li> <li>• My opinions were included in the plan.</li> <li>• The child’s ideas or needs were considered in the plan.</li> </ul>	<ul style="list-style-type: none"> <li>• I asked the family if they had any questions or needed any clarification about the information presented by the professionals</li> <li>• I believe the agency was open to the family’s ideas and decision making abilities</li> <li>• The child’s ideas or needs were considered in the plan</li> <li>• I believe the child welfare agency decision makers did not have a predetermined outcome for this family meeting</li> </ul>
<b>Inclusion and Respect</b>	<ul style="list-style-type: none"> <li>• Members of Mom’s side of the family were invited to the family meeting.*</li> <li>• Members of Dad’s side of the family were invited to the family meeting.*</li> <li>• Professionals were invited to the family meeting.*</li> <li>• Other people who feel “like family” (neighbors, friends) were invited to the family meeting.*</li> <li>• I helped determine when and where the family meeting would be held.*</li> <li>• I felt safe at the family meeting.</li> <li>• The facilitator respected me.</li> <li>• CPS staff respected me.*</li> </ul>	<ul style="list-style-type: none"> <li>• Ratio of family and like family contacted compared to number of service providers and others contacted</li> <li>• Ratio of family and like family who attended compared to number of service providers and others who attended</li> <li>• I felt safe at the family meeting</li> <li>• The child welfare agency staff were respectful to the family during the meeting</li> </ul>
<b>Preparedness</b>	<ul style="list-style-type: none"> <li>• I understand Child Protective Services (CPS)’s concerns about the child.</li> <li>• I understand the purpose of the family meeting.</li> <li>• I feel prepared by the facilitator to attend the family meeting.</li> </ul>	<ul style="list-style-type: none"> <li>• I had a clear understanding of the agency’s concerns about the child.</li> <li>• I had a clear understanding of the purpose of the family meeting</li> </ul>
<b>Transparent Planning</b>	<ul style="list-style-type: none"> <li>• CPS staff clearly told us CPS’s concerns that the plan would need to address.*</li> <li>• The plan includes things for family members to do.</li> <li>• The plan includes things for CPS to do.</li> <li>• The plan includes a way to know if everyone is doing their part.</li> </ul>	<ul style="list-style-type: none"> <li>• The child welfare agency staff clearly told all participants the agency’s concerns that the plan would need to address.</li> <li>• The plan had things for the family members to do</li> <li>• The plan had things for the child welfare agency to do</li> </ul>



Domain	Participant Survey Questions	Facilitator Survey Questions
	<ul style="list-style-type: none"> <li>• The facilitator worked with the family and CPS to reach a final plan that all could agree on.</li> <li>• The plan was fully approved at the family meeting.</li> </ul>	<ul style="list-style-type: none"> <li>• I facilitated discussions between the family and the agency for them to reach consensus on the plan</li> <li>• Was the plan fully approved at the meeting? (Y/N)</li> </ul>

\* Family and like-family participant responses only

### Process Evaluation Participants

As aforementioned, the target population for the NPLH project were families receiving in-home child welfare services in each of the three project sites who were referred for an FGC (or a FUM in Larimer County). Members of the target population were included in the process evaluation in addition to referring child welfare workers, FGDM coordinators, supervisors of both groups, and child welfare managers and administrators.

### Process Evaluation Data Sources and Collection

For the process evaluation, the analytic plan was the same across the sites. This differs from the outcome analysis which employed different statistical methods to maximize each sites' strengths in using a RCT or comparison group design. It is important to note that there is overlap between measures used in the process and outcome evaluations.

**In regards to qualitative data, the project relied heavily on focus groups and key interviews conducted throughout the life of the project. For all focus groups and notes were taken and discussions audio recorded. Content analysis techniques were used to capture themes that addressed each research question. Please see**

Table 5 below for a schedule of the focus groups that were conducted in each site.

**Table 5. NPLH Focus Group Schedule**

<b>Site</b>	<b>Date</b>	<b>Target Audience</b>
Texas	January 2012	9 groups of caseworkers, facilitators, supervisors, administrators, and family members
South Dakota	December 2011	6 groups of caseworkers, facilitators, supervisors, and family members
Larimer County, CO	January 2012	7 groups of caseworkers, facilitators, and supervisors
South Dakota	March 2013	3 groups of caseworkers, supervisors, and administrators
Texas	March 2013	1 group of administrators
South Dakota	November 2014	3 groups of caseworkers, facilitators, and supervisors
Texas	January 2015	7 groups of caseworkers, facilitators, and supervisors

In addition to qualitative data, descriptive data analyses from various surveys informed the process evaluation as well. In particular, the NPLH Fidelity Surveys were key to the process evaluation in both Texas and Larimer County, CO.

## **B. Outcome Evaluation Methodology**

The planned outcome evaluation for the study consisted of an experimental or quasi-experimental longitudinal design, specific to the practice site, and covering a minimum of 18 months of potential follow-up. The data collection period began in October 2012 and referrals to the project ceased on April 30, 2014 in Larimer County, CO and June 30, 2014 in Texas. A no-cost extension was sought and approved to allow for the maximum data collection period as well as to allow for data analysis. The final set of administrative data (e.g. SACWIS data) was pulled on these two groups to track services and outcomes, such as maltreatment recurrence and placements, in June 2015. During the ~24 month data collection period, interim results were used for formative work with the sites and to inform technical assistance and training efforts. The final six months of the no cost extension project period were be used to consolidate data, complete the

analysis, and prepare the evaluation report. The core research questions of the NPLH outcome evaluation were:

1. Are children in families in the focus population who experience FGDM interventions less likely to experience placement compared to children in the control group?
2. If children are placed out of home, are they more likely to be placed with relatives compared to the control group?
3. Are families in the population who experience FGDM interventions as likely as families in the control group to experience child maltreatment re-reports or re-reports with substantiation?
4. Are families who experience FGDM processes more satisfied with their experiences with child welfare compared to children in the control group?
5. For all of the outcomes identified above (placement, relative placement, reporting, and satisfaction) are families less likely to have disparate experiences based on race or ethnicity compared to families in the control group?

### Outcome Evaluation Design

Design for the outcome evaluation refers to the quasi-experimental and experimental approach to the study in each project site. The ability to implement a control or comparison in each of the sites represents an important evaluation capacity for the project and has the potential to address key questions concerning FGDM model effectiveness. The planned designs for each site are described below, however, it should be noted that adjustments to the design, particularly in South Dakota were necessary once data collection began and will be described in greater detail in Section V.C. Evaluation Plan Revisions. Texas, Region 3 agreed to participate in a randomized control trial (RCT). Rapid City, SD and Larimer County, CO agreed to participate in quasi-experimental studies, an intent-to-treat and propensity score match design, respectively. Described below, each site's preliminary design had both common and unique features with respect to populations and assignment.

### Outcome Evaluation Data Sources and Collection

The data collection protocol for the NPLH project was extensive and varied by site. Data for the evaluation was derived from multiple sources including focus groups, interviews,

administrative data, and surveys. Local evaluation support was used to help address evaluation implementation and data quality concerns by site. The text below describes the data sources that were collected, the instruments or data collection approach, and the timing. Please see Appendix C. NPLH Instrument List with References for a full reference list of instrument sources and rationale for inclusion.

*Focus Groups and Key Stakeholder Interviews.* Semi-structured focus groups with FGDM staff were conducted in all three years of the project to understand aspects and/or strategies for FGDM meeting preparation and family engagement as well as to understand how agencies handled barriers to implementing FGDM meetings. Key information was extracted from transcribed focus group audio recordings and included herein to provide contextual information to support conclusions and findings.

*Surveys.* A range of survey data was collected from various study participants to answer process and outcome evaluation research questions. For any survey administered, there were three key principles that were followed to ensure high data quality. The first was that the survey data must be linkable to the administrative data. The second was that, to the extent possible, any survey administered would not be overly burdensome to agency staff study participants. To this end, the determination of what survey instruments to utilize for the project was done through extensive conversations with the sites to ensure that the data obtained via any given proposed survey could not be obtained by other means to minimize respondent burden and avoid duplicative documentation. The third principle was the data collected regarding family characteristics would be an aspect of the assessment process for the study families and would apply to both the intervention and control groups so that the evaluation team would investigate

the question of whether certain case characteristics or family demographics impacted the effectiveness of FGDM.

**General Staff Survey.** The General Staff Survey was administered to referring case workers, FGDM staff, and supervisors of both at the time that staff received the mandatory NPLH Evaluation Training prior to the beginning of data collection. It was used to obtain demographic information for those staff that were expected to have any involvement with NPLH family meetings. This instrument was adapted from one already in use in Larimer County for the evaluation of the county's differential response initiative. Many of the items and scales were validated based on studies of worker and supervisory decision making and disparities conducted in Texas. Scales pertaining to skills, tenure, perceptions of family meeting usefulness and effectiveness were included. In addition, an organizational climate and culture scale was included to better understand how attitudes, beliefs, and experiences may influence relations between the intervention and the outcomes. Finally, questions about job satisfaction and the relationship between it and family meeting experiences were also included.

**Caregiver Survey Pre and Posttest.** The Caregiver Survey consisted of demographic questions, including those around economic hardship and child disability, the Protective Factors Survey, and child behavior questions to gather a more comprehensive understanding of the characteristics of families being referred for family meetings, both in the treatment and control group. The survey was meant to be administered as a pre and posttest in order to assess change over time, particularly in reference to the protective factors and child behavior. In addition, the posttest included questions about service provision and satisfaction with child welfare services.

Beginning in July 2013, incentives of \$20 Target gift cards were offered to caregivers in the Texas site for returning a Caregiver Pre or Posttest Survey to the evaluation team.

**Participant and Facilitator Fidelity Surveys.** To better understand the conditions under which FGDM affects child outcomes, it was important to examine the characteristics of the meetings and the degree to which fidelity was achieved. To this end, three versions of a fidelity tool were developed. The three versions are: 1) Coordinator/Facilitator Fidelity, 2) Participant Fidelity, and 3) Follow-Up Participant Fidelity. Additionally, satisfaction questions were included in the follow-up version of this survey to assess family members' degree of satisfaction with CPS services, in addition to their retrospective reflections of their experiences with the family meeting.

The Coordinator/Facilitator Fidelity survey was completed online by the Coordinator at the conclusion of the family meeting. Coordinators were asked to capture their perspectives about the preparatory activities on a hardcopy of the survey, to later be transferred into the Survey Monkey online platform when they completed the balance of the survey post-family meeting. The Participant Fidelity survey was introduced by FGDM facilitators at the start of each meeting, and all participants in attendance were asked to complete the first page which captured their experiences with the preparation process and some basic demographic information. At the conclusion of the family meeting, all participants were asked to complete the second page of the survey. A large envelope was available for participants to securely place their surveys. The envelope was sealed and mailed to the Kempe Center. Lastly, the Follow-Up FGDM Fidelity survey was mailed to Texas participants who provided their contact information, which was an optional field on the participant fidelity survey.

**Case-Specific Questionnaire.** When a case closed or was transferred to out-of-home services, the lead caseworker received an email with a Survey Monkey link and a request to complete the case-specific questionnaire. The questionnaire captured information regarding what services

were utilized by families, whether service provision occurred as a result of participation in a family meeting, and the caseworkers' perceptions of the families' level of social support.

*Meeting Logs.* The NPLH Meeting Logs consisted of site-specific Excel spreadsheets, housed on a secure SharePoint site hosted by the evaluation team that were designed to provide live tracking of case assignment to the treatment and control group for all NPLH referrals. The spreadsheets contained: contact information for the primary caregiver(s) to aid in facilitator outreach and information around family meeting(s) received. These documents allowed for real-time communication between the evaluation team and FGDM program staff about who is being referred to and receiving the enhanced-dosage FGDM model and aided in data tracking and QA efforts.

*Administrative Data.* Administrative data from the Colorado and Texas SACWIS systems were used to answer many of the questions related to the outcome evaluation for both treatment and control groups. In addition, the administrative data was used for the propensity score matching process in Larimer County. The evaluation team worked with designated IT staff in Texas and Colorado to define the specifications for the extract files which included re-referrals, out-of-home placements, and other key outcomes as well as demographic information. Other extract data associated with the events included event dates, and identifiers for children, families, cases, and households.

*Cost Data.* Cost data were collected via MS Excel spreadsheets that were completed by Texas and Larimer County, the two sites for whom the cost evaluation was executed. Costs were computed primarily based on a functional analysis of staffing resources and an estimation of indirect (non-personnel) costs. Direct costs attributable to project FGCs were tracked in the site-specific Meeting Logs while direct paid service costs, where available, are obtained from



administrative SACWIS data based on expenditure records. Ultimately, inconsistent tracking in the meeting logs resulted in unusable estimates of direct costs attributable to FGCs; however, these costs were deemed marginal and included in the broader, non-personnel cost accounting.

Agency staff costs were estimated based on the agency budgeting model and represent the average cost per position type including fringe benefits. Other overheads specific to the FGC program was obtained from the site budgetary models or drawing from existing research, where needed. Casework staff time estimates were derived from workload studies or site-specific estimates of workload and were used to provide an average time per in-home service per month, per case. An assumption was made that this time did not vary. Finally, meeting time was estimated from the Facilitator Fidelity Survey as well as the Meeting Logs. Additional time for various agency staff participating in the meetings was included; agency staff costs were assigned based on the position of the participating staff and the average staff cost per position as determined from the site's budget model.

#### Data Collection Targets and Response Rates

Based on preliminary discussions with the site project leads, targets for referrals to the project were determined based on the size of the county and their capacity to provide FGDM services to the target population. In addition, the evaluation team determined target response rates for each piece of survey data collected. Table 6, below, indicates the referral and response rates for each piece of survey data collected in the Colorado and Texas sites. It should be noted that depending on the survey, the denominator for the response rate varied. For Caregiver Pre-test data rates were calculated using the number of study referrals. For Fidelity data, the total number of meetings held was used. For Caregiver Posttest data, the denominator employed was the total number of pretest surveys received. And finally, for the Case-Specific Questionnaire the

number of closed in-home cases was used. Due to extremely low referral rates in the South Dakota site, the study design was shifted to that of a process evaluation only in that site and data collection for a number of survey pieces was suspended (for more information see section V.C. Evaluation Plan Revisions).

**Table 6. NPLH Referrals and Response Rates**

<b>Data Source</b>	<b>N level</b>	<b>Dallas County, TX</b>	<b>Tarrant County, TX</b>	<b>Larimer County, CO</b>
<b>Referrals</b>	Case-level n	270	275	541
<b>Caregiver Survey Pretest</b>	Case-level n (response rate)	150 (57%)	105 (38%)	35*
<b>Participant Fidelity</b>	Meeting-level n (response rate)	73 (97%)	83 (70%)	145** (26%)
<b>Facilitator Fidelity</b>	Meeting-level n (response rate)	57 (77%)	96 (86%)	154** (28%)
<b>Caregiver Survey Posttest</b>	Case-level n (response rate)	46 (30%)	29 (28%)	n/a*
<b>Case-Specific Questionnaire</b>	Case-level n (response rate)	140 (64%)	171 (67%)	128 (34%)

\*Data collection for the Caregiver Survey was suspended in Larimer County as of Nov. 1, 2013 (for more information see Section V.C. Evaluation Plan Revisions). As a result, Caregiver Posttest data was also not collected in that site.

\*\*These data were collected from both FUMs and FGCs conducted with the study population during the project period. Fidelity data were collected on up to 4 FUMs and up to 2 FGCs, resulting in up to a maximum of 6 fidelity surveys being administered per case. Please see Figure 1 for a description of meeting sequences that study families received in Larimer County.

### C. Evaluation Plan Revisions

#### South Dakota

Despite anticipating comparatively low numbers of referrals in Rapid City due to the small size of the county, targets were not met in that site. Within the first 6-months of data collection, in consultation with the site leads, site-specific changes to the evaluation design were made in the hopes of increasing referrals. First, an additional NPLH study population was added in an attempt to increase study referral rates: out-of-home service cases where reunification within 6 months has been deemed likely. In addition, the randomizer was suspended to increase the

number of families receiving FGCs. As a result, there were two intervention groups and no control group under the revised study design, which shifted the focus of the evaluation to that of a formative evaluation more so than an outcome evaluation. It was hoped that these changes would generate not only an increase in study participants but a greater degree of data around fidelity, in particular, a key focus of the NPLH cross-site evaluation. However, in total, over the life of the project only 20 families were referred to the NPLH project (e.g. referred for an FGC in either of the 2 target populations). The extremely small *n* precluded an outcome evaluation in South Dakota, however, in an attempt to understand the reasons for the low referrals, rich process data was gathered in the site. Please see Appendix G for a detailed process report for the South Dakota site.

#### Larimer County, CO

Due to low Caregiver Survey response rates (18%) and challenges reconciling data such that many surveys received appeared to be from non-study participants while simultaneously missing for confirmed study participants, the decision was made to drop this survey from the data collection protocol for the Larimer County site in November 2013. Prior to this decision, an extensive follow-up effort was implemented by the evaluation team involving six phone calls to caregivers in an attempt to solicit a survey response. This effort was largely unsuccessful despite being highly resource intensive. As this survey was intended to be used as a source of matching data for the Propensity Score Match design this shortcoming resulted in a greater reliance on administrative (SACWIS) data to conduct the Propensity Score Match. In addition, data quality issues were endemic to this site's data collection efforts, in part due to the heavy reliance on administrative data to track study participants, as the site opted against dedicating site staff resources towards manually completing and maintaining a Meeting Log. As such, the evaluation

team implemented a quality assurance process for Larimer County in an attempt to mimic the Meeting Log document by creating a “Verified Participant Log” using various administrative data extracts and reports provided by Larimer County. This process provided a solution to the ongoing data quality issues experienced in this site and enhanced the ability to successfully reconcile data, which strengthened the evaluation team’s ability to report meaningful and valid findings.

#### D. Evaluation Trainings and Engagement

Leading up to the launch of the evaluation in late Fall 2012, in-person visits were made to each of the three sites by evaluation staff to provide live evaluation trainings for all referring workers, facilitators, and supervisors. These trainings covered a project overview, goals of the evaluation and the specific tasks required of any staff member who might be touched by the project, namely referring caseworkers and FGDM facilitators. Most significantly, staff roles in survey data collection/administration and data entry in the Meeting Log tracking documents were emphasized. At this time, the General Staff Survey was administered to all staff in attendance as well. Over the life of the project, as new staff were on boarded to their respective agencies, or as staff transferred into new roles within their respective agencies that necessitated involvement with the NPLH project (e.g. referring workers and facilitators), evaluation staff at the Kempe Center offered evaluation training webinars to orient and train new staff to the project and their role in it. At these times, the General Staff Survey was also administered to those new staff.

## VI. Process Evaluation Results

### A. General Child Welfare Staff Survey

The successful implementation of an intervention in child welfare often depends on agency ownership, confidence in intervention effectiveness, and the ability to implement with fidelity. Because child welfare worker attitudes can impact referral rates to and participation in FGCs, results from the General Child Welfare Staff Survey (GSS) were used to assess child welfare professionals' perspectives about the effectiveness of this intervention. These perceptions are important because, beyond referral and participation, they also reflect organizational support of FGCs and implementation fidelity. Further, the GSS was used to assess staff orientation toward family preservation or child safety, which may have implications for decision-making within an agency. (Full descriptive results from the GSS are available in Appendix D). The sample of staff surveyed consisted of any staff that had a role in the NPLH evaluation, their supervisors, and FGC coordinators and trainers. In all, 301 staff members responded to the survey (58 percent from Dallas and Tarrant Counties in Texas; 33 percent from Larimer County, Colorado; and 9 percent from Rapid City, South Dakota).

*Perceptions of Effectiveness.* General Staff Surveys representing staff from across the three study sites were examined to determine whether staff characteristics were associated with perceptions about the effectiveness of family meetings. Using Bayesian Model Averaging (described in depth in Williams et al., 2015) techniques, factors considered included: job responsibility, years of experience, worker ethnicity, job satisfaction, orientation toward child safety versus family preservation, workload, supervisor competence, leadership, organizational culture, perceptions of community services, perceptions of families, and site. Ultimately, as shown in Appendix H, the model suggested that of those factors considered, respondents who

carried a caseload who also reported more confidence in service availability were more likely to perceive family meetings as effective (compared to caseload carriers who did not have faith in available services), while among those respondents who did not carry a caseload, those who believed that families had an ability to develop a case plan expressed greater confidence regarding the effectiveness of family meetings compared to those who did not trust in families' capabilities to develop plans.

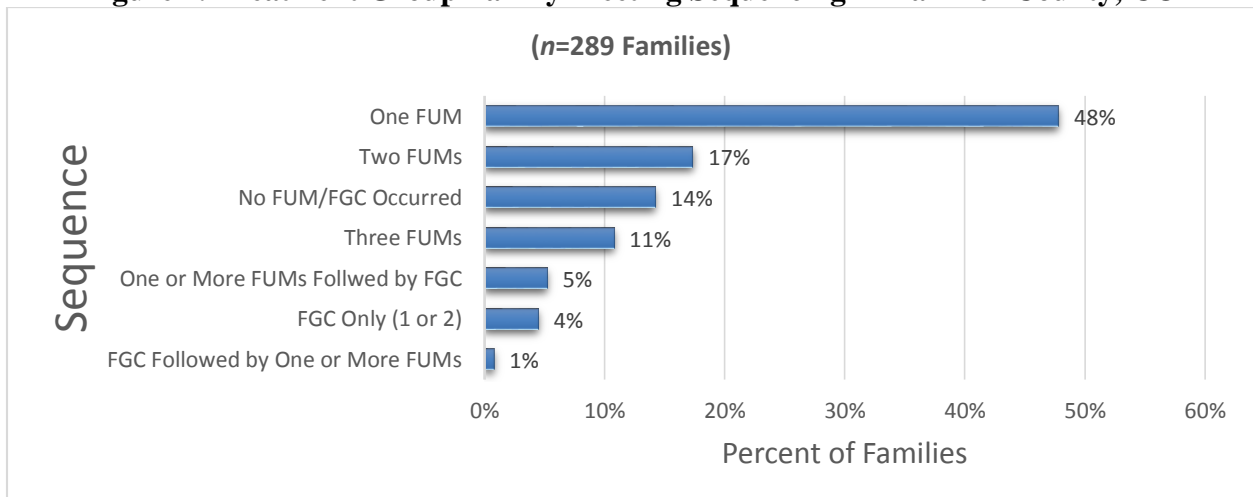
*Proclivity toward Child Safety or Family Preservation.* Another analysis involved using regression techniques to analyze what predicts a worker's orientation toward child safety versus family preservation. Factors examined included worker age, race, gender, education, job title, caseload, work experience, job satisfaction, confidence in service providers, and organizational climate and culture, which was measured using ratings of agency leadership, supervision, and shared vision with fellow workers in a staff member's unit. The results suggested that fewer years of experience working in the field of child welfare and carrying a caseload (often associated with front-line staff who do home visits) were associated with a stronger child safety orientation. Conversely, staff members in a supervisory role, those that have been in the field longer, or those who may no longer work directly with families, had more of a family preservation orientation. The results also suggested that a higher rating of shared vision with workers in one's unit was also related to greater likelihood of orientation toward child safety. The other measures of organizational climate and culture, as well as the other factors examined (i.e., race, gender, education, job satisfaction, and confidence in service providers), were not found to be significantly associated with staff orientation.

## B. Fidelity Results - Larimer County

### Description of Larimer County Family Unity Meetings and Family Group Conferences

*Number, Type, and Sequence of Meetings.* Treatment group families involved in the Larimer County site could participate in one or more FUM or FGC meetings during the course of the study and the sequence of the types of meetings experienced varied across families. The sequencing of the meetings in Larimer is reflected below in in Figure 1. The Larimer treatment group study sample was largely composed of families receiving one or more FUMs (82%); those receiving one or more FGCs only constituted a mere 5% of the final sample. In addition, 14% of treatment group families ( $n = 41$ ) who were referred for a FUM or FGC did not participate in either meeting type during the study period. About one third of families participated in more than one meeting (34%,  $n = 98$ ).

**Figure 1. Treatment Group Family Meeting Sequencing in Larimer County, CO**



### *Length of Time between Initial Referral and First FUM or FGC*

The median number of days between date of initial referral and date of first FUM was 31 days. The mean number of days was much greater, 68 days (SD: 90.06), because it took more than 90 days in 34% of cases ( $n = 87$ ) before the first FUM was held. The minimum number of

days to date of first FUM was 0; the maximum was 531 days. The median number of days between date of initial referral and date of first FGC was 39 days, with a mean of 60 days (SD: 44.62). The minimum number of days to date of first FGC was 14; the maximum was 188 days. It is important to note that it is highly likely that most, if not all, of these families received an FSRT (not included in data collection) before they participated in a FUM or FGC.

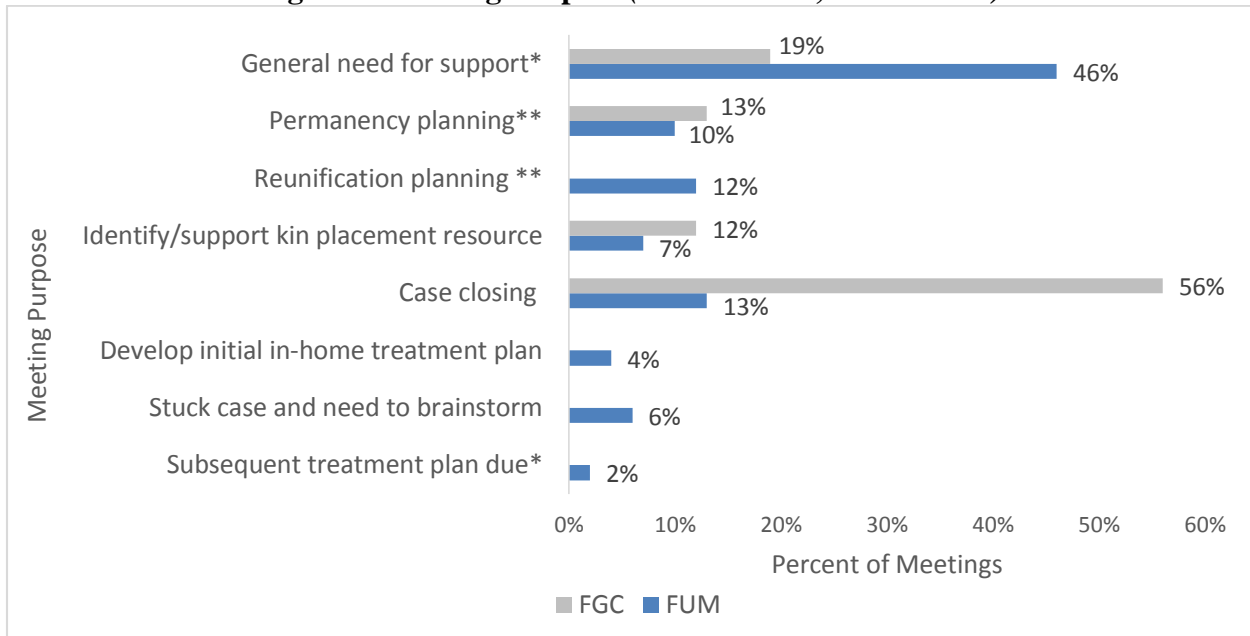
### Larimer County Facilitator Fidelity Results

A total of 13 facilitators provided fidelity data for FUM meetings, whereas a total of 6 facilitators provided fidelity data for FGC meetings. Facilitators of FUMs held a median of 8 meetings each (minimum: 1 meeting; maximum: 33 meetings). Facilitators of FGCs held a median of 2 meetings each (minimum: 1 meeting; maximum: 8 meetings).

*Meeting Purpose.* Facilitators were asked to indicate the reason for the family meetings. The results are shown in Figure 2. The most frequent reason for having a FUM was a general need for support (46%), followed by case closing (13%), reunification planning (12%), and permanency planning (10%). In contrast, the most frequent reason for having an FGC was that the case was closing (56%), followed by a general need for support (19%), permanency planning (13%), and identifying or supporting a kin placement resource (12%). None of the 17 FGCs were held to facilitate reunification planning and, in general, case planning efforts were conducted in FUMs, not FGCs.



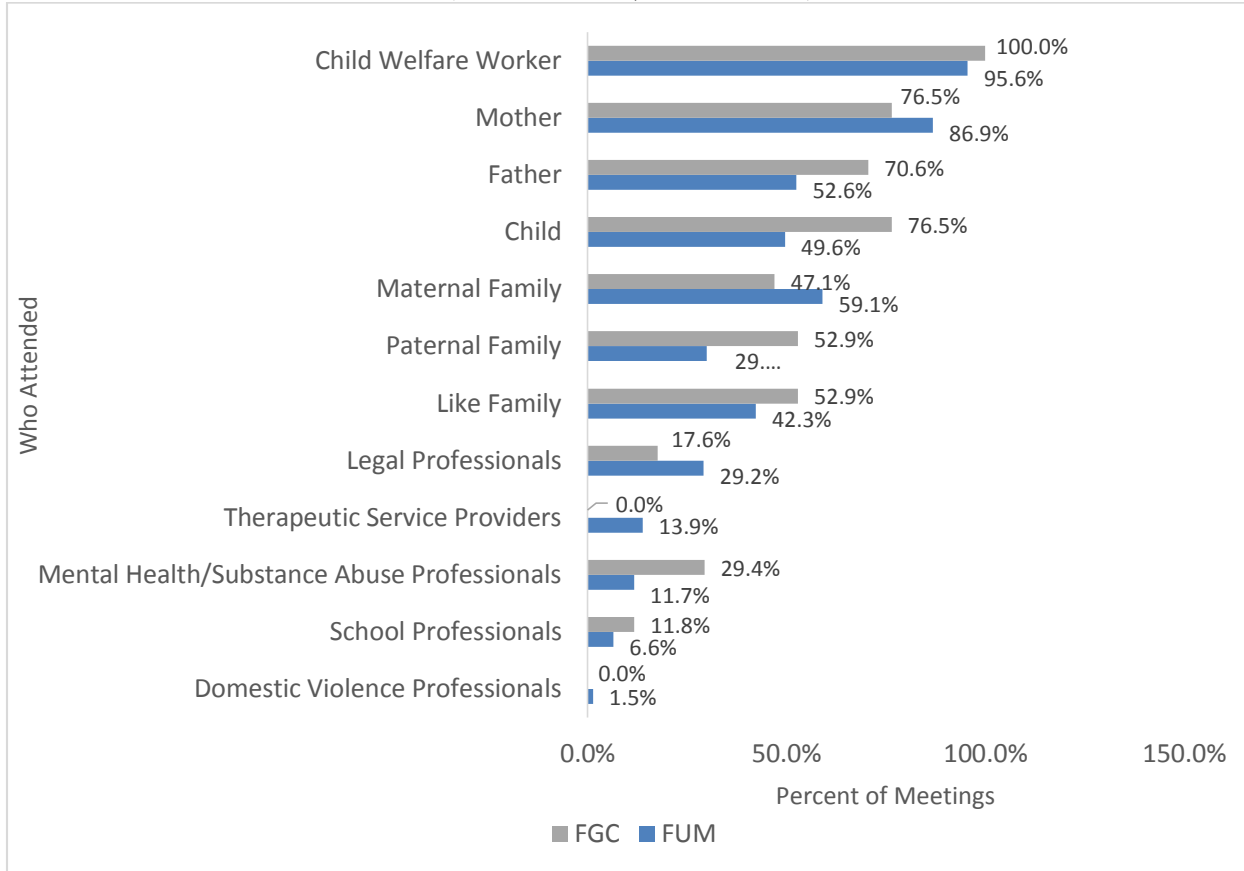
**Figure 2. Meeting Purpose (n=137 FUMs; n=17 FGCs)**



\* Child in home; \*\*child in out-of-home care

*Meeting Attendance.* Facilitators reported that a median of 6 people attended FUM meetings (minimum: 1; maximum: 18), whereas a median of 7 people attended FGCs (minimum: 6; maximum: 13). As shown in Figure 3, child welfare workers attended 100% of FGCs and 96% of FUMs. Mothers also typically attended meetings, with attendance rates of 77% for FUMs and 87% for FGCs. Fathers attended around 71% of FGCs and 53% of FUMs. Maternal family members were more likely to attend FUMs than FGCs (attendance rates of 59% and 47%, respectively), whereas the reverse was true for paternal family members (attendance rates of 30% and 53% for FUMs and FGCs, respectively). Professionals other than child welfare workers attended FUMs or FGCs 30% or less of the time, with mental health/substance abuse professionals showing the highest rates of attendance in FGCs (29%) and legal professionals showing the highest rates of attendance in FUMs (29%).

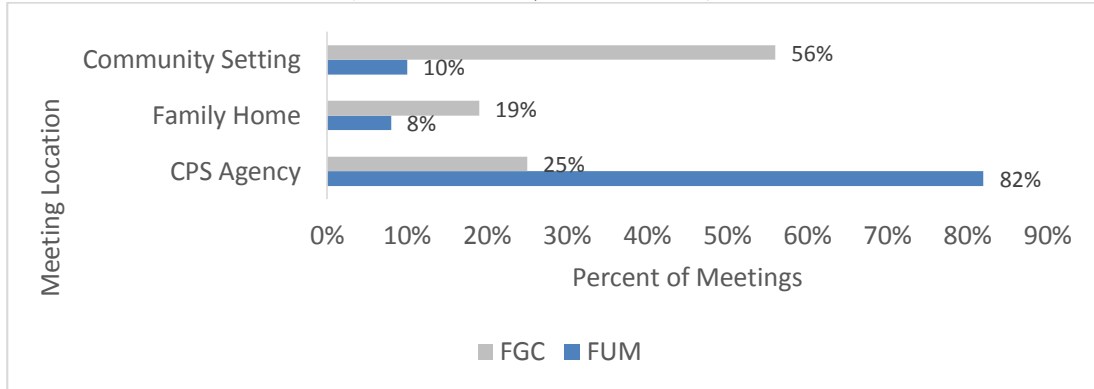
**Figure 3: Rates of attendance among different categories of participants  
(n=137 FUMs; n=17 FGCs)**



One of the practice standards identified in the FGDM literature is for coordinators to organize meetings that have twice as many family and like-family participants compared to service professionals in attendance. In Larimer County, this 2:1 family to professional ratio standard was met for 40 percent of FUMs and almost 49 percent of FGCs. Due to the small number of FGCs statistical tests to address differences in characteristics of FUMs compared to FGCs were not carried out.

*Meeting Location.* As shown in Figure 4, most FGCs (56%) took place in a community setting such as a community center, library, or place of worship, whereas most FUMs took place at the CPS agency (82%).

**Figure 4. Location of Family Meetings**  
(*n*=137 FUMs; *n*=17 FGCs)



*Length of Meetings and Private Family Time.*

On average, FUMs took 90 minutes (*SD* = 22 minutes), though facilitators reported durations ranging from 50 minutes to three hours. Most meetings (89%) were completed in two hours or less. Private family time was reported for only 4 FUMs (3% of FUMs), which is typically not a core component of the FUM. Families spent a median of 80 minutes in private family time (mean: 78 minutes; *SD*: 15). The minimum amount of private family time was 1 hour (*n* = 1), while the maximum was 90 minutes (*n* = 2).

On average, FGCs took 2 hours 12 minutes (*SD* = 34 minutes), with facilitators reporting durations ranging from 90 minutes to three hours 10 minutes. Around half of meetings (53%) were completed in two hours or less. Private family time was reported for 10 FGCs (59% of FGCs). Families spent a median of 65 minutes in private family time (mean: 67 minutes; *SD*: 21.1). The minimum amount of private family time was 30 minutes (*n* = 1), while the maximum was 90 minutes (*n* = 2).

*Use of Primary Language.* Facilitators indicated that 98% of FUMs were held in the family’s primary language; the remaining 2% were not held in the family’s primary language. Three

percent of FUMs were held in Spanish. One hundred percent of FGCs were held in the family's primary language, which was English.

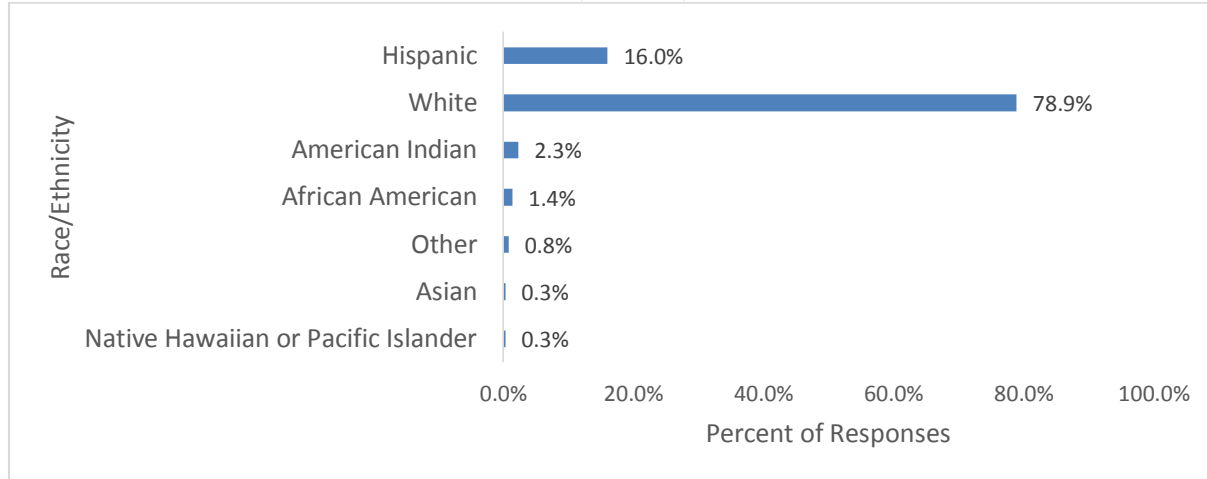
*Plan Approval.* Facilitators indicated that the family plan developed at the meeting was approved for 93% ( $n = 127$ ) of FUMs, and 71% ( $n = 12$ ) of FGCs. The lower percentage of plan approvals at FGCs, in comparison to FUMs, may reflect that many of the FGCs were held at case closure, when there may not have been a “live” decision or plan to be approved.

#### Larimer County Participant Fidelity Surveys: Respondent Characteristics

Participant surveys were received for 127 cases from a total of 685 cases at the participant-level. This means that some respondents completed multiple fidelity surveys, as they participated in family meetings at multiple points in time. Respondents to cases with more than 4 FUMs were excluded from these analyses to control for the potential effects of respondent fatigue, leaving a sample of 124 cases from a total of 663 respondents representing 160 total FUMs and FGCs available for analysis. Of the respondents, 71% could be classed as family or like-family/fictive kin, whereas 29% could be classed as professionals. The respondents included 29 children/youth, with a median age of 13 years (minimum: 8 years; maximum: 18 years). Excluding children/youth, the majority of family members, 66%, were maternal family members, and most respondents (73%) were female.

The racial/ethnic profile of respondents to the participant survey is shown in Figure 5. Hispanic participants made up 16% of the sample. A total of 34 participants did not respond to the question, and thus are not shown in Figure 5.

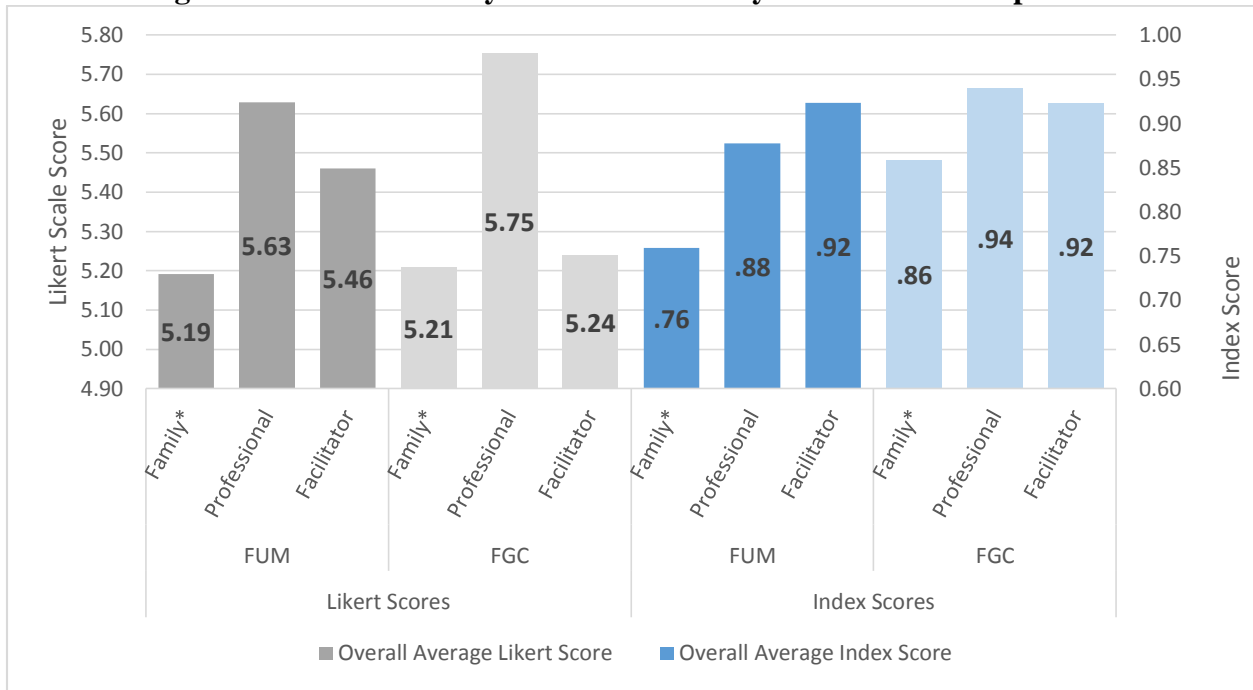
**Figure 5. Race/Ethnicity of Respondents to Participant Survey  
(n = 629)**



### Larimer County Fidelity Likert and Index Results

The overall mean fidelity Likert and FI scores for Larimer County are presented below in Figure 6. The figure identifies average scores by participant and meeting type (FUM and FGC). Both types of scores provide insights into the respondents’ perspectives about the process, content, and tone of the meeting, and how this reflects the core principles and practices of the family meeting processes. As described in the methodology section, above, the fidelity questions’ response options ranged from Strongly Disagree (1) to Strongly Agree (6). As the figure indicates, the mean Likert score was above five for all respondent types, indicating that on average, respondents “Agreed” that the elements constituting fidelity were present in the meeting across all of the questions asked of them. In the Index, scores reflect the average number of questions where the responses indicate their perception that certain fidelity-related practices occurred in the meeting. Differences in the relative magnitude of the Likert and FI scores are related to the metrics used to calculate them. In other words, the Likert scores indicate some degree of the *strength* of agreement in the questions being asked while the FI scores reflect the *presence or absence* of agreement.

**Figure 6. Larimer County Likert and Fidelity Index Score Comparison**



\*Note: Family category includes like-family respondents.

### Larimer County Fidelity Domains

In addition to the FI scores, as noted in the Methodology section, questions from the fidelity tools were separated into domains (preparedness, inclusion and respect, family leadership, and transparent planning) to provide a more nuanced look at fidelity based on the core FGDM principles and practices (see Table 3 for a listing of survey items composing each domain). There are three ways to look at the results: compare FI and domain scores across the respondent types, compare FI scores and domain scores within respondent types (not presented here), and compare scores between meeting types (FUMs vs. FGCs) for Larimer County (also not presented here). For the comparison of FI and domain scores across respondent types, tests of significance are important to indicate if the scores are meaningfully different.

Overall, for FUMs and across respondent types, the facilitators gave overall fidelity and domain fidelity the highest scores compared to service professionals and family respondents and family respondents gave the lowest average index and domain scores across all measures compared to everyone else. Of all the domains, the family and like-family survey respondent group gave the *transparent planning* domain the best average fidelity score. While the differences between the domain scores were slight, family and like-family respondents rated the *family leadership* and *inclusion and respect* domains the lowest for FUMs. Service professionals and facilitators also rated the *family leadership* domain the lowest of all the domains for Larimer FUMS. In contrast, service professionals and facilitators rated the *inclusion and respect* domain the highest among all of the domains, indicating disparate perceptions of performance on this domain.

The fidelity domain scores for FGCs generally follow the same patterns as the FUM results, with some exceptions. Families gave the lowest mean scores across respondent groups, except in two domains—*preparedness* and *transparent planning*—where the facilitator rated fidelity slightly lower. Of the domains, family respondents rated *preparedness* the highest and *inclusion and respect* the lowest, service professionals rated *inclusion and respect* the highest and *family leadership* the lowest, and the facilitators rated *inclusion and respect* the highest and *transparent planning* the lowest.

Finally, between the two meeting types and across respondents, FGCs received higher fidelity index and domain scores than FUMs, in general. The one exception to this is that facilitators rated FUMs higher than FGCs on *transparent planning*.

**Table 7. Larimer County Average Fidelity Index and Domain Fidelity Scores**

<b>Larimer FUMs</b>						
	<b>Family/Like-Family (n=290)</b>		<b>Service Professionals (n = 142)</b>		<b>Facilitator (n = 105)</b>	
	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation
<b>FI Score</b>	0.759	0.196	0.880	0.170	0.923	0.125
Preparedness (PD)	0.755	0.244	0.897	0.169	0.915	0.249
Inclusion and Respect (IRD)	0.750	0.183	0.936	0.190	1.000	1.000
Family Leadership (FLD)	0.750	0.256	0.827	0.235	0.884	0.202
Transparent Planning (TPD)	0.783	0.251	0.872	0.237	0.947	0.110
<b>Larimer FGCs</b>						
	<b>Family/Like-Family (n=58)</b>		<b>Service Professionals (n =18 )</b>		<b>Facilitator (n = 17)</b>	
	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation
<b>FI Score</b>	0.859	0.123	0.940	0.067	0.923	0.090
Preparedness (PD)	0.940	0.107	0.960	0.010	0.923	0.188
Inclusion and Respect (IRD)	0.804	0.126	0.974	0.092	1.000	1.000
Family Leadership (FLD)	0.878	0.148	0.917	0.144	0.981	0.069
Transparent Planning (TPD)	0.872	0.152	0.930	0.126	0.846	0.217

### Larimer County Paired Samples Fidelity Results

In order to determine if the differences observed between respondent types were significant, paired sample t-tests were examined. In total 685 Fidelity Surveys were received across 160 unique meetings.<sup>2</sup> Table 8 below shows the number of meetings, by type and by participant, for

<sup>2</sup> For this analysis, the full sample of surveys across the 160 unique meetings because the evaluation team wanted to leverage the highest number of surveys to be better able to detect differences.



which surveys were received for the same meeting. Using matched pairings enables a comparison of scores between participant types, whether or not a survey from a respondent in the omitted group was received. For some meetings, surveys were received from two but not all three of the respondent types. As the numbers indicate, family/like-family respondents and facilitators were the largest pairing group for both FUMs and FGCs. By comparing scores, the extent to which people involved in a given meeting have similar perceptions of the presence or absence of fidelity in that meeting can be illustrated.

**Table 8. Matched Survey Pairings (Meeting Level)**

FUM (n=71)			FGC (n=13)		
Family & Professional	Family & Facilitator	Professional & Facilitator	Family & Professional	Family & Facilitator	Professional & Facilitator
71	102	74	13	17	13

Table 9, below, provides the paired samples scores and p-values for each domain and available matched pair and reflects data from both meeting types combined, due to the low frequency of FGC meetings. A negative value in the mean column indicates that the first group listed in the pair type had a lower score than the second group. Where these differences are statistically significant at the  $p < .05$  level, the p-values (reported in the last column) are in bold. For example, a comparison of the mean FI scores reported by family/like-family versus professional respondents, indicated that the family FI scores were lower by an average of -0.12 and that this difference is statistically significant.

As Table 9 indicates, there are significant differences with respect to the FI scores between family/like-family members and professionals as well as between family members and the meeting facilitator. Differences between professionals and the meeting facilitator on the average FI scores were not significant.

Family and like-family respondents also consistently reported statistically significant, lower scores than either service providers or facilitators each of the four domains. Still, facilitators reported statistically significant and higher average scores on the *family leadership* and on the *inclusion and respect* domains compared to the service professionals. Service professionals and facilitators gave virtually identical scores on the *preparedness* domain, on average. Differences between service professionals and facilitators on the average FI scores and on *transparent planning* were not statistically significant, though the results suggest a trend of service professionals rating these fidelity elements lower (i.e., p-value was less than .10). The greatest disparity in mean scores appeared on family and facilitators' ratings of *inclusion and respect*.

**Table 9. Larimer County Paired Samples t-tests**

	<b>Domain</b>	<b>Pair Type</b>	<b>Mean</b>	<b><i>t</i></b>	<b><i>n</i></b>	<b>Sig. (2-tailed)</b>
Pair 1	FI Score	Family - Professional	-.12	<b>-5.70</b>	108	<b>.000</b>
Pair 2		Family – Facilitator	-.13	<b>-6.21</b>	119	<b>.000</b>
Pair 3		Professional - Facilitator	-.03	-1.68	87	.097
Pair 4	Family Leadership	Family - Professional	-.09	<b>-3.02</b>	108	<b>.003</b>
Pair 5		Family – Facilitator	-.12	<b>-4.29</b>	119	<b>.000</b>
Pair 6		Professional - Facilitator	-.06	<b>-2.05</b>	87	<b>.044</b>
Pair 7	Inclusion & Respect	Family - Professional	-.19	<b>-8.69</b>	108	<b>.000</b>
Pair 8		Family – Facilitator	-.24	<b>-12.84</b>	119	<b>.000</b>
Pair 9		Professional - Facilitator	-.06	<b>-3.18</b>	87	<b>.002</b>
Pair 10	Preparedness	Family - Professional	-.13	<b>-4.94</b>	108	<b>.000</b>
Pair 11		Family – Facilitator	-.12	<b>-3.80</b>	119	<b>.000</b>
Pair 12		Professional - Facilitator	.00	.02	87	.982
Pair 13	Transparent Planning	Family - Professional	-.10	<b>-3.33</b>	108	<b>.001</b>
Pair 14		Family – Facilitator	-.11	<b>-3.75</b>	119	<b>.000</b>
Pair 15		Professional - Facilitator	-.05	-1.79	87	.077

\* Results in bold are significant where  $\alpha = 0.05$

## C. Fidelity Results - Texas DFPS Region 3

### Description of Texas Region 3 Family Group Conferences

*Number of FGC Meetings.* The practice model in Texas Region 3 was such that families were referred for one FGC during the course of the study; additional FGCs in the FBSS stage of service were rare and, as such, not tracked as part of the study. According to the Meeting Log, overall 73% of families ( $n = 196$ ) who were offered an FGC actually participated in one.

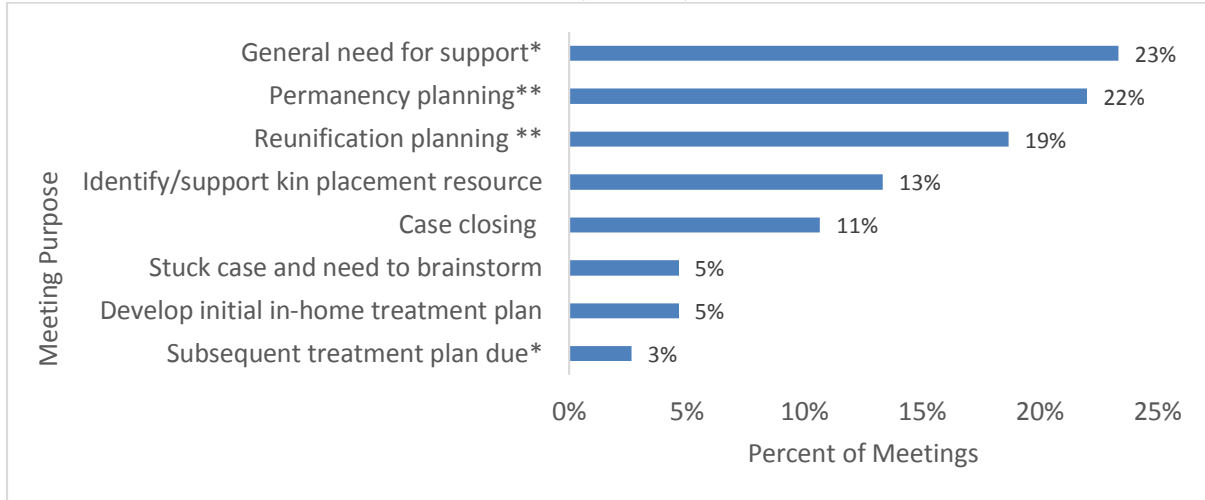
*Length of Time between Referral and Meeting.* On average, meetings were held within 41 days of a referral ( $SD = 31.7$  days) in Dallas and Tarrant Counties. However, meetings occurred as quickly as 20 days in 17.8% of the cases ( $n = 35$ ) while 13.3% ( $n = 26$ ) took place over 60 days after the referral.

### Texas Region 3 Facilitator Fidelity Results

A total of 16 facilitators provided fidelity data. These facilitators held a median of 10 meetings each (minimum: 2 meetings; maximum: 22 meetings).

*Meeting Purpose.* Facilitators were asked to report the purpose of each FGC held. The results are shown in Figure 7, below. The most common reason for having the FGC was a general need for support for a child at home (23%), followed by permanency planning (22%) and reunification planning (19%) for a child in placement. One respondent did not indicate the reason for the meeting (0.7%).

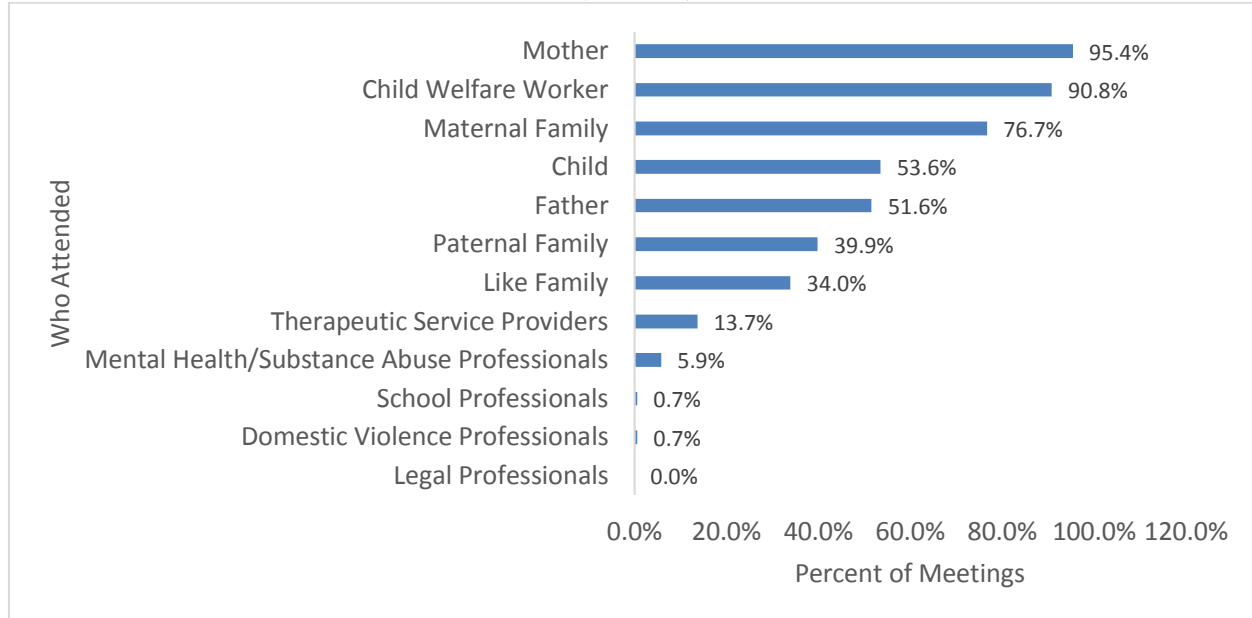
**Figure 7. Purpose of FGCs  
(n = 153)**



\* Child at home; \*\* Child in placement

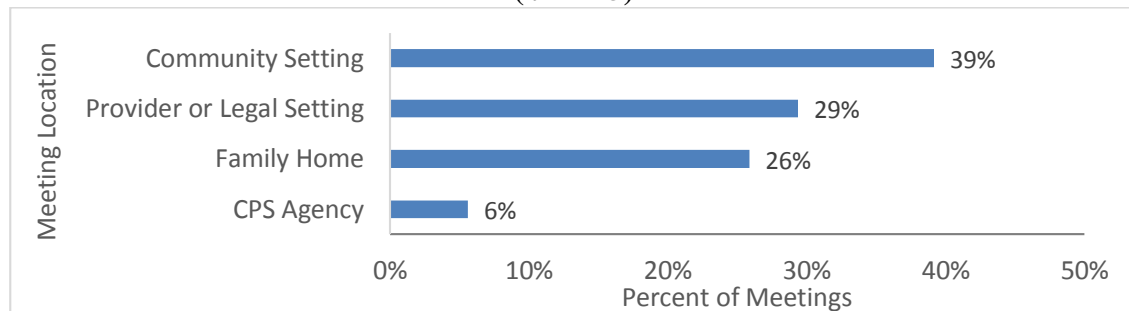
*Meeting Attendance.* Facilitators reported that a median of 8 people attended FGCs in Region 3 (minimum: 2; maximum: 21). Overall, the standard of having twice as many family or like-family participants compared to service professionals was met in 45 percent of the FGCs. As shown in Figure 8, mothers and child welfare workers had the highest rates of attendance (95% and 91%, respectively). Mothers and maternal family members had higher rates of attendance (95% and 77%, respectively) than fathers and paternal family members (52% and 40%, respectively). Professionals other than child welfare workers had relatively low rates of attendance, with therapeutic service providers being most likely to attend meetings (14%).

**Figure 8. Rates of Attendance among Different Categories of Participants  
(n = 153)**



*Meeting Location.* As shown in Figure 9, most FGCs (39%) took place in a community setting, such as a community center, library, place of worship, or recreation center. About a quarter (26%) took place in the family’s home.

**Figure 9. Location of FGCs  
(n = 145)**



*Length of Meetings and Private Family Time.* On average, meetings took two hours ( $SD = 40$  minutes), though facilitators reported durations ranging from one to four hours. Most meetings (66%) were completed in two hours or less; only 15 percent of meetings took three hours or more.

Private family time was reported for 118 meetings. Families spent a median of 30 minutes in private family time (mean: 34.15 minutes; *SD*: 17.46). The minimum amount of private family time was 10 minutes ( $n = 12$ ), while the maximum was 90 minutes ( $n = 2$ ).

*Use of Primary Language.* Facilitators indicated that 99.3% of the meetings were held in the family's primary language (response was missing for 1 meeting). Nine percent of the meetings ( $n = 14$ ) were held in Spanish, six percent in both Spanish and English ( $n = 9$ ) and one meeting was held in Mandarin.

*Plan Approval.* Facilitators indicated that the family plan was approved at the FGC for 95% ( $n = 145$ ) of the 153 meetings.

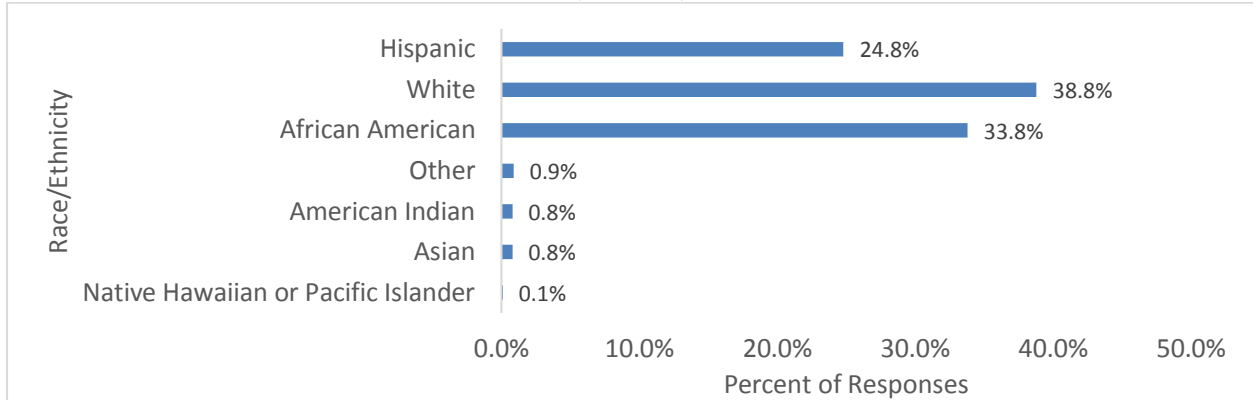
#### Texas Region 3 Participant Fidelity Surveys: Respondent Characteristics

Participant surveys were received for 156 FGCs from a total of 834 respondents. Of the respondents, 85% could be categorized as family or like-family/fictive kin, whereas 15% could be categorized as professionals. The respondents included 9 children, with a median age of 12 years (minimum: 9 years; maximum: 15 years). Excluding children, the majority of family members, 62%, were maternal family members. Most respondents (70%) were female. About one-quarter of the sample (26%) was male, and 4% did not respond to the question about gender.

The racial/ethnic profile of respondents to the participant survey is shown in

Figure 10. Hispanic participants made up 25% of the sample. A total of 45 participants did not respond to this question, and therefore are not shown in Figure 10.

**Figure 10. Race/Ethnicity of Respondents to Participant Survey  
(n = 789)**

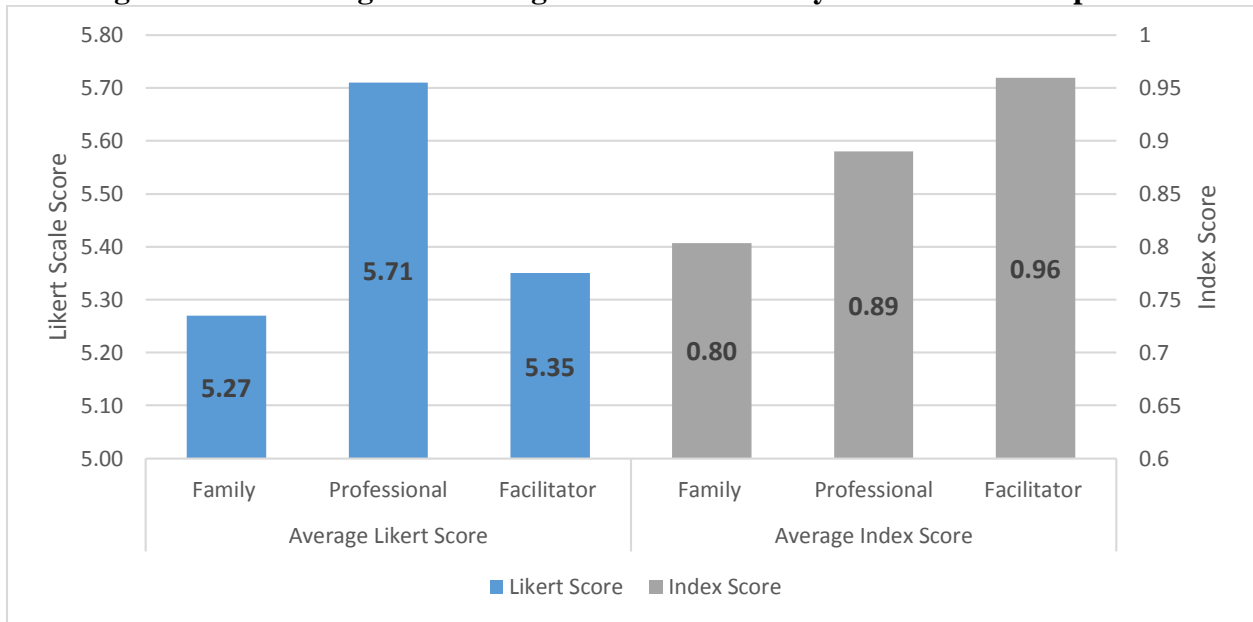


### Texas Region 3 Fidelity Index

The mean score comparison between Likert and FI scores, analyzed by respondent type, are illustrated below in Figure 11. Generally, the family participants rated overall fidelity the lowest, professional participants rated overall fidelity higher than family, while facilitators rated overall fidelity the highest. On average for all FGCs, families rated overall fidelity a 0.80 on the index, indicating at least moderate fidelity to the model. As explained above, differences in the relative magnitude of the Likert and FI scores are related to the metrics used to calculate them. In other words, the Likert scores indicate some degree of the *strength* of agreement in the questions being asked while the FI scores reflect the *presence or absence* of agreement.



**Figure 11. Texas Region 3 Average Likert and Fidelity Index Score Comparison**



### Texas Region 3 Fidelity Domains

In addition to the FI scores, as noted in the Methodology section, questions from the fidelity tools were separated into domains to provide a more nuanced look at fidelity based on the core FGDM principles.

As show in Table 10, within respondent types, family respondents gave the *inclusion and respect* domain (IRD) the lowest score while service professionals and facilitators both indicated that *family leadership* domain (FLD) was the weakest of the domains. In contrast, service professionals and facilitators rated *inclusion and respect* (IRD) the highest of the domains while family respondents indicated that the *transparent planning* (TPD) was the strongest domain.

**Table 10. Texas FGC Average Fidelity Index and Domain Scores by Respondent Type**  
(*n* = 124)

	Family/Like Family ( <i>n</i> = 516)		Service Professionals ( <i>n</i> = 122)		Facilitator ( <i>n</i> = 124)	
	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation
<b>FI Score</b>	0.80	0.13	0.89	0.16	0.96	0.08
Preparedness (PD)	0.81	0.19	0.92	0.12	0.99	0.06
Inclusion and Respect (IRD)	0.76	0.14	0.95	0.18	1.00	1.00
Family Leadership (FLD)	0.82	0.15	0.84	0.21	0.91	0.19
Transparent Planning (TPD)	0.84	0.17	0.88	0.23	0.98	0.06

### Texas Region 3 Paired Samples Fidelity Results

In order to determine if the differences observed between respondent types were significant, paired sample *t*-tests were examined. Using matched pairings enables a comparison of scores between participant types, whether or not a survey from a respondent in the omitted group was received. For some meetings, surveys were received from two but not all three of the respondent types. Overall, there were 63 FGCs where family or like-family respondents and a service professional submitted surveys, 124 FGCs in which at least one family or like-family respondent and facilitator submitted surveys, and 63 FGCs where perspectives from service professionals could be compared to perspectives of facilitators. As the numbers indicate, family/like-family respondents and facilitators were the largest pairing group. By comparing scores, the extent to which people involved in a given meeting have similar perceptions of the presence or absence of fidelity in that meeting can be illustrated.

Table 11 provides the paired samples scores and *p*-values for each domain and available matched pair. Again, a negative value in the mean column, indicates that the first group listed in the pair type had a lower score than the second group. Where these differences are statistically

significant at the  $p < .05$  level, the p-values (reported in the last column) are in bold. For example, a comparison of the mean FI scores reported by family/like-family versus professional respondents, indicated that the family FI scores were lower by an average of -0.09 and that this difference is statistically significant.

Overall significant differences exist with respect to FI scores across all three matched pairs; between family members and professionals, family members and the meeting facilitator, and professionals and the meeting facilitator. While the family and service providers reported similar domain scores for the *family leadership* and *transparent planning* domains, all other comparisons of domain scores by respondent type were identified as statistically significant. The greatest disparity in mean scores appeared on family and facilitators' ratings of *inclusion and respect*.

**Table 11. Texas Paired Samples *t*-test**

	<b>Domain</b>	<b>Pair Type</b>	<b>Mean</b>	<b><i>t</i></b>	<b><i>n</i></b>	<b>Sig. (2-tailed)</b>
Pair 1	FI Score	Family - Professional	-.09	<b>-3.41</b>	63	<b>.001</b>
Pair 2		Family - Facilitator	-.16	<b>-11.17</b>	124	<b>.000</b>
Pair 3		Professional - Facilitator	-.07	<b>-3.00</b>	63	<b>.004</b>
Pair 4	Family Leadership	Family - Professional	-.018	-0.53	63	.599
Pair 5		Family - Facilitator	-.09	<b>-4.21</b>	124	<b>.000</b>
Pair 6		Professional - Facilitator	-.07	<b>-2.08</b>	63	<b>.042</b>
Pair 7	Inclusion & Respect	Family - Professional	-.19	<b>-6.05</b>	63	<b>.000</b>
Pair 8		Family - Facilitator	-.24	<b>-15.15</b>	124	<b>.000</b>
Pair 9		Professional - Facilitator	-.05	<b>-2.30</b>	63	<b>.025</b>
Pair 10	Preparedness	Family - Professional	-.11	<b>-4.59</b>	63	<b>.000</b>
Pair 11		Family - Facilitator	-.17	<b>-8.97</b>	124	<b>.000</b>
Pair 12		Professional - Facilitator	-.07	<b>-4.13</b>	63	<b>.000</b>
Pair 13	Transparent Planning	Family - Professional	-.04	-1.13	63	.263
Pair 14		Family - Facilitator	-.15	<b>-8.66</b>	124	<b>.000</b>
Pair 15		Professional - Facilitator	-.10	<b>-3.35</b>	63	<b>.001</b>

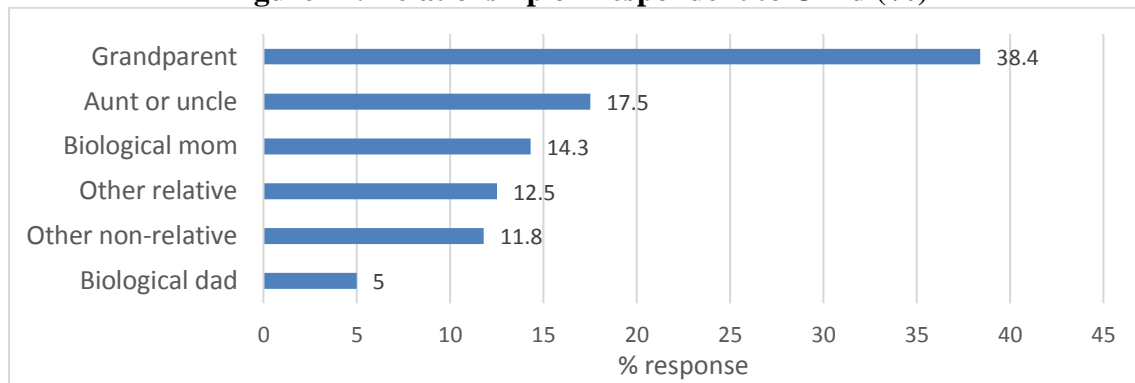
\* Results in bold are significant where  $\alpha = 0.05$

### Texas Region 3 Post-Test FGDM Fidelity

FGDM fidelity, post-meeting and from family members' perspectives, has been understudied. A goal of this project was to gain a greater understanding of family member's perspectives of family meetings after time (approximately 3-6 months) had passed from the initial meeting where a decision or plan was made. This time lapse provides family with an opportunity to see the extent to which plans made at the FGC were implemented as intended or modified, and the degree to which agreed upon services and supports are delivered. In turn, it is reasonable to expect family members' perceptions of what occurred at the FGC to change, based on these experiences.

Therefore, post-test fidelity surveys were mailed to family and like-family respondents of the fidelity survey at the FGC, if they provided their willingness to be contacted by the evaluators. In total, 120 family and like-family individuals provided post-fidelity surveys. As Figure 12 indicates, the majority of these respondents were relatives, with grandparents (38.4%) as the highest respondent group, followed by aunts or uncles (17.5%). "Other relative" included: brother's grandchild/grandson, cousin, godmother/father, great grandmother/parent, step grandmother, step nephew, stepdad, and stepdaughter's little sister. "Other non-relative" included: caregiver/babysitter, family friends/neighbors, godmother, mom's significant other, grandfather figure.

**Figure 12. Relationship of Respondent to Child (%)**



Respondents were asked to reflect on their experiences at the FGC, and rate their level of agreement with the statements where 1= “Strongly disagree” and 6 = “Strongly agree.” As Table 12 shows, respondents, on average, rated the notion of family autonomy or leadership in decision making the lowest, with an average score of 3.67 on the item, “Paid professionals at the meeting did not tell our family how to solve the agency’s concerns.” This may reflect family’s retrospective perception that the agency was directing or steering the plan in a certain direction. Closely related are a few of the other lowest rated statements, including “the family completed parts of the plan they agreed to at the meeting” and “we received the services that we put in the plan,” both registering an average score of 4.72. Additional analyses comparing scores from the individual meetings may be helpful to understanding whether there is measurable change over time related to fidelity, and if there are differences between scores on individual items, the relationship of those scores to case outcomes.

**Table 12: Post-Test Mean Scores of Fidelity Survey Items**

<b>Fidelity Survey Item</b>	<b><i>n</i></b>	<b>Mean</b>	<b>Median</b>	<b>SD</b>	<b>Does not apply (<i>n</i>)</b>
2. The right people were at the family meeting.	119	5.02	5.00	1.18	0
3. Paid professionals at the meeting did not tell our family how to solve the agency's concerns.	108	3.67	4.00	1.77	4
4. Others listened to my opinions about what was best for the child at the meeting.	117	5.03	5.00	.99	2
5. The facilitator worked with the family and the child welfare agency to reach a final plan that all could agree on at the meeting.	119	5.17	5.00	.93	0
6. The decisions made at the meeting were the best decisions for the child.	118	5.05	5.00	1.19	1
7. The facilitator respected me at the family meeting.	119	5.44	6.00	.73	2
8. The child welfare agency staff respected me at the family meeting.	115	5.30	5.00	.88	3
9. The family completed the parts of the plan they agreed to do at the family meeting.	117	4.72	5.00	1.47	3
10. The child welfare agency completed the parts of the plan they agreed to do at the family meeting.	119	5.02	5.00	1.23	2
11. My caseworker and I discussed the plan developed at the family meeting during our visits.	94	4.79	5.00	1.37	22
12. The plan developed at the family meeting was implemented.	110	4.81	5.00	1.28	6
13. The plan developed at the family meeting needed to be changed.	106	2.89	2.00	1.58	12
14. We received the services that our family put in the plan.	98	4.72	5.00	1.35	16
15. I would recommend the family meeting process to others.	117	4.98	5.00	1.26	2

In addition to the asking these respondents to rate their perceptions of the presence of fidelity elements that were part of their family meeting, the follow-up survey also asked general questions about their experiences with family meetings as well as some general satisfaction items. For these scales, statements were rated from a 1 = “Strongly disagree” to 6 = “Strongly agree.” As shown in Table 13, while all means scores hover in the slightly agree to moderately agree category, the three that were the lowest related to their perceptions in being better able to parent (mean=4.48), and meet their basic needs (mean=4.06) because of their experiences with

CPS, and their likelihood of calling CPS if they needed help in the future (mean=4.54). The two highest rated statements related to CPS recognizing the family strengths (mean=4.8) and listening to the family (mean=4.75).

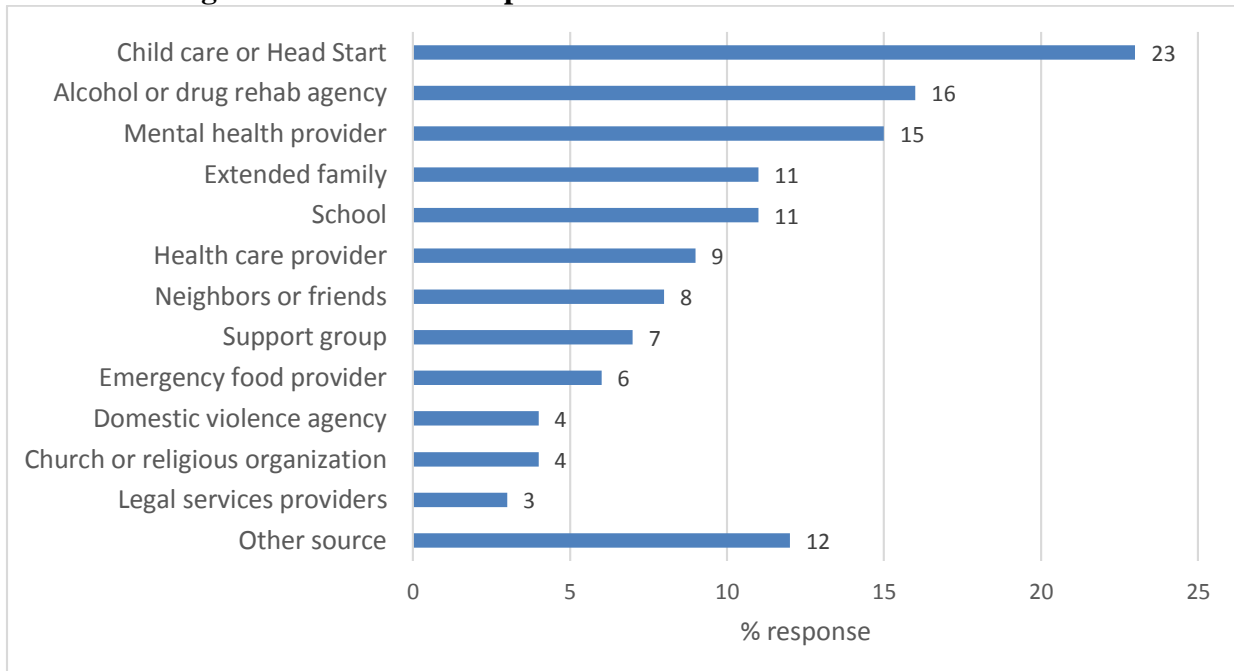
**Table 13: Post-Test Mean Scores of Satisfaction Items**

<b>Fidelity Survey Item</b>	<b><i>n</i></b>	<b>Mean</b>	<b>Median</b>	<b>SD</b>
19. CPS provided my family with enough information about their concerns.	117	4.69	5.00	1.27
20. CPS listened to what my family had to say.	118	4.75	5.00	1.32
21. CPS understood my family's needs.	119	4.65	5.00	1.34
22. CPS recognized the things that my family does well.	113	4.80	5.00	1.19
23. CPS considered my family's culture when working with us.	115	4.65	5.00	1.42
24. CPS considered my family's opinions before making decisions about us.	117	4.58	5.00	1.48
25. I am a better parent or caregiver because of my experience with CPS.	110	4.48	5.00	1.50
26. My children are safer because of our experience with CPS.	113	4.58	5.00	1.52
27. I am better able to provide necessities like food, clothing, shelter, or medical services because of my experience with CPS.	109	4.06	5.00	1.82
28. CPS provided services to meet my family's needs.	114	4.66	5.00	1.40
29. Overall, I am satisfied with how my family was treated by CPS.	116	4.72	5.00	1.42
30. Overall, I am satisfied with the help my family received through CPS.	117	4.74	5.00	1.37
31. Overall, I am satisfied with the plan for my child(ren).	114	4.75	5.00	1.42
32. I would call CPS if my family needed help in the future.	117	4.54	5.00	1.67

A series of questions were asked to ascertain the role of the CPS agency in delivering or connecting families to services. Based on a listing of services, and as shown in Figure 13, a sum of sources of help received was computed. Mean number of services received was 1.72 (SD: 1.20); the median number was 1, minimum was 1, and maximum was 7 (based on  $n = 75$ ). The

three most frequently identified services received were child care/Head Start (23%), alcohol or drug rehabilitation (16%), and mental health treatment (15%).

**Figure 13. Source of Help Received as a Result of CPS Involvement**

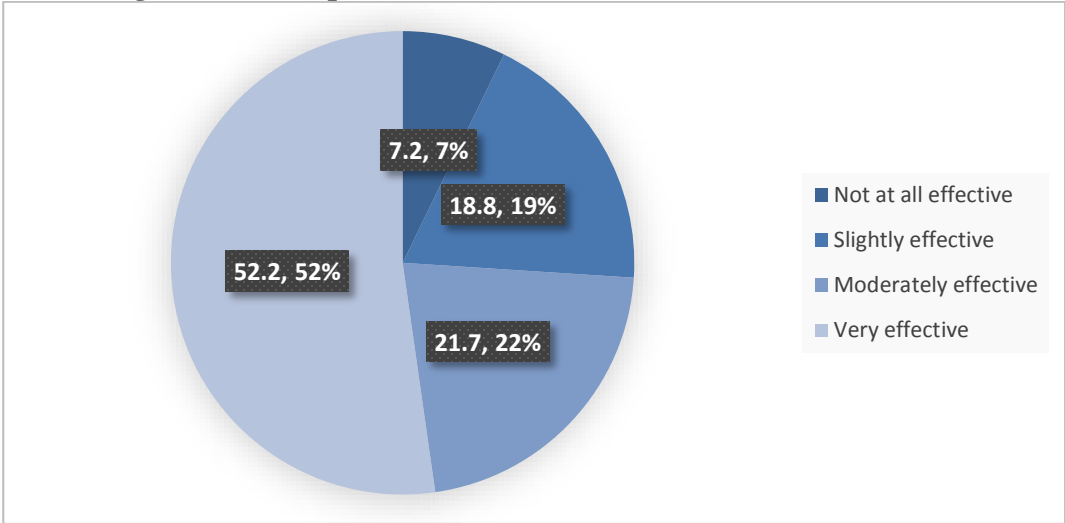


\*Other sources ( $n = 1$  in each instance): Bus passes; Daycare, Parenting classes, Psychologist, Therapy and parenting classes.

Next, those who received services ( $n=69$ ), on a scale of 1 = “Not at all effective” to 4 = “Very effective,” were asked to rate their perception of effectiveness of help or services received in assisting them with their problems. With a mean average score of 3.19 (SD .99), as shown in Figure 14, close to 75% of respondents believed the services they received to be very or moderately effective; a small percentage (7.2%) rated the services as not effective at all. In a more nuanced analysis, for each source of help (Figure 14), there was no association between whether or not help was received and overall perception of effectiveness of help or services received. There was also no association between the number of services received and perception of effectiveness (Spearman’s  $r = -.14$ ,  $p = .28$ ,  $n = 65$ ).

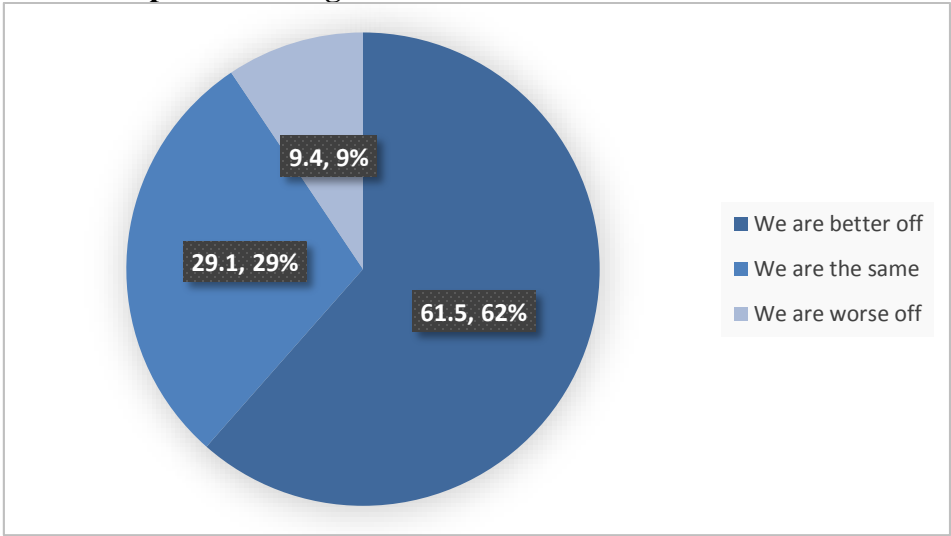


**Figure 14. Perception of Effectiveness of Services Received (%)**



Lastly, as a way to measure family perceptions of improvements given their involvement with CPS, respondents rated whether their family was better or worse off because of their experiences with CPS. As shown in Figure 15, of the 117 respondents, over half (61.5%) responded they were better off, while 29.1% responded they were the same, and 9.4% reported being worse off.

**Figure 15. Perception of Being Better or Worse Off Due to CPS Involvement (%)**



## VII. Outcome Evaluation Results

### A. Outcome Analysis - Larimer County

#### Larimer County Target Population

The Larimer County target population consisted of families open to in-home, “ongoing” services who were referred to any type of family meeting (FSRT, FUM, or FGC) during the study period and could either be in the “high risk” or “family assessment response” track as Larimer County is a differential response organized child welfare system. The intervention sample was constructed from among those families who received at least one FUM or FGC while receiving ongoing in-home services during the study period. Once assigned to the intervention group, a comparison sample was selected based on propensity score matching among cases from the study population who did not receive a FUM or FGC during the study period. The rationale behind these eligibility criteria were that in Larimer County, all investigations (called high risk assessments) require that an FSRT meeting be held, such that almost all of both the comparison and intervention group cases received at least one FSRT. Family meetings (FSRTs, FUMs, and FGCs) are such an ingrained and ubiquitous feature of Larimer County’s child welfare practice that stipulating no family meetings as the condition for the comparison group was expected to yield an insufficient number of comparison cases. In fact, according to an administrative data set on family meeting service authorizations compiled by Larimer County, during the study period (October 20, 2012-April 30, 2014), there were 1,078 FSRTs, 1,267 FUMs, and 180 FGCs held.

## Larimer County Sample

Propensity score matching (PSM) was used to examine FGDM outcomes in Larimer County. PSM is a commonly-used quasi-experimental design when randomized control designs are not possible. Briefly, PSM estimates the effect of a treatment by accounting for the covariates that predict receiving the treatment; that is, systematic differences in baseline characteristics between treated and untreated participants (Austin, 2011). The unit of analysis for the Larimer County outcome evaluation is children. The ‘treatment’ group consisted of 466 individual children in 262 families where the families were involved in at least one FGC or FUM meeting during a six- to nine-month period after the case was opened to in-home services. The matched control group consisted of 287 individuals in 190 families. The control individuals were selected from an overall sample of 392 individuals in 241 families who did not receive an FGC or FUM in the six- to nine-month period after the case was opened.

Intervention and comparison children were matched on the basis of propensity scores, which collapse a set of background covariates into a single summary measure (the propensity score), representing an estimate of the probability of receiving treatment. By matching on the basis of propensity scores, the comparison group included individuals who had the same probability of receiving an FGC or FUM meeting as those in the treatment group who actually received an FGC or FUM meeting. PSM involves several analysis steps, namely, (1) selecting background covariates to be included in the propensity score model, (2) estimating propensity scores and using the scores for matching, and (3) evaluating the propensity score model by examining balance diagnostics. Please see

Appendix E. Larimer County Propensity Score Match for a detailed synopsis of the propensity score matching process conducted in Larimer County.

### Larimer County Study Participant Characteristics

#### **Characteristics of children involved in the study, drawn from the SACWIS administrative data, are shown in**

Table 14. Characteristics are shown for the treatment group ( $n = 466$  individuals) and PSM-matched control group ( $n = 287$  individuals) separately, as well as for the combined sample ( $n = 753$  individuals). Comparisons between the treatment and comparison groups were made using chi-square tests for continuous and ordinal variables, and independent  $t$ -tests for continuous variables. Sample characteristics for the treatment and comparison groups prior to matching are provided in Appendix F: Larimer County Child Characteristics Prior to PSM.

#### *Demographic characteristics*

The population overall included slightly more males (54.4%) compared to females (46.5%). This pattern was mirrored in the treatment sample (55.5% male). Within the comparison group, the sample was evenly split on gender. The gender difference between the treatment and comparison group sample was not significant.

The mean age of children in the population was 6.41 years (SD: 4.87). The difference between the mean ages of children in the treatment and comparison groups was not statistically significant, although children in the comparison group were, on average, slightly younger (mean: 6.22 years; SD: 4.95) than children in the treatment group (mean: 6.73 years; SD: 4.72). All age groups (0-1 year, 2-5 years, 6-10 years, and 11+ years) were approximately equally represented in the sample, although mirroring the mean scores by group, the proportion of young children was slightly higher in the treatment group (0-5 years: 49.6%) than in the comparison group (0-5 years: 41.4%).

In both the treatment and comparison group, just over one-quarter of children were Hispanic (26.2% in the sample overall). Within the treatment group, other minority racial/ethnic groups included African American (.9%), and non-Hispanic mixed race (3.4%). Within the comparison group, other minority racial/ethnic groups included Asian (.7%) and non-Hispanic mixed race (3.1%).

As described above, the treatment group of 466 children were part of 262 families, whereas the comparison group consisted of 287 children in 190 families. The average number of children in treatment group families (mean: 2.45; SD: 1.37) was significantly greater than the average number of children in comparison group families (mean: 2.03; SD: 1.05).

#### *Maltreatment allegations and risk factors*

In both groups, neglect allegations were almost 4 times more common than abuse allegations. The rate of caregiver history of child maltreatment was significantly higher in the treatment group (33.8%) than in the comparison group (21.3%). Rates of substance abuse and domestic violence were also higher in the treatment group (50.2% and 56.3% for substance abuse and domestic violence, respectively) than in the comparison group (46.3% and 54.4% for substance abuse and domestic violence, respectively); however, these differences were not statistically significant.

#### *Prior CPS involvement*

In terms of prior CPS involvement, almost three-quarters of the sample had previous reports of suspected maltreatment. Nine percent of children in the treatment group and 12.5% of children in the comparison group had no previous accepted reports of maltreatment, but they did have at least one prior report that had been screened out. Around two-thirds of the sample (68.9% of the treatment group; 64.1% of the comparison group) had at least one prior report that had been

accepted for further investigation. A categorical version of this variable (no prior reports vs. at least one prior report) was also created; in the sample overall, 77.4% of children had a prior report of maltreatment (accepted or screened out), whereas 22.6% of children had no prior reports. The treatment and comparison groups did not differ significantly on rates of prior reports of maltreatment. To further gauge prior CPS involvement, we considered whether children had previously been involved in a Family Assessment Response (FAR). A significantly greater proportion of children in the comparison group had previously been involved in FAR (70.4%) compared with the treatment group (54.9%).

**Table 14: Child Characteristics, Larimer County, After Propensity Score Matching (PSM)**

Categorical/ordinal variables		Treatment (n = 466)		Comparison (n = 287)		$\chi^2$ (df)	p	Total (n = 753)	
		n	%	n	%			n	%
<i>Demographic characteristics</i>									
Gender	Male	257	55.2	209	50.9	1.31 (1)	.25	403	54.4
	Female	209	44.8	141	49.1			350	46.5
Age	0-1 years	103	22.1	54	18.8	5.47 (3)	.14	157	20.8
	2-5 years	128	27.5	65	22.6			193	25.6
	6-10 years	127	27.3	98	34.1			225	29.9
	11+ years	108	23.2	70	24.4			178	23.6
Ethnicity	African American	4	.9	0	0	5.87 (5)	.32	4	.5
	Asian	0	0	2	.7			2	.3
	Caucasian	324	69.5	198	69			522	69.3
	Hispanic	120	25.8	77	26.8			197	26.2
	2 or more non-Hispanic	16	3.4	9	3.1			25	3.3
Unknown	2	.4	1	.3			3	.4	
<i>Maltreatment allegations and risk factors</i>									
Abuse allegations	Yes	106	22.7	53	18.5	1.95 (1)	.16	159	21.1
	No	360	77.3	234	81.5			595	78.9
Neglect allegations	Yes	367	78.8	227	79.1	.01 (1)	.91	594	78.9
	No	99	21.2	60	20.9			159	21.1
Caregiver childhood history A/N	Yes	157	33.8	61	21.3	13.6 (1)	< .01	218	29.0
	No	307	66.2	226	78.7			533	70.8
Risk of substance abuse	Yes	233	50.2	133	46.3	1.07 (1)	.30	366	48.6
	No	231	49.8	144	53.7			385	51.1
	Yes	261	56.3	156	54.4			.26 (1)	.61

Risk of domestic violence	No	203	43.8	131	45.6			334	44.4
<i>Prior CPS involvement</i>									
Priors – ordinal	No priors	103	22.1	67	23.3	2.86 (2)	.55	170	22.5
	No accepted priors, but at least 1 prior that was screened out	42	9.0	36	12.5			78	10.4
	At least 1 prior that was accepted	321	68.9	184	64.1			505	67.1
Priors - categorical	Yes	363	77.9	220	76.7	.16 (1)	.69	583	77.4
	No	103	22.1	67	23.3			170	22.6
FAR	Yes	256	54.9	202	70.4	17.79 (1)	< .01	458	60.8
	No	210	45.1	85	29.6			295	39.2
		<b>Treatment</b>		<b>Control</b>				<b>Total</b>	
<b>Continuous variables</b>		<b>n</b>	<b>Mean (SD)</b>	<b>n</b>	<b>Mean (SD)</b>	<b>t (df)</b>	<b>p</b>	<b>n</b>	<b>Mean (SD)</b>
<i>Demographic characteristics</i>									
Age		466	6.22 (4.95)	287	6.73 (4.72)	1.41 (751)	.16	753	6.41 (4.87)
Number of children		466	2.45 (1.37)	287	2.03 (1.05)	-4.72 (717)	< .01	753	2.29 (1.27)

Note: Caregiver childhood history A/N: Caregiver history of childhood abuse or neglect. Prior - ordinal: Prior involvement with CPS (0 = no priors; 1 = no accepted priors but  $\geq$  prior that was screened out; 2  $\geq$  1 prior that was accepted), Priors - categorical: Prior involvement with CPS (0 = no prior involvement; 1 = prior involvement); FAR: Prior involvement in Family Assessment Response (FAR).

In summary, there were few significant differences between the treatment and comparison groups on the variables used for PSM matching. However, rates of caregiver history of childhood maltreatment were significantly higher in the treatment group than in the comparison group, and the number of children on a case was greater, on average, in the treatment group. Furthermore, rates of previous FAR were higher in the comparison group than in the treatment group. Because the comparisons shown in Table 14 indicate that there remained significant differences on these variables after matching, it is necessary that they be used as covariates in the outcome analyses.

### Larimer County Outcome Results

Administrative data from Larimer County was used to analyze effects of FGDM on screened-in re-referrals to the child welfare system and removals (out-of-home placements) after the initial referring case had closed. Case closure was taken into account to ensure that outcomes occurred

after, rather than during, the initial case and to account for the fact that multiple FUMs and FGCs over an extended period may have occurred for the treatment families.

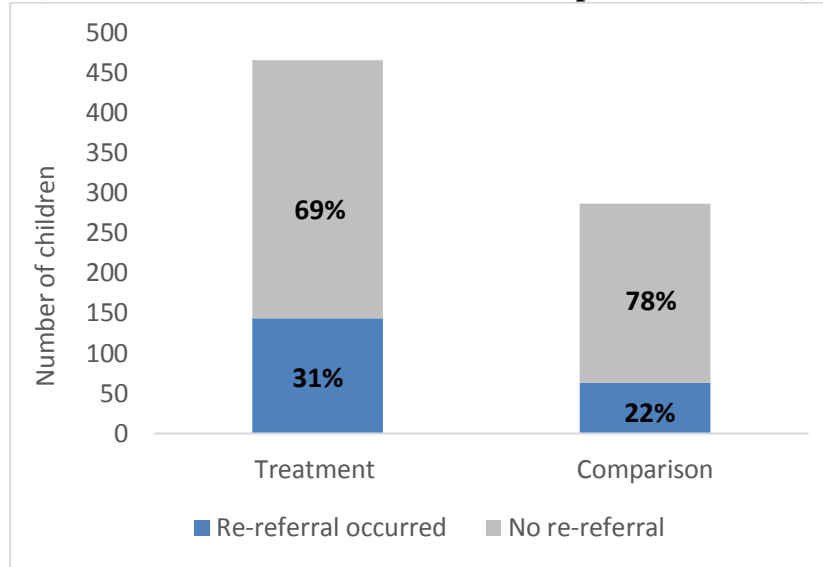
Within the matched sample of 753 individuals, the initial case had closed for 466 individuals (61.9%) by the last date of the administrative data collection period, July 3, 2015. Initial cases took significantly longer to close in the treatment group (mean number of days to case closure = 274 days; SD = 178.16) compared to the comparison group (mean = 149 days; SD = 121.43). As a note the majority of cases in the comparison group would have received an initial FSRT as did most of the treatment group. Additionally, case closure by the last date of data collection was less likely to have occurred in the treatment group ( $n = 34, 7.3\%$ ) than the comparison group ( $n = 9, 3.1\%$ ).

#### *Screened-in re-referrals*

Screened-in re-referrals after initial case closure occurred for 207 individuals within the matched sample. Figure 16 shows the number and proportion of children with screened-in re-referrals separately for the treatment and comparison groups. Based on an analysis that does not include any other characteristics of children, families, or the case, individuals in the treatment group were significantly more likely to have a screened-in re-referral after case closure ( $n = 144, 30.9\%$ ) compared with individuals in the comparison group ( $n = 63, 22\%$ ) ( $\chi^2 = 7.14[1], p < .01$ ). The average number of days to screened-in re-referral from case closure was similar in the treatment and comparison groups ( $t = .68[92]; p = .50$ ).



**Figure 16. Screened-In Re-Referral Occurrence Among Families  
(with  $n=466$  treatment and  $n = 287$  comparison children)**



It should be noted that these simple, unadjusted comparisons do not take into account any covariates that may affect the outcomes, or account for the clustering of children in families. In addition, we know there are still some remaining significant differences between the treatment and comparison groups used for the analysis. Thus, these results should not be taken to evaluate the effects of FGDM and are shared for descriptive purposes only. We compared rates of screened-in re-referrals in the treatment and comparison groups using Cox regression models in SAS. The dependent variable was time (number of days) to screened-in re-referral, estimated as the difference in days between the date of screened-in re-referral and date of initial case closure among individuals who received a screened-in re-referral, and difference in days between the data collection end date and the date of initial case closure among individuals who did not receive a screened-in re-referral. We used Cox regression methods to account for the fact that some cases were still open at the time the study ended.

Cox regression models used data from the whole sample of 858 individuals prior to matching (see Appendix E), and included both treatment status and the estimated propensity scores as predictors. We opted to use this approach, *covariate adjustment using propensity scores* (Austin, 2011), due to the relatively low frequency of the outcome of interest. Although covariate adjustment using propensity scores increases the risk that systematic differences in baseline characteristics between treatment and comparison cases remain, compared with simply using the PSM-matched sample, it eliminates the loss of cases resulting from “unmatched” propensity scores.

The primary predictor in our cox regression models was group (coded as a dummy variable, where 1 = treatment and 0 = comparison). Propensity scores were included as a covariate to account for the probability of assignment to the treatment or control group based on the baseline characteristics (described in Appendix E). Age, gender, Hispanic ethnicity, caregiver history of childhood abuse or neglect, number of children on a case, and prior involvement in FAR were included as covariates to account for any significant differences between the two groups that remained even after the matching procedure. For all cases in the analysis the risk period was defined as the point from case closure to date of the screened in re-referral, or when the re-referral event did not take place, the child was considered a censored observation. Because initial cases took significantly longer to close in the treatment group than the comparison group, we also included length of case as a covariate. This variable was used to determine whether the length of an open case may act as a risk factor for screened-in referrals independent of whether individuals were in the treatment or comparison group. Finally, case (family) was modeled as a fixed effect to account for the clustering of children in families.

Results from the Cox regression model for screened-in re-referrals are shown in Table 15. This model used 812 out of the 858 individuals available for analysis. The ‘event,’ screened-in re-referral, occurred for 355 cases (43.7%); in other words, censoring occurred for 457 cases (56.3%). The overall model was significant, indicating that time to screened-in re-referrals was predicted by variables in the model ( $-2 LL: 4271.63, df = 11, p < .01$ ). Group (treatment vs. comparison) was not a significant predictor of time to screened-in re-referrals (using  $p < .05$ ), when adjusting for the other covariates included in the model. In other words, receipt of FGCs had no significant effect on time to a screened-in re-referral. Time to screened-in referral was significantly predicted by gender, length of case, caregiver history of childhood maltreatment, and case size. Specifically, children were more likely to have a shorter time to subsequent re-referral if they were female, had cases that were opened longer, had a caregiver with history of childhood maltreatment, and came from larger families.

**Table 15. Cox Regression Model Predicting Time to Screened-In Re-Referrals ( $n = 812$ )**

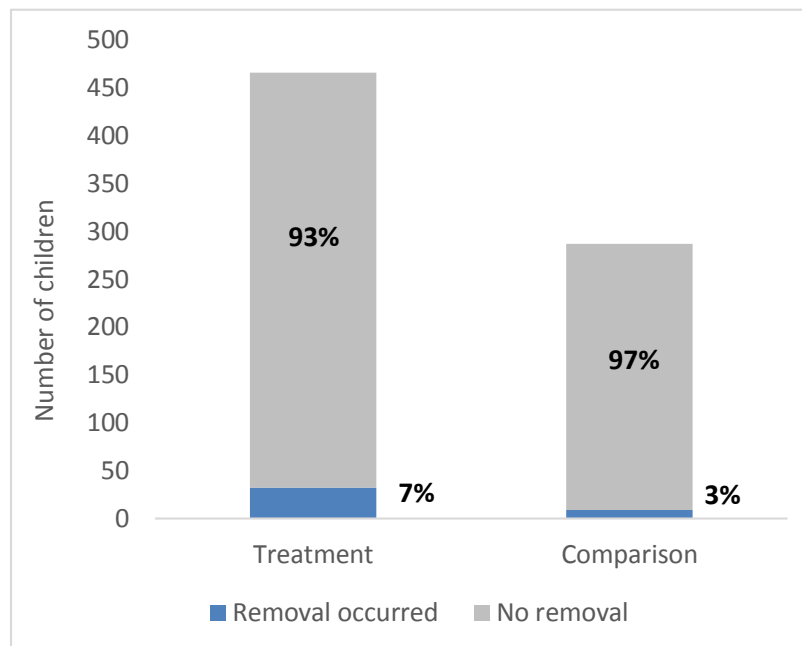
	Coefficient	SE	$\chi^2$	p	Hazard ratio
Group	-6.62	29.88	.05	.82	.001
Gender	-.24	.11	4.86	.03	.78
Age	.01	.01	1.42	.23	1.02
Hispanic ethnicity	.14	.12	1.39	.24	1.15
Number of days from case referral to case closure	.16	.03	40.16	< .01	1.18
Propensity score	-6.47	30.00	.05	.83	.002
Caregiver childhood history A/N	.42	.16	7.25	< .01	1.52
FAR	.21	.13	2.66	.10	1.23
Number of children	.19	.05	16.78	< .01	1.21
Cluster	.000	.00	.54	.46	1.00

Note: SE: standard error; Group: Treatment (coded 1) or control (coded 0) group, Gender: Male (coded 1) or female (coded 0); Hispanic ethnicity: Hispanic ethnicity (coded 1) or non-Hispanic ethnicity (coded 0); Number of days from case referral to case closure: Coded in increments of 60 days, creating 10 bins; the last bin reflected cases that were open for 600+ days (minimum number of days: 5; maximum number of days: 825). Caregiver childhood history A/N: Caregiver history of childhood abuse or neglect, coded as positive (1) or negative (0); FAR: Prior involvement in Family Assessment Response (FAR), coded as positive (1) or negative (0). Reference categories (in parentheses) are as follows: Group (Intervention), Gender (Female), Hispanic ethnicity (Hispanic), Caregiver childhood history A/N (positive history), FAR (FAR).

*Removals (out-of-home placements)*

Removal from the home after initial case closure occurred for 41 individuals within the matched sample. Figure 18 shows the number and proportion of children who were removed from the home separately for the treatment and comparison groups. Children in the treatment group were significantly more likely to be removed from the home ( $n = 32$ ; 6.9%) compared with children in the comparison group ( $n = 9$ ; 3.1%) ( $\chi^2 = 4.80[1]$ ,  $p = .03$ ). The average number of days to removal from case closure was significantly shorter in the comparison group than in the treatment group ( $t = -3.72[93]$ ;  $p < .01$ ). Specifically, removal from the home occurred at a mean of 185 days ( $SD = 129.09$ ) after the closure date for individuals in the treatment group ( $n = 32$ ), and at mean of 51 days ( $SD = 83.66$ ) after the closure date for individuals in the comparison group ( $n = 9$ ). Given persistent differences between treatment and comparison groups following the matching procedures, these results are reported for descriptive purposes only. They do not measure the impact of FGCs or FUMs on removals. Multivariate models are required to assess this relationship.

**Figure 17. Removal Occurrence Among Families  
(with  $n=466$  treatment and  $n = 287$  comparison children)**



Due to the small difference between treatment and comparison groups in the frequency of removals, the overall low frequency of removals, and the resultant lack of power, Cox regression models were not conducted. In an attempt to understand the higher rate of removals within the treatment group, however, we compared the children who were removed from the home in the treatment and comparison group on the type of placement they subsequently received. Children in the control group were most likely to enter kinship care (72.7%), rather than non-relative foster care (9.1%) or congregate care (18.2%). Rates of kinship care were lower among children in the treatment group who were removed from the home (43.8%), and roughly similar to rates of foster care (40.6%). However, these differences in type of placement across the treatment and comparison group were not statistically significant ( $\chi^2(2) = 3.88, p = .14$ ).

## B. Outcome Analysis - Texas Region 3

### Texas Region 3 Target Population and Sample

The Texas sample consisted of families open for Family Based Safety Services (FBSS) in Dallas and Tarrant Counties who were referred by their FBSS caseworker for a Family Group Conference (FGC). Each family in the population was randomly assigned as a member of the intervention (treatment) group or the control group based on the capacity of their respective county to provide FGCs at the time of the referral, though not all families accepted the offer to hold a family meeting. Since these families were receiving FBSS for indeterminate periods prior to assignment, key events (e.g., placements and re-reports) and time in service associated with these pre-assignment service periods were taken into account along with other data into the comparative analysis. However, evaluation outcomes were assessed following the initial FGC referral date, with the end point for follow-up defined as the end of the study (a maximum of 32

months). The outcomes analysis included 270 families from the treatment group and 272 families from the control group.

### Texas Region 3 Study Participant Characteristics

Characteristics of children and parents involved in the study, drawn from the SACWIS administrative data, are shown in

Table 16. Characteristics are shown for the treatment group and control group separately, as well as for the combined sample. Differences were tested for on a number of characteristics between the two groups to ensure that the randomization process functioned properly and potential biases in the selection of families into one group versus another were not present. Across the two groups, there were few notable differences, indicating that the randomization methodology was successful. No significant differences existed across age, gender, or race/ethnicity of children and parents between the groups.

The average age of children listed on the case file was 5.4 years old for the treatment group and 6.0 years old for the control group; for parents, the average age was 28.4 for the treatment group and 29.3 for the control group. Children were 49.1% female in the treatment group and 50.7% female in the control group, and parents were 48.5% female in the treatment group and 49.1% female in the control group. Children were 31.6% Hispanic, 34.5% African American, and 25.0% white (“other” and “unknown” categories omitted here) in the treatment group and 35.6% Hispanic, 33.7% African American, and 23.0% white in the control group. Parents were 26.5% Hispanic, 33.8% African American, and 33.1% white in the treatment group and 29.7% Hispanic, 33.8% African American, and 32.2% white in the control group.<sup>3</sup>

---

<sup>3</sup> Race and ethnicity categories were defined using the classification scheme of the U.S. Census Bureau, whereby Hispanic trumps race categories and other race categorizations indicate the selection of a single race as well as non-Hispanic.

**Table 16: Child and Parent Demographics, Dallas and Tarrant Counties (Texas)**

		Treatment		Control		$\chi^2$ (df)	p	Total	
		n	%	n	%			n	%
Children's Age	0-1 year	219	29.3	204	25.1	5.89 (3)	.12	423	27.1
	2-5 years	227	30.3	235	28.9				
	6-10 years	158	21.1	191	23.5				
	11+ years	144	19.3	184	22.6				
Children's Gender	Male	377	50.9	401	49.3	0.40 (1)	.53	778	50.0
	Female	364	49.1	413	50.7				
Children's Race/Ethnicity	Hispanic	236	31.6	291	35.6	4.19 (4)	.38	527	33.7
	African American	258	34.5	275	33.7				
	White	187	25.0	188	23.0				
	Other	34	4.5	27	3.3				
	Unknown	33	4.4	36	4.4				
Parent's Age	Under 21 years	73	12.5	61	11.0	7.05 (5)	.22	134	11.7
	21-25 years	155	26.5	145	26.1				
	26-30 years	153	26.1	134	24.1				
	31-35 years	114	19.5	107	19.3				
	36-40 years	56	9.6	52	9.4				
	Over 40 years	35	6.0	56	10.1				
Parent's Gender	Male	302	51.5	283	50.9	0.05 (1)	.83	585	51.2
	Female	284	48.5	273	49.1				
Parent's Race/Ethnicity	Hispanic	155	26.5	165	29.7	4.24 (4)	.38	320	28.0
	African American	198	33.8	188	33.8				
	White	194	33.1	179	32.2				
	Other	9	1.5	7	1.3				
	Unknown	30	5.1	17	3.1				

Beyond individual descriptive characteristics, case-level characteristics (see Table 17) were also tested for along a number of dimensions and observed only one significant difference: whether or not a referral to a community provider for domestic violence services was made. Treatment group families were referred to community domestic violence service providers more frequently than control group families (40.0% vs. 30.9%, respectively). Otherwise, case characteristics were similar across groups regarding race/ethnicity, previous involvement with child protective services, the type of alleged maltreatment, referrals to services, the number of children and adults on a case, and the average risk scores for families. The average number of children listed on the case was 2.8 for the treatment group and 3.0 for the control group, and the



average number of parents was 2.2 in the treatment group and 2.1 in the control group. Families' risk scores averaged 3.5 and 3.6 on caregiver capability for treatment and control groups respectively, 3.1 and 3.2 on child vulnerability, 3.3 and 3.4 on home/social environment, 2.6 and 2.6 on response to intervention, 3.4 and 3.5 on maltreatment, 2.9 and 2.9 on protective capability, and 3.1 and 3.2 on quality of care.

Further, 47.9% of families in the treatment group listed a child previously investigated for alleged abuse or neglect compared to 46.1% of control group families, and 31.5% of treatment group families listed a child previously *confirmed* as a victim of abuse or neglect compared to 31.0% of control group families. Among treatment group families, 54.1% were being investigated for alleged abuse on the initiating referral as compared to 57.6% among control group families. Eighty-eight percent of treatment group families were investigated on the initiating referral for neglect compared to 91.9% of control group families. Lastly, no significant differences were observed for community providers of mental health services across treatment and control group families (75.6% vs. 70.6%, respectively), DFPS-paid mental health services (4.1% vs. 6.6%), community providers of substance abuse services (60.7% vs. 57.4%), or DFPS-paid substance abuse services (8.9% vs. 8.8%).

**Table 17: Case-Level Demographics, Dallas and Tarrant Counties (Texas)**

		Treatment		Control		$\chi^2$ (df)	p	Total	
		n	%	n	%			n	%
Race/Ethnicity	Hispanic	59	22.2	66	24.5	1.22 (3)	.75	125	23.4
	African American	80	30.1	80	29.7			160	29.9
	White	67	25.2	58	21.6			125	23.4
	Multiple Races or Ethnicities	60	22.6	65	24.2			125	23.4
Prior Investigation-Child on Case Was Alleged Victim	Yes	128	47.9	125	46.1	0.18 (1)	.67	253	47.0
	No	139	52.1	146	53.9			285	53.0
	Yes	84	31.5	84	31.0			0.01 (1)	.91

Prior Investigation-Child on Case Was Confirmed Victim	No	183	68.5	187	69.0			370	68.8
Alleged Abuse of Child on Case	Yes	144	54.1	156	57.6	0.64 (1)	.42	300	55.9
	No	122	45.9	115	42.4			237	44.1
Alleged Neglect of Child on Case	Yes	234	88.0	249	91.9	2.27 (1)	.13	483	89.9
	No	32	12.0	22	8.1			54	10.1
Referral to DV Service Provider	Yes	108	40.0	84	30.9	4.92 (1)	.03	192	35.4
	No	162	60.0	188	69.1			350	64.6
Referral to MH Service Provider	Yes	204	75.6	192	70.6	1.70 (1)	.19	396	73.1
	No	66	24.4	80	29.4			146	26.9
DFPS-Paid MH Service Prior to Initiating Referral	Yes	11	4.1	18	6.6	1.73 (1)	.19	29	5.4
	No	259	95.9	254	93.4			513	94.6
Referral to SA Service Provider	Yes	164	60.7	156	57.4	0.64 (1)	.42	320	59.0
	No	106	39.3	116	42.6			222	41.0
DFPS-Paid SA Service Prior to Initiating Referral	Yes	24	8.9	24	8.8	0.01 (1)	.98	48	8.9
	No	246	91.1	248	91.2			494	91.1

	Treatment		Control		<i>t</i> ( <i>df</i> )	<i>p</i>	<i>n</i>	Total Mean (SD)
	<i>n</i>	Mean (SD)	<i>n</i>	Mean (SD)				
		2.8		3.0				2.9
Number of Children	267	(1.9)	271	(1.9)	1.33 (536)	.19	538	(1.9)
		2.2		2.1	-1.81			2.1
Number of Parents	267	(0.8)	269	(0.8)	(534)	.07	536	(0.8)
		3.5		3.6				3.6
Caregiver Capability Risk Score	267	(0.7)	271	(0.6)	1.17 (536)	.24	538	(0.7)
		3.1		3.2				3.2
Child Vulnerability Risk Score	267	(0.7)	271	(0.7)	1.34 (536)	.18	538	(0.7)
		3.3		3.4				3.4
Home/Social Environment Risk Score	267	(0.8)	271	(0.8)	0.33 (536)	.75	538	(0.8)
		2.6		2.6	-0.30			2.6
Response to Intervention Risk Score	267	(1.0)	271	(0.9)	(536)	.76	538	(1.0)
		3.4		3.5				3.4
Maltreatment Risk Score	267	(0.7)	271	(0.7)	0.47 (536)	.64	538	(0.7)
		2.9		2.9				2.9
Protective Capability Risk Score	267	(0.9)	271	(0.8)	0.83 (536)	.41	538	(0.9)
		3.1		3.2				3.1
Quality of Care Risk Score	267	(0.8)	271	(0.7)	1.78 (536)	.08	538	(0.8)

Note: SD = Standard Deviation; DV = Domestic Violence; MH = Mental Health; SA = Substance Abuse

## Texas Region 3 Outcome Analysis Results

Families were referred to the study from the in-home services population of Dallas and Tarrant County, Texas, when the child welfare agency deemed a family meeting appropriate. While not all families accepted the offer to hold a family meeting, for the purposes of this study, intention-to-treat (ITT) analyses are used. ITT analyses are inclusive of individuals who declined a family meeting, so as to remain conservative in the estimation of the treatment effect and avoid inflating the effect through the removal of families who decline the service and might be otherwise different than those who accept it. Administrative data from Texas DFPS was used to analyze effects of FGCs on re-referrals to the child welfare system and removals (out-of-home placements).

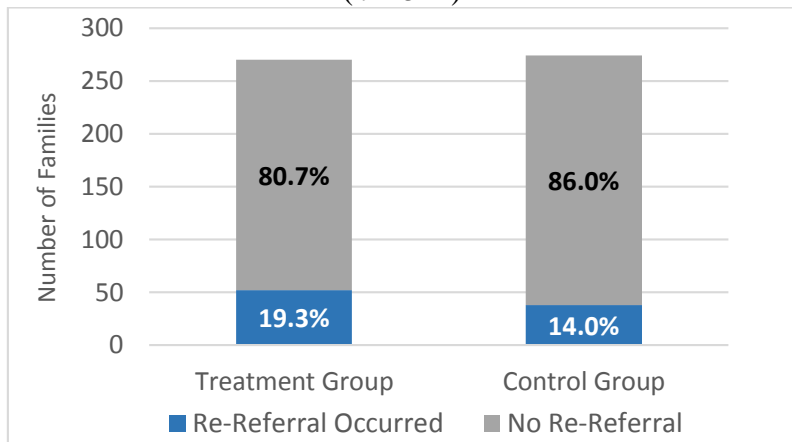
*Re-Referrals and Removals.* Utilizing the date of the initiating referral, the date of the FGC, and the dates of subsequent referrals to DFPS or removals of children from their homes, we assessed the impact of FGCs on re-referrals and removals. Because the dates of family meetings were only tracked for FGCs, it was necessary to construct an appropriate time frame for the opportunity for re-referral and/or removal among control group families. Treatment group families participated in an FGC, on average, 41 days following their referral to said meeting, implying that 41 days passed before a family had the opportunity for the meeting to “take effect.” In order to afford every family the same opportunity structure, we also applied this 41-day window, or grace period, to control group families and treatment group families who refused an FGC. We counted any re-referral to Texas DFPS or removal of a child from his/her home subsequent to an FGC or subsequent to the aforementioned 41-day window as an affirmative outcome in these analyses. Several re-referrals (4 treatment group, 6 control group) that fell before the FGC was held or within this 41-day window were *not* counted as re-referrals in this

analysis. Texas DFPS pulled the final administrative data for analysis on June 30, 2015; all families were tracked for a follow-up period of at least 14 months, with some families (e.g., those who entered the study in October 2012) being tracked for up to 32 months.

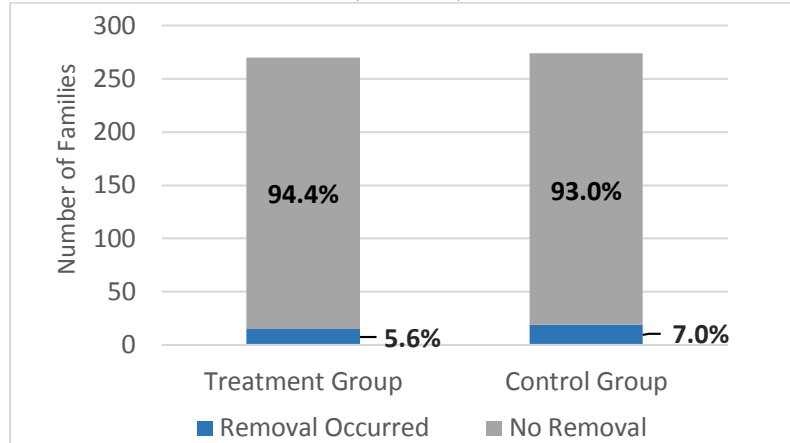
Figure 18 and

Figure 19 display the recurrence of referrals to Texas DFPS and removals of children from their homes, respectively, for the treatment and control groups. Notably, no significant differences existed for either outcome. For the treatment group, 52 families (19.3%) experienced a re-referral, whereas only 38 families (14.0%) in the control group did ( $\chi^2 = 2.74$ ,  $df = 1$ ;  $p = 0.10$ ). And concerning removals, 15 (5.6%) treatment group families had a child placed out-of-home, while 19 (7.0%) control group families did ( $\chi^2 = 0.47$ ,  $df = 1$ ;  $p = 0.49$ ). Neither of these differences were significant.

**Figure 18. Re-Referral Occurrence Among Families**  
( $n = 542$ )

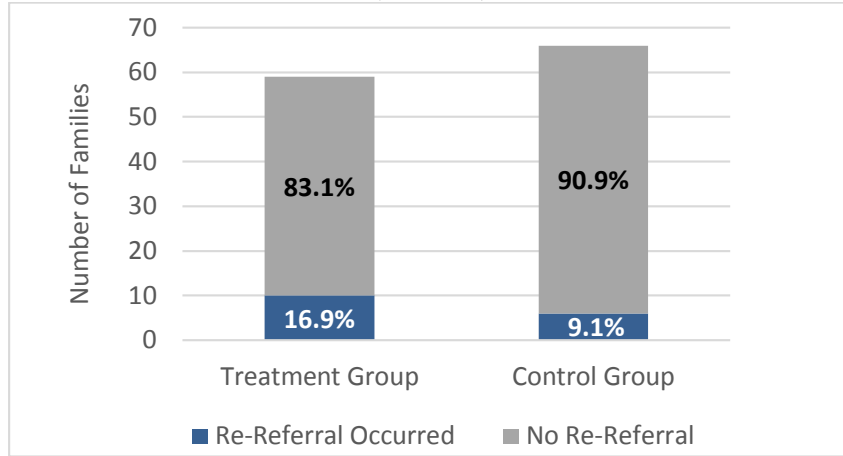


**Figure 19. Removal Occurrence Among Families**  
(*n* = 542)



The same treatment-control group analyses were conducted on each outcome (i.e., re-referrals and removals) for four different sub-groups of families based on race and ethnicity: Hispanic, Black, white, and multiple race/ethnicity families. The results for these subgroup analyses should be interpreted with caution as between group differences need to be examined in a multivariate context to elucidate which factors, among a variety of contributing factors (e.g., prior referrals), significantly predict re-referrals and removals. Figure 20 and Figure 21 display the recurrence of referrals to Texas DFPS and removals of children from their homes, respectively, *among Hispanic families* for the treatment and control groups. For the treatment group, 10 families (16.9%) experienced a re-referral, whereas only 6 families (9.1%) in the control group did ( $\chi^2 = 1.72$ ,  $df = 1$ ;  $p = 0.19$ ). Further, for the treatment group, 2 families (3.4%) had a child placed out-of-home, while 3 families (4.5%) in the control group did ( $\chi^2 = 0.11$ ,  $df = 1$ ;  $p = 0.74$ ); neither finding demonstrated statistical significance.

**Figure 20. Re-Referral Occurrence among Hispanic Families  
(n = 125)**



**Figure 21. Removal Occurrence among Hispanic Families  
(n = 125)**

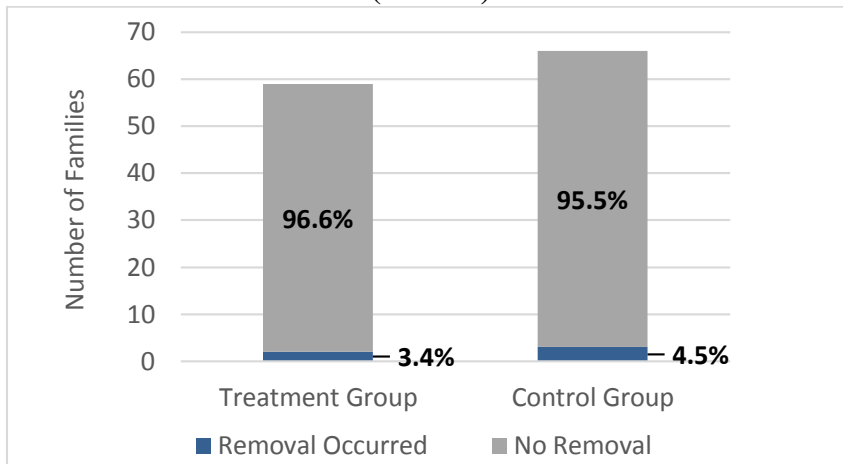
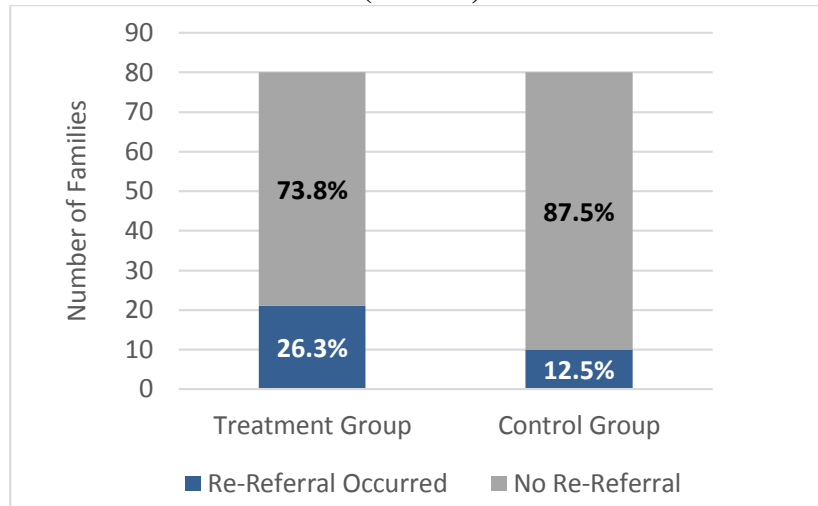


Figure 22 and Figure 23 display the recurrence of referrals and removals of children from their homes, respectively, *among African American families* for the treatment and control groups. For the treatment group, 21 families (26.3%) experienced a re-referral, whereas only 10 families (12.5%) in the control group did ( $\chi^2 = 4.84, df = 1; p = 0.03$ ). Notably, this difference was statistically significant, suggesting that African American families selected into the treatment group were more likely to be re-referred to Texas DFPS than African American families selected into the control group. Concerning removals, no statistically significant difference existed. For

the treatment group, 4 families (5.0%) had a child placed out-of-home, while 8 families (10.0%) in the control group did ( $\chi^2 = 1.44, df = 1; p = 0.23$ ).

**Figure 22. Re-Referral Occurrence among African American Families\***  
( $n = 160$ )



\*Significant difference at  $p < 0.05$ .

**Figure 23. Removal Occurrence among African American Families**  
( $n = 160$ )

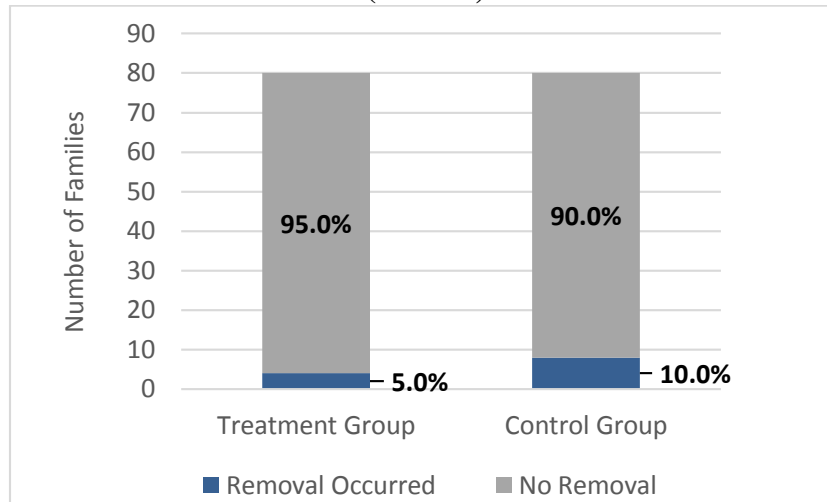
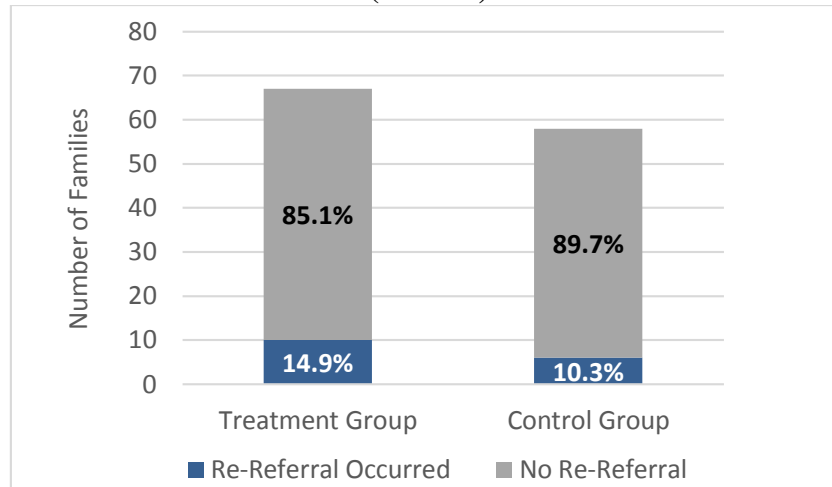


Figure 24 and Figure 25 display the recurrence of referrals and removals of children from their homes, respectively, among white families for the treatment and control groups. For the treatment group, 10 families (14.9%) experienced a re-referral, whereas only 6 families (10.3%) in the control group did ( $\chi^2 = 0.58, df = 1; p = 0.45$ ). Further, for the treatment group, 4 families



(6.0%) had a child placed out-of-home, while 2 families (3.4%) in the control group did ( $\chi^2 = 0.43$ ,  $df = 1$ ;  $p = 0.51$ ); neither finding demonstrated statistical significance.

**Figure 24. Re-Referral Occurrence among White Families**  
( $n = 125$ )



**Figure 25. Removal Occurrence among White Families**  
( $n = 125$ )

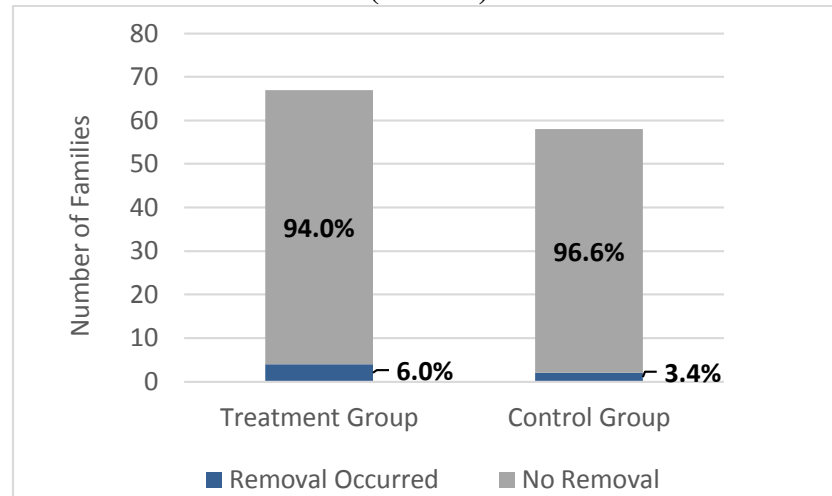
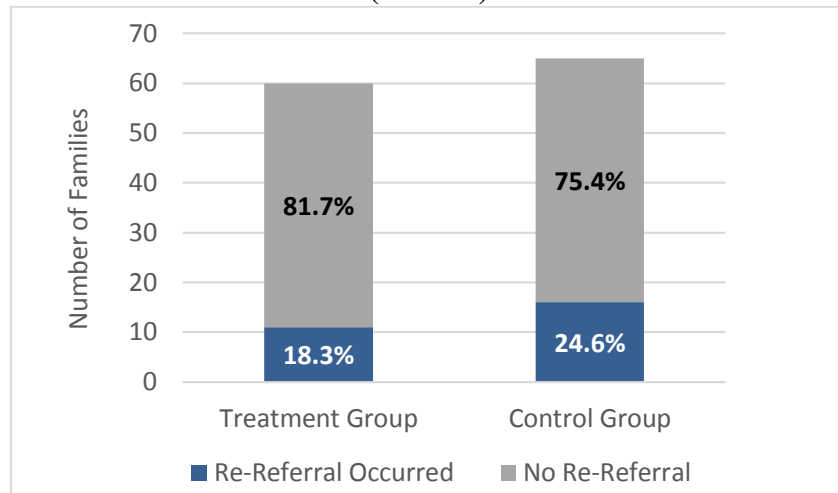


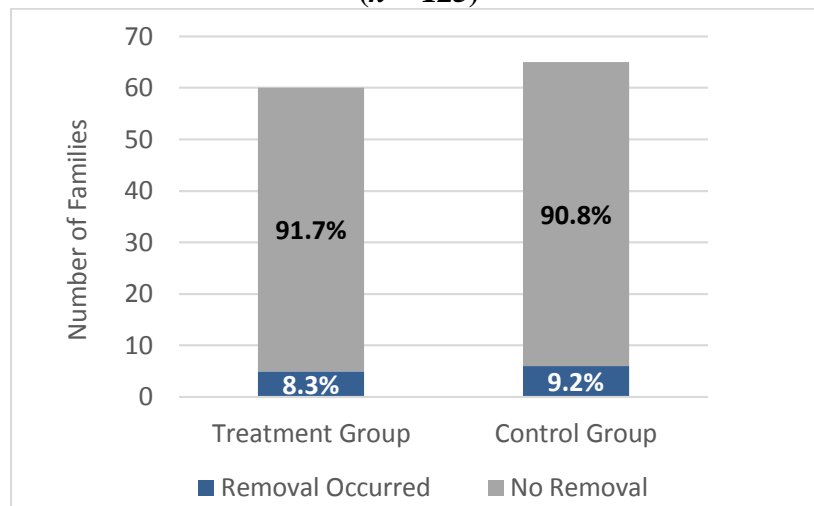
Figure 26 and Figure 27 display the recurrence of referrals and removals of children from their homes, respectively, among families with family members of *multiple race or ethnic groups* for the treatment and control groups. For the treatment group, 11 families (18.3%) experienced a re-referral, whereas 16 families (24.6%) in the control group did ( $\chi^2 = 0.73$ ,  $df = 1$ ;

$p = 0.39$ ). Further, for the treatment group, 5 families (8.3%) had a child placed out-of-home, while 6 families (9.2%) in the control group did ( $\chi^2 = 0.03$ ,  $df = 1$ ;  $p = 0.86$ ); neither finding demonstrated statistical significance.

**Figure 26. Re-Referral Occurrence among Multi-Racial/Ethnic Families**  
( $n = 125$ )



**Figure 27. Removal Occurrence among Multi-Racial/Ethnic Families**  
( $n = 125$ )



### C. Cost Analysis

The following cost analyses utilize information on the salaries, fringe benefits, and time allotment of personnel involved in the administration of Family Group Conferences. This *cost*

*allocation* is distinct from a cost-benefit or cost-effectiveness analysis as it only calculates the total annual cost of FGC provision (and the subsequent cost per meeting); this analysis cannot, by design, assess cost-effectiveness because the marginal increase in cost to provide FGCs over services as usual has not been measured and the lack of significant differences between treatment and control groups regarding re-referrals and out-of-home placements make it impossible to generate cost savings estimates resulting from the intervention. These cost estimates represent the cost to CPS and do not include the cost of FGCs to other systems (e.g., mental health counselors invited to participate in the meetings).

Table 18 presents the estimated personnel and non-personnel overhead costs for FGCs across Dallas, Tarrant, and Larimer Counties.<sup>4</sup> Costs were calculated for all personnel involved in the delivery of FGCs, including FGC coordinators, caseworkers, administrative staff, and their supervisors. The cost of FGC coordinators' time was calculated slightly differently for Texas and Larimer County. In Texas, the cost of FGC coordinators' time was calculated using their average annual salary multiplied by the number of coordinators and the percentage of time spent in FBSS FGCs under the purview of the NPLH project versus other family meeting types. In Larimer County, the cost of FGC coordinators' time was calculated using their hourly rate multiplied by the average number of hours worked per FGC and the number of FGCs held under the NPLH project in a given year. For both Texas and Larimer County, the cost of caseworkers' time was calculated using their hourly rate multiplied by the average time spent on FGCs (in the meeting and preparation time) multiplied by the number of FGCs under the purview of the NPLH project held in a given year. The cost of supervisors' time was calculated using their annual salary

---

<sup>4</sup> Only the costs of FGCs under the purview of the NPLH project are considered. The costs of *other* family meeting types (e.g., Family Team Meetings in Texas or Family Safety and Resource Team Meetings in Larimer County) are excluded.

multiplied by the FTE (full-time equivalent) allotted for supervising family meeting coordinators multiplied by the percentage of family meetings held that were FGCs under the purview of the NPLH project versus other family meeting types. And, the cost of administrative (clerical) staff was calculated using their annual salary multiplied by the FTE allotted to supporting FGCs under the purview of the NPLH project.

Fringe benefits and non-personnel costs were included to represent the entire cost of FGC delivery. Fringe benefit costs were calculated using the payroll rates used to budget benefit costs based on employment classifications. Non-personnel costs, which typically include contracted services, supplies and materials, durable equipment, rent and facilities, training, other direct costs, and indirect overhead, were estimated at 25% because not all of the non-personnel costs could be itemized. Twenty-five percent was derived from two seminal cost studies (Burwick et al., 2014; Corso & Filene, 2009) that demonstrated these costs average between 24-28% (though non-personnel cost estimates have been shown to range from 11% to 46%).

**Table 18. Total Estimated Annual Cost (to CPS) of FGCs across Dallas, Tarrant, and Larimer Counties**

	Site (County)		
	Dallas	Tarrant	Larimer
<i>CPS Personnel Salary</i>	\$ 224,091	\$ 170,560	\$ 92,885
<i>CPS Personnel Fringe Benefits</i>	\$ 67,227	\$ 51,168	\$ 18,577
<i>Non-Personnel Costs to CPS</i>	\$ 72,830	\$ 55,432	\$ 27,866
<b>Total:</b>	\$ 364,148	\$ 277,160	\$ 139,328
<b>Number of Meetings Held Annually</b>			
	93	119	151 <sup>a</sup>
<b>Estimated Cost per Meeting:</b>	\$ 3,916 <sup>b</sup>	\$ 2,329	\$ 923

<sup>a</sup> The number of meetings held in Larimer County is based off of an estimate of the average number of family meetings held per month multiplied by the percentage of family meetings that are FGCs. Approximately 12-13 FGCs were held each month.

<sup>b</sup> The higher *per meeting* cost of FGCs in Dallas County is largely driven by the significantly greater percentage of time Dallas staff devote to FBSS FGCs versus other family meeting types.

## D. Additional Analyses

Two additional analyses considering implementation issues have been conducted to date. An examination of Texas Pre- and Post- Caregiver Survey data and an analysis of the items on the Case-Specific Questionnaire were conducted. A brief summary of each is presented here.

### Caregiver Follow-Up Survey Analysis

Additional analyses were conducted examining data from the Caregiver Pre- and Post-Test/Follow-Up surveys in Texas Region 3 (as this was the only site where posttest data was collected, as discussed above). Caregiver Follow-Up surveys were sent to all caregivers in both the treatment and control groups, who had completed a Caregiver Pretest, following the closure of their case. The survey asked questions regarding a target child's child psychosocial behaviors, the caregiver's emotional response to their first and last contacts with CPS staff, protective factors, services received, and satisfaction measures. Of the 259 original Caregiver Survey Pretest respondents from Texas, 75 follow-up surveys were obtained, representing a 29% response rate. Caregivers from the control and treatment groups were almost equally represented in the final pre-post sample; thirty-nine, or 52% of the post-test surveys were received from respondents assigned to the treatment group, while 48% (n = 36) came from the control group. There was no difference in the likelihood of returning a follow-up survey for the control and treatment group caregivers. Further, a response bias analysis found that the post-test respondents were no different from the pre-test only respondents on demographics and psychosocial scale data collected in the pre-test survey.

*Control vs. Treatment Group Results.* T-tests and ANOVAs were employed to examine whether caregivers from the treatment and control groups differed from one another on their answers to questions on: protective factors, their emotional responses to the first and last meeting

with CPS staff; whether or not they received any of 17 services; dynamics of their interactions with CPS staff; and their satisfaction with their experience. Other than the control group reporting higher levels of receipt of mental health services (43% vs 15%, respectively;  $t(62) = 2.626, p = .011$ ), no other significant differences were detected. Comparisons of respondents on their pre-post SDQ scores indicated that while overall changes in scores over time were not statistically significant ( $F(1) = 1.458, p = .234$ ), the treatment vs control groups' changes in scores over time were different and statistically significant ( $F(1) = 4.413, p = .042$ ). Specifically, while on average the control group's scores *decreased* (i.e., reportedly improved) between the pre- and post-test (from 4.6 to 3.5, or from "very high" to "high"), the treatment group families reported an overall *increase* in average SDQ scores (from 2.7 to 3.0, or from "close to average" to "high") (See Appendix C for references describing this scale).

*Treatment Group Meeting Recipients vs. Others.* While the above analysis utilized the ITT design, some treatment group families did not actually receive a family meeting. Thus, a sensitivity analysis examining the same questions about caregiver emotional responses, services usage, satisfaction outcomes, and protective factor and SDQ scores was employed to compare responses from treatment group families who actually received a meeting with all other study participants (i.e., control group families *and* those treatment group families that did not experience a meeting during the study period). The *t*-tests, ANOVAs, and repeated measures analyses detected only one difference between the groups related to the protective factor concerning attachment. Here, significant effects for time ( $F(1) = 4.039, p = .049$ ) and for an interaction between time and group status ( $F(1) = 5.369, p = .024$ ) were identified. On average, caregivers in the treatment group who experienced a family meeting were associated with an increase in their protective factor score reflecting attachment (from 6.1 to 6.6), compared to the

other study participants, whose average scores were fairly steady over time (from 6.4 to 6.6). That said, the post-test scores for the two groups on this factor were virtually identical and indicated that both sets of respondents felt they demonstrated this protective factor “Very Frequently.”

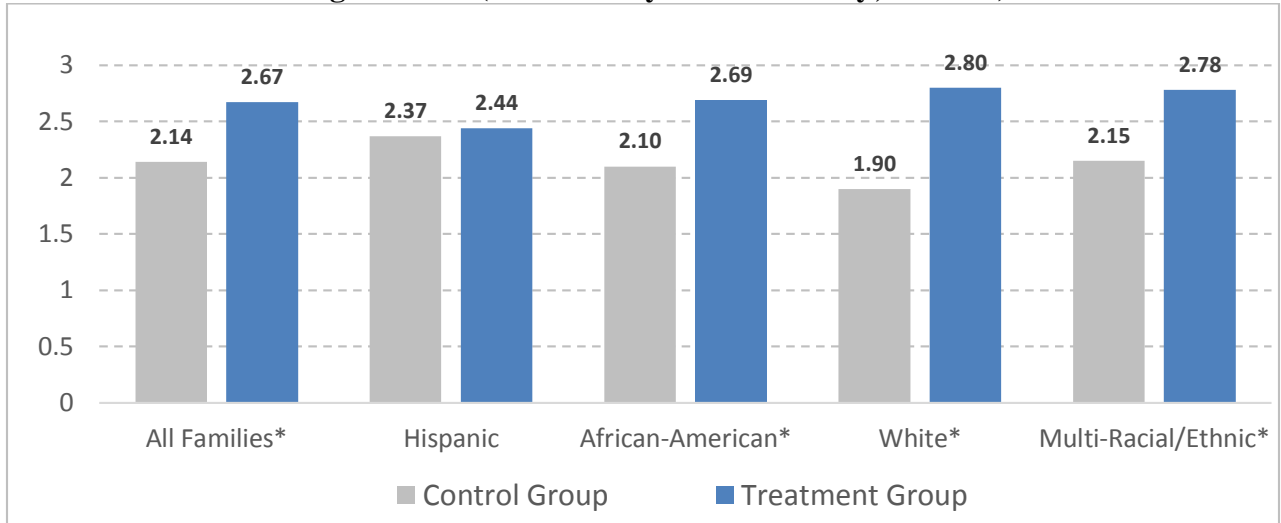
### Case-Specific Questionnaire – Texas Region 3

*Improvement Experienced in Social Supports.* At case closure, child welfare staff were asked to complete a case-specific questionnaire describing the service needs, the level of service provided, and the degree of improvement experienced by families on a variety of domains, including: material needs, substance abuse, parent developmental disability, parent physical disability, child developmental disability, child physical disability, parent mental health, child mental health, parenting skills, domestic violence, child education, medical care, and social supports. Few differences existed across these domains, however, one domain exhibited a consistent and striking difference: social supports (or more specifically, the degree of improvement families experienced in social supports [e.g., extended family, friends, neighbors, etc.]).

Figure 28 displays the degree of improvement in social supports noted by child welfare staff for families ( $n = 300$ ) in the treatment and control groups in Texas (both Dallas and Tarrant County combined). Similar analyses were not conducted for Larimer County due to the significantly lower response rate for the site’s Case-Specific Questionnaires (34%; see Table 6). Case-Specific Questionnaires were not completed for all families (e.g., families whose case with DFPS was not closed). The analyses revealed that treatment group families had a significantly higher degree of improvement in social supports than their control group counterparts. Further, as demonstrated in Figure 28, this significant difference (as indicated by an \*) existed among

Black families, white families, and multi-race/multi-ethnic families, though not among Hispanic families.

**Figure 28. Average Rating of Degree of Improvement in Social Support Experienced among Families (Total and by Race/Ethnicity; N = 300)**





## VIII. Evaluation Discussion

### A. Challenges

There are three overarching challenges that posed problems to the evaluation that arose over the duration of the project: low study referral rates, low survey response rates, and data quality issues. In terms of referral rates, none of the three sites hit the targets that were set in consultation with leadership in each site and informed by historical FGC referral rates and capacity estimations for FGDM staff. The difficulty in referring eligible cases to the study was such that the outcome evaluation ultimately had to be abandoned in South Dakota (as detailed in Appendix G). While not as extreme in Texas or Colorado, lower-than-anticipated referrals in both sites, and particularly around FGCs in Colorado may have impacted the lack of significant findings discovered in either site due to insufficient sample sizes.

One of the sources of this problem was the strong influence of typical practice expectations. Typical practice in both sites is for workers to utilize discretion in making an FGC referral. As an example of how this may have affected study enrollments, focus group findings informed the evaluation team that, in Texas, resistance to the RCT, stemming from a lack of discretion during the study owing to the study randomization processes, interfered with worker determination regarding who would receive an FGC. This resulted in some workers not referring to the study at all and others holding their own, informal family meetings in situations where their cases had been randomized into the control group. The impact of this problem was not measurable given the unexpected nature of the problem, but no doubt contributed to the smaller than expected study samples in Texas.

From an evaluation standpoint, these anecdotally identified effects on the study design raise issues regarding the ability of studies, such as this, to fully understand the impact of worker

biases. It has been recognized that with respect to FGDM practice in general that worker buy-in is often problematic. Biases regarding the effectiveness of FGDM, concerns about the time to carry out FGDM processes, workers not believing in family capacity to plan for themselves, and other types of biases are all possible contributors to issues with FGDM implementation and have been noted in other studies. However, it is also clear that these biases are not easily addressed by evaluation procedures, and in fact, these procedures may result in unintended consequences as the example from Texas suggests.

Regarding the low survey response rates, and as aforementioned, this issue was most impactful in Larimer County for the Caregiver Survey in that the survey had to be dropped from the propensity score analysis process as well as the pre-post outcomes analysis for protective factors and child behavior. However, lower-than-target response rates were experienced for most surveys in both Texas and Colorado which leads not only to issues around statistical power, but also to concerns about respondent bias in that the data gleaned from these surveys may not be representative of the study population as a whole.

Finally, data quality issues particularly in Larimer County had significant implications for who was ultimately included in the study as various data sets received from that site contained data on conflicting cases and clients. As a result, a quality assurance process was implemented by the evaluation team that conservatively culled the full list of potential study participants (found in any of the files received) to allow only for those who appeared in all of the relevant files to be included in the study. This resulted in approximately 200 potential study participants being culled from the matching process and ultimately the study as a whole. Although this compounded the aforementioned issues around low referral rates, the evaluation team was

concerned about issues of data quality and contamination if those ‘unknown’ cases were allowed into the study.

## B. Limitations

As in any study, but especially one as complex as this, there are limitations in the design, data collection, and execution of the evaluation. Some of those that are most salient for this study are described in this section. First, for some aspects of data collection, response rates were low, which can bias study results if non-responders are different than responders. Second, while the sample size was large for a study of this kind, it was lower than projected, and, in some cases, the result is a lack of statistical power to be able to detect significant differences in low frequency occurring outcomes, such as child removals.

In terms of outcomes, results did not reveal significant differences in the TX and CO sites for treatment and comparison/control groups for re-reports for the sample as a whole. Several possible explanations can contribute to this lack of significant findings. First, as stated previously, low frequency events pose a challenge to detecting effects. It is possible discernable effects would emerge if the sample size were larger. Second, in CO in particular, multiple types of family meetings were offered, which may dilute the differences between the two groups in the intervention received. Similarly, we have some reason to believe that worker concern about families being assigned to the control group in TX may have resulted in offering ‘unofficial’ family meetings given their commitment to the practice in some cases. This could bias the control group. Finally, the benefits of family meetings may take more time to be realized or may be present in unmeasured outcomes. For example, significant impacts on social support for families that received FGCs in TX were observed. The benefits of social support on the ultimate outcomes of interest may manifest over a longer time period beyond the duration of this study

and also affect other outcomes. Such outcomes may include child well-being, which was not measured for this evaluation.

The PSM methodology used in Larimer was based on the available administrative data, other variables which may have achieved more precise matching were not available. The descriptive results suggest that the control group may not have been equivalent to the treatment group even after the best matching procedures were employed.

## IX. Conclusions

Despite the limitations, it could be that FGCs are not a cost-effective approach for this stage of service and/or are contingent on other factors that lead to effective outcomes – such as the quality and availability of services in the community. As Parker, Bush and Harris (2014) note, there should be caution in thinking about when and how much of an effect should be expected for a given intervention. It may be that families need more than one family meeting for effects to be produced, that concerted efforts supporting reinforcement of a family’s resources (social, material, or otherwise) may be a critical step towards successful outcomes, that resources in the community may not be of sufficient quality or availability to achieve the desired outcomes, and/or that the impact may not be discernable for years to come. In summary, these findings need to be interpreted in the context of the sites themselves and other research to better make sense of them.

### A. Major Takeaways

- Low frequency events, such as child removals, can pose a challenge to detecting significant effects.
- Fidelity index scores indicate overall favorable responses from both family and professionals, but professionals and facilitators generally had slightly higher fidelity ratings.
- Orientation toward child safety versus family preservation depended on job type, years of experience, and shared vision.
- Higher ratings of perceived FGC effectiveness depended on worker type, perceptions of local services, and belief in families’ ability to construct plans to address issues.
- No statistically significant differences between treatment and comparison/control groups in screened-in re-referrals were found in Larimer or Texas.
  - In the case of Larimer, FSRTs may produce some of the same results as FUMs and FGCs for this stage of service (as indicated by lack of significant findings between treatment and comparison groups in Larimer).

- No statistically significant differences between treatment and comparison groups in removals were found in Larimer or Texas. Impacts from intermediate outcomes such as more comprehensive, family-centric plans and increases in social support may take longer to be reflected in long-term outcomes.
- Other unmeasured benefits may be present, such as child well-being, positive impacts on agency culture, consistency with agency values, and family engagement over time.
- Propensity score matching is a very complex methodology with important limitations. For this study important questions emerged regarding the difficulty of obtaining appropriate control samples and matching variables.

In summary, this large-scale, multi-site, multi-method evaluation showcased some relevant and unexamined findings related to family meetings. A new fidelity index was created and tested for this project. The FGCs in TX and CO showed favorable fidelity to the model from both family members and the child welfare staff who participated. And, the psychometric analyses performed on the instrument in this project have informed further development of this tool, which will be a useful asset to the field in both practice and research.

Staff attitudes and buy-in are such a critical component of effective practice, as has been demonstrated from studies of implementation science and organizational culture. The worker survey findings shed light on staff attitudes as pertaining to FGCs. Higher ratings of FGC effectiveness depend on whether workers carry a caseload, perceptions of local services, and belief in family abilities to construct plans to address issues. Similarly, worker orientation to child safety versus family preservation also depends on whether or not a worker carries a caseload, years of experience, and perceptions of a shared vision within the organization.

In terms of outcomes, in TX, it was found that families receiving an FGC were perceived by their caseworkers as having a greater amount of social support at the close of their case than families who did not receive an FGC. The importance of social support in the context of child

maltreatment has been well established (Thompson, 2015). This finding is an important one. However, significant effects were not found on the likelihood of a re-report or an out-of-home placement for the full samples in TX. Possible explanations are provided below. Two racial and ethnic differences in these outcomes were detected, but, as stated in the report, they need additional exploration to rule out spurious results.

## B. Contributions

As exemplified by comparisons with previously conducted research on many child welfare interventions, large-scale and rigorous experimental and quasi-experimental designs in child welfare are relatively rare. Specifically, this study contributes new knowledge and understanding of various family meeting models, as implemented by public child welfare agencies in different locations. Obstacles such as cost, time, and organizational culture and individual attitudes around the ethics of such designs are not insignificant for this field. This evaluation showcased how such a study could be conducted on a commonly-used practice in child welfare in an ethical and relatively efficient way. Federal, foundation, and local funding and established guidelines need to continue to push for and support this type of research to advance the field. This evaluation also highlights the use of SACWIS data for evaluation purposes, which leads to greater cost and time efficiency in carrying out evaluations of this sort and is aligned with recent efforts in human services fields to link and use “big” data for predictive modeling.

From the start of the project, until likely well beyond the project end-date, we have produced peer-reviewed publications on the findings. To date, three journal articles have been published, with one more currently under review. Publication in peer-reviewed journals lends credibility to the research findings, is a standard for contributing to the evidence-base and to be included in evidence-based clearinghouses. It is how knowledge gets transmitted through social worker

training and education in universities. It is also the basis on which further evidence and research scholarship is built.

The three project agencies benefited directly from their participation. They learned about their practice, their effectiveness, limitations, strengths, and organizational culture of family-centered practice. And, they deserve to be recognized as a learning organization and a contributor of additional knowledge for the field for their willingness to participate in such an evaluation. Finally, it is an ethical imperative for the field that commonly-used practices, programs, and interventions be rigorously evaluated. Families and children deserve services that both reflect the philosophy of an agency regarding how best to meet their needs and that are demonstrated to either be effective, or at a minimum, not harmful. Child welfare practice has a long way to go in terms of increasing the menu of evidence-based programs and their adoption in the field, but the Children's Bureau's, Casey Family Programs,' and the sites' support of this evaluation highlights their commitment to fulfilling this need in order to ultimately improve outcomes for children and families.

### C. Next Steps

The data collected in the NPLH project contain more information than can possibly be analyzed within the scope of the project. Multivariate analyses may reveal different patterns of results for different types of families. Linking worker, fidelity, and administrative data will lead to a wealth of analyses that examine how workforce and practice characteristics influence outcomes. Different outcomes can be examined, as well, such as substantiated re-reports, the length of placements in out-of-home care, type of placements, and services received. The investigators on this project are committed to continuing to examine and publish research related to these topics to expand the knowledge base around FGCs and FUMs.



## X. Recommendations

The project team provides the following recommendations to those implementing similar projects, to the Children’s Bureau, and to the broader child welfare field. These recommendations are formulated from our collective implementation and evaluation experiences with this project, over the past four years and are found in Table 19, below.

**Table 19: Recommendations**

<b>For administrators of future, similar projects</b>	
a. Automate the referral process	Even when policy exists that require workers to refer families to meetings, based on a core set of criteria, we have found that some workers are less likely to refer. One solution, particularly for an RCT designed evaluation project, is to automate the referral which removes worker discretion.
b. Launch data collection as quickly as possible	Even though the evaluation team used data from previous years to determine and project the sample size, and even though we had originally planned for an 18-month data collection period, we found that the projected referrals did not keep pace with actual referrals. This impacted our final sample size. Therefore, our recommendation is to consolidate the pre-implementation phase into a shorter timeframe which would allow for a longer data collection period.
c. Target more families to receive the intervention to ensure adequate sample size for the evaluation	Related to the recommendation above, we encourage other projects to identify their sample, based on various sources of data, and then determine if there are ways to increase the number of families referred, if staffing patterns permit. This could mean adapting the criteria or creating an automated referral source.
d. For projects implementing a randomized control design, monitor, as closely as possible, the practices within the control group to assess the efficacy of the RCT.	At our RCT site, we found that workers did not particularly like the notion that the families they believed would benefit most may not receive the intervention. Through our focus groups, we learned that some workers decided to hold their own “family meetings” for some families assigned to the control group (not to receive the FGC). The degree to which this occurred is unknown, but does serve as a reminder for evaluators to monitor assignment

	processes to ensure that the rigor of the RCT is not compromised.
e. Use project data, throughout and at various points, as continuous quality improvement processes.	Providing key stakeholders (staff, FGDM coordinators, supervisors and others) with data throughout the evaluation could have had many benefits, including: increasing staff buy-in to the benefits of the evaluation; and providing staff with the data that could have been used to review and possibly improve their practice. Continuous feedback loops should be embedded into these types of evaluation.
<b>For project funders (The Children’s Bureau)</b>	
f. Lengthen the grant period from 3 to 5 years	As has been noted by other grantees in this cluster, three years to implement FGDM and evaluate it doesn’t appear to provide sufficient time to enroll a large enough sample of families and follow them for a sufficient period of time post-family meeting to find differences in short- and long-term outcomes between groups.
g. Adopt the international definition of family group decision making, requiring grantees to implement the six core elements in practice.	In 2010, guidelines on the core elements of FGDM were promulgated by an international committee of experts. This 2011 Family Connections funding announcement provided great latitude in defining FGDM, which resulted in such variations in practice that clearly did not meet the core values, principles, and values of the FGDM process. This has great implications for the child welfare field, in that, there are now some evaluations of FGDM with limited to no model fidelity. The evaluative findings—both positive and negative—will be connected, however, to FGDM
h. Require researchers and evaluators to provide statistical power analyses	It is critical that power analyses are conducted prior to data collection to ensure that the proposed study has adequate power for testing the study hypotheses. For evaluations of FGDM, the required sample size will vary in relation to the prevalence rate of the outcome of interest (e.g., removals). Power analyses should take into account potential challenges such as difficulties in recruitment and sample attrition.
i. Encourage researchers in this grant cluster to explore outcomes beyond safety and permanency, meaning beyond re-referrals, repeat maltreatment and placement.	The funding announcement required evaluators to measure the standard child welfare outcomes of safety, permanency and well-being which tends to be the most elusive of the three. ACF also provided latitude to grantees to select instrumentation to meet their projects’ core questions. We believe, in addition to these outcomes, there are more nuanced analyses that could benefit the child welfare field. For example, are members of the family group more likely

	to seek help because of their participation in family meetings? From an interpretation angle, should we expect re-referral and placement rates to increase because families are providing an extra layer of surveillance? What if the child welfare field was to review those outcomes as positive indicators?
j. Acknowledge survey response rates and the difficulty within child welfare evaluations of getting ample responses from families	Evaluations such as these invest a significant amount of time and fiscal resources to collect data/surveys from staff and families, yet typically at the end of the evaluation have low response rates, particularly for families. This needs to be taken into account in the funding announcement and additional technical assistance provided to boost response rates, if there are tested, successful strategies.
k. Review suggested measures/instrumentation through the social desirability lens	The child welfare population is not a voluntary population; even families receiving non-court ordered in-home services may not perceive that they have any choice in their involvement with child welfare. The results of the Protective Factors Survey would lead the readers to believe that the study's parents/caregivers are parenting well, despite their involvement with child welfare. Or, it could be interpreted that parents, given their involvement and perhaps perceived threats that can accompany their involvement, may be less likely to provide an honest appraisal of their parenting strengths and struggles.
<b>For the child welfare field</b>	
l. Critically examine the purpose of various family meeting models that are implemented.	Some family meeting models are positioned in both timing and process to ensure that agency-decisions include family and others recognize the family as the nexus-construct for decisions about their families. Having clarity about the various models, their purposes, their core elements, and how they build on one another is very important for both agency professionals and family members, who oftentimes partake in multiple meetings.
m. Invest in collecting data from families about their experiences with family meetings.	Similar to the work of other evaluators of FGDM fidelity, this evaluation showed that agency professionals' perceptions of empowerment, leadership, inclusion and respect may likely be different than what the family experiences. Even without a formal evaluation process, soliciting family input about their experiences with FGDM (both fidelity and satisfaction) can support an agency in improving these practices and likely provides the most-illustrative guidance available. Fidelity measures can be used to support supervision, coaching and program enhancement activities.

<p>n. Streamline the number and type of family meetings being implemented by the agency/community.</p>	<p>Many child welfare agencies implement various family meeting processes at different decision making junctures. Involving any family—even in a limited form—at every possible decision making point is good practice. However, we have found that agencies tend to create complicated, even bureaucratic family meeting structures that even the most seasoned professionals cannot fully distinguish. There is a tendency for agencies that have two or more family meeting processes in their service delivery system to prioritize family meeting processes that are expedient for the agency but not fully inclusive of the family group. In addition, a continuum of meetings, particularly when the elements are not that distinguishing, are likely to confuse even the most sophisticated consumer.</p>
<p>o. Ensure that FGCs are not relegated to being implemented at a point in the case when there is no “live” decision to be made.</p>	<p>FGCs are a decision making process whereby active decisions should be made by the family group with support of the agency providers. It appears as if some FGCs are held at the point of case closure, likely as a way to create plans that wrap support around the family.</p>
<p>p. Use the evaluation results to alter practices, as warranted, in strategic and informed ways. Constantly be thinking about application and what additional information is needed to inform decision making.</p>	<p>At a minimum, agencies implementing any type of family meeting can, with relative ease and minimal expense, implement fidelity surveys that capture the perspectives of all participants to determine the extent to which the family meeting aligns with the principles and core practices. Analyzing fidelity data using practice constructs—like we did in this evaluation—provides useful information to the many professional stakeholders who are positioned to adjust practices.</p>

## References

- Allan, H., Fleming, L., Letts, E., Maher, E., Merkel-Holguin, L., & Deluca-Forzley, D. (2014, May 2). *Family group decision making for in-home services: Issues in implementation and fidelity*. Presentation at the 19<sup>th</sup> National Conference on Child Abuse & Neglect.
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *36*, 399-424. doi: [10.1080/00273171.2011.568786](https://doi.org/10.1080/00273171.2011.568786)
- Burwick, A., Zaveri, H., Shang, L., Boller, K., Daro, D., & Strong, D. A. (2014). *Costs of early childhood home visiting: An analysis of programs implemented in the Supporting Evidence-Based Home Visiting to Prevent Child Maltreatment Initiative*. Princeton, NJ: Mathematica Policy Research.
- Corso, P. S., & Filene, J. H. (2009). Programmatic cost analysis of the Family Connections program. *Protecting Children*, *24*(3), 78-88.
- Parker, R., Bush, J., & Harris, D. (2014). Important methodological issues in evaluating community-based interventions. *Evaluation Review*, *38*(4), 295-308.
- Texas Department of Family and Protective Services. (February 2014). DFPS 2013 Data Book. Retrieved July 28, 2015 from [https://www.dfps.state.tx.us/About\\_DFPS/Data\\_Books\\_and\\_Annual\\_Reports/2013/default.asp](https://www.dfps.state.tx.us/About_DFPS/Data_Books_and_Annual_Reports/2013/default.asp).
- Thompson, R. A. (2015). Social support and child protection: Lessons learned and learning. *Child Abuse & Neglect*, *41*, 19-29.
- U.S. Census Bureau. (May 2015). State & county QuickFacts: Dallas County, TX. Retrieved July 28, 2015 from <http://quickfacts.census.gov/qfd/states/48/48113.html>.

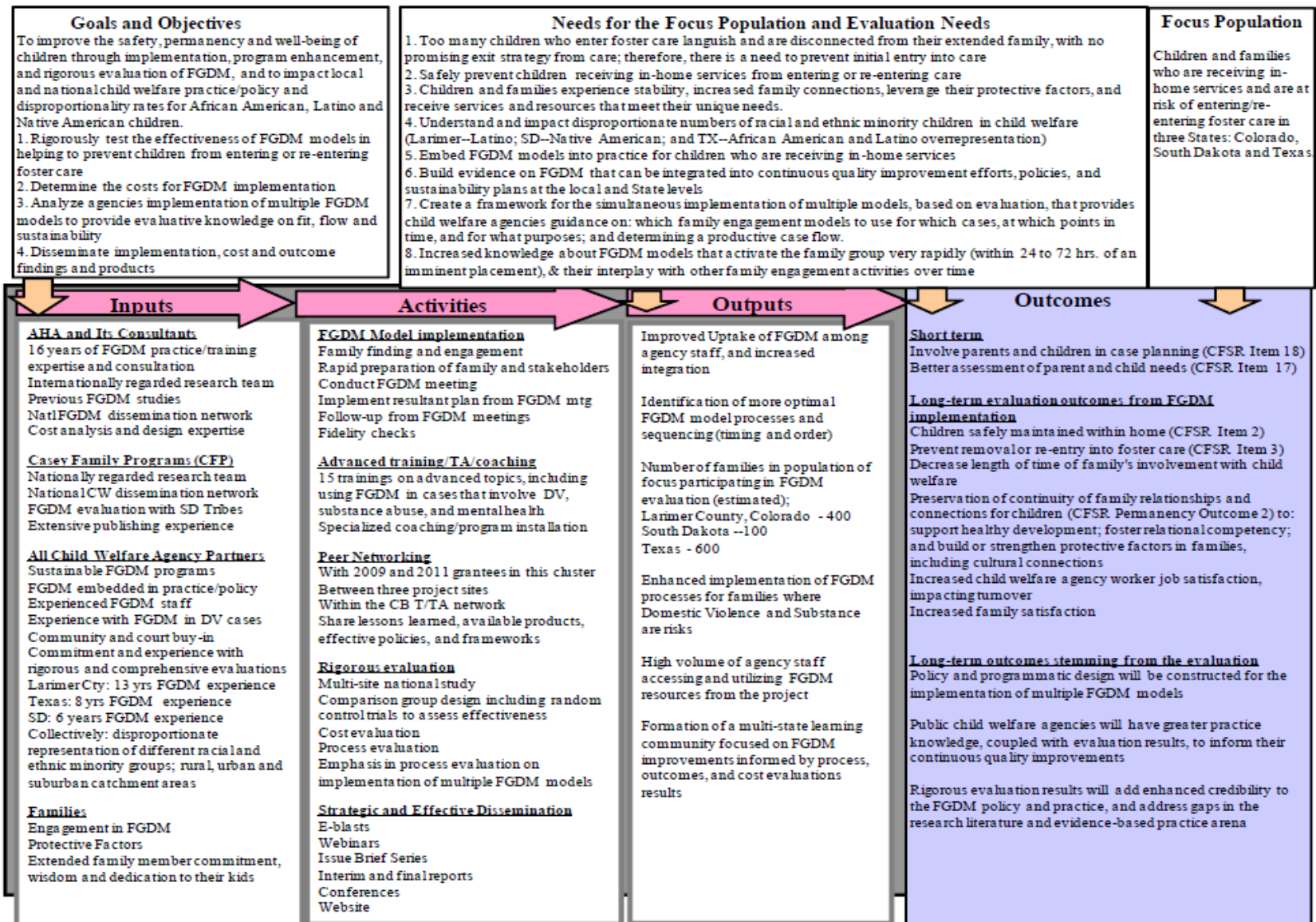
U.S. Census Bureau. (May 2015). State & county QuickFacts: Larimer County, CO. Retrieved July 28, 2015 from <http://quickfacts.census.gov/qfd/states/08/08069.html>.

U.S. Census Bureau (May 2015). State & county QuickFacts: Pennington County, CO. Retrieved July 28, 2015 from <http://quickfacts.census.gov/qfd/states/46/46103.html>.

U.S. Census Bureau. (May 2015). State & county QuickFacts: Tarrant County, TX. Retrieved July 28, 2015 from <http://quickfacts.census.gov/qfd/states/48/48439.html>.

Williams, J., Merkel-Holguin, L., Allan, H., Maher, E., Fluke, J., and Hollinshead, D., (September 2015), Factors associated with staff perceptions of the effectiveness of family group conferences. *Journal for the Society of Social Work and Research*, 6(3), 343-366.

## Appendix A. NPLH Logic Model



## Appendix B. NPLH Training and Technical Assistance

<b>TEXAS REGION 3</b>					
<b>TRAINING OR SERVICE (Coaching or TA)</b>	<b>Training Days</b>	<b>Audience (Supervisors, caseworkers; FBSS, FGDM)</b>	<b>Dates</b>	<b>Number of Participants</b>	<b>County (Dallas/ Tarrant / Both)</b>
Preparation Is Crucial: The Nuances to Coordinating Family Conferences	2	FGDM staff and supervisors	4/3-4/4	25	Both
Family Group Decision Making: At First Glance	2 1-day sessions	FBSS Caseworkers & Supervisors	4/17-4/18	70	Tarrant
Family Group Decision Making: At First Glance	2 1-day sessions	FBSS Caseworkers & Supervisors	4/19-4/20	70	Dallas
The Referring Worker: A Key Role in the FGDM Process	5 1-day sessions	FGDM & FBSS Caseworkers and Supervisors, Program Administrators and Directors	5/14-5/18	200	Both
Transforming Conflict into Partnership in the FGDM Process	2 1-day sessions	FBSS Caseworkers & Supervisors	7/18-7/19/12	100	Dallas
NPLH Evaluation Training	4 .5-day sessions	FGDM & FBSS Caseworkers and Supervisors, Program Administrators and Directors	9/11-14/12	250	Both
Round 1: On-Site FGDM Coaching	3.5	FGDM Specialists	4/3/13-4/5/13	12	Dallas
Round 1: Session 1 Phone FGDM Coaching	-	FGDM Specialists	4/26/13	12	Dallas
Round 1: On-Site FGDM Coaching	3.5	FGDM Specialists	5/7-5/10-13	11	Tarrant
FGDM Supervisors Coaching & Consultation	1	FGDM Sups, leads	5/21/2013	6	Both
FGDM Supervisors Coaching & Consultation	1	FGDM Sups, leads	5/22/2013	1	Tarrant



FGDM Supervisors Coaching & Consultation	1	FGDM Sups, leads	5/23/2013	1	Dallas
Round 2: On-Site FGDM Coaching	3.5	FGDM Specialists	7/8-7/11/13	7	Dallas
Round 2: On-Site FGDM Coaching	3.5	FGDM Specialists	7/12, 15-16/13	11	Tarrant
Round 2: Session 1 Phone FGDM Coaching	-	FGDM Specialists	9/9/13	11	Tarrant

**SOUTH DAKOTA**

<b>TRAINING OR SERVICE (Coaching or TA)</b>	<b>Training Days</b>	<b>Audience (Caseworkers, supervisors; IFA, FGDM)</b>	<b>Dates</b>	<b>Number of Participants</b>	<b>Region (Rapid City/Sioux Falls)</b>
Introductory Overview of FGDM	3	IFA caseworkers and supervisors	3/27-3/29	34	Both
Preparation is Crucial: The Nuances to Coordinating Family Conferences	2	FGDM coordinators and supervisors	5/1-5/2	12	Both
Managing Emotions as an FGDM Coordinator/Facilitator	1	FGDM coordinators and supervisors	5/3	12	Both
On-site coaching	4	FGDM Supervisor and Coordinators	5/21-5/24	3	Rapid City
FGDM: At First Glance	2 .5-day sessions	Stakeholders and staff	7/24	50	Sioux Falls
The Referring Worker: A Key Role in the FGDM Process	2	IFA caseworkers	7/25-7/26	70	Sioux Falls
The Referring Worker: A Key Role in the FGDM Process	2	IFA caseworkers	8/14-8/15	70	Rapid City
NPLH Evaluation Training	2 .5-day sessions	IFA and Ongoing workers and supervisors; FGDM Coordinators and Supervisor	9/24-25	30	Rapid City

Using FGDM to Reduce Disproportionality in Child Welfare	2 1-day trainings	Caseworkers, supervisors, FGDM	4/30-5/1/13	59	Rapid City/Sioux Falls
Individual coaching sessions	2	FGDM coordinators and supervisor	7/1-7/2/13	4	Rapid City
<b>LARIMER COUNTY</b>					
<b>TRAINING OR SERVICE (Coaching or TA)</b>	<b>Training Days</b>	<b>Audience (Caseworkers, Supervisors, FGDM facilitators)</b>	<b>Dates</b>	<b>Number of Participants</b>	
FGDM: At First Glance	1	Agency staff and community stakeholders	4/25	22	
NPLH Evaluation Training	3 .5-day sessions	Facilitators, Intake, Family Assessment Response, and Ongoing caseworkers and supervisors	9/11-9/12	60	
Engaging Fathers in the Child Welfare Process	2 1-day trainings	Supervisors, FGDM Staff, Caseworkers	4/17-4/18/13	29	
The Impact of Power: Exploring Issues of Domestic Violence in FGDM	2	Supervisors, FGDM Staff, Caseworkers	6/20-6/21/13	20	
Involving Children in FGDM	2 1-day trainings	FGDM Staff	7/24-7/25/13	15	
Using FGDM to Reduce Disproportionality in Child Welfare	2 1-day trainings	Supervisors, FGDM Staff, Caseworkers	9/4-9/5/13	12	

## Appendix C. NPLH Instrument List with References

- 1. General Staff Survey<sup>5</sup>** (*1x to caseworkers/coordinators electronically*) (27 items)
    - includes tenure and duties, skills, job satisfaction, FGDM knowledge and attitudes, organizational culture and climate<sup>6</sup>, services, risk aversion, demographics.
    - Rationale*: The purpose of this survey is to assess the worker, coordinator and supervisor characteristics of staff involved with FGDM which can then be correlated with meeting referral rates, family satisfaction, fidelity, etc. Much of the content for this instrument was derived from the QIC-DR worker survey and adapted for an FGDM lens.
  
  - 2. Caregiver Survey** (*1x to caregivers via paper and pencil*) (53-55 items)\*
    - a. Caregiver demographics** (19 items)
      - including: gender, age, race/ethnicity, language spoken in home, education, employment status, public assistance, household income, economic hardship, marital status, housing status, adults in household, children in household.
      - Economic Hardship<sup>7</sup>** (2 items)
        - Rationale*: The purpose of these items is to assess the psychological sense of disparity between needs and resources, which can be viewed as the essence of stress (a key factor that affects child and family functioning). This measure may function as an important moderator of the intervention's effect on child and family outcomes (e.g., the extent to which FGDM services influence outcomes might vary as a function of caregivers' levels of economic hardship).
- 
- <sup>5</sup> Adapted from the QIC-DR General Worker Survey developed by Gary Siegel and Tony Loman at IARSTL.
- <sup>6</sup>Supervisor Competence scale. New York State Social Work Education Consortium. (2005). *Workforce Retention Survey*. Albany, NY: University of New York, Albany.  
McCarthy, M. Retrieved 11.15.11 from [http://www.ocfs.state.ny.us/ohrd/swec/pubs/swconsort05%20\\_Final.pdf](http://www.ocfs.state.ny.us/ohrd/swec/pubs/swconsort05%20_Final.pdf)
- Leadership scale. Potter, C. C., Comstock, A., Brittain, C., & Hanna, M. (2009). Intervening in multiple states: Findings from the western regional recruitment project. *Child Welfare*, 88(5), 169–185.
- Shared Vision/ Professional Orientation scale. Westbrook, T., Ellett, A., & Deweaver, K. (2009). Development and validation of a measure of organizational culture in public child welfare agencies. *Research on Social Work Practice*, 19(6), 730-741.
- Ellett, A. (2009). Intentions to remain employed in child welfare: The role of human caring, self-efficacy beliefs, and professional organizational culture. *Children and Youth Services Review*, 31(1), 78-88. doi: 10.1016/j.chilyouth.2008.07.002.
- <sup>7</sup>Conger, R. D., Ebert-Wallace, L., Sun, Y., Simons, R. L., McLoyd, V. C., & Brody, G. H. (2002). Economic pressure in African American families: A replication and extension of the Family Stress Model. *Developmental Psychology*, 38, 179-193.
- Conger, R. D., & Elder, G. H., Jr. (1994). *Families in troubled times: Adapting to change in rural America*. Hillsdale, NJ: Aldine.
- Conger, R. D., Ge, X., Elder, G. H., Jr., Lorenz, F. O., & Simons, R. L. (1994). Economic stress, coercive family process and developmental problems of adolescents [Special issue on children and poverty]. *Child Development*, 65, 541-561.

**b. Protective Factors Survey<sup>8</sup>** (20 items)

-Domains: family functioning and resiliency, social support, concrete support, child development/parenting knowledge, and nurturing and attachment.

-*Rationale*: Assessment of protective factors and the impact of FGDM on enhancing them was a focal point of the grant application. This instrument was a James Bell and Associates recommendation for the cross-grantee evolution and is being used in most other grantee evaluations.

**c. Child behavior**

\*skip pattern to direct respondents to age-appropriate child behavior subscale

-*Rationale*: It is hypothesized that the intervention will have an indirect effect on child functioning through parenting behaviors. Because FGDM services aim to change parent behavior and do not explicitly focus on changing child behaviors, we propose to use one measure of child adjustment, externalizing symptoms, at baseline and follow-up to assess change. While externalizing behaviors represent only one dimension of child adjustment, we hypothesize that caregivers will be more likely to see change for these behaviors more readily than less observable outcomes (e.g., internalizing symptoms, socio-emotional competence).

**C1. – “Infant Externalizing Scale”<sup>9</sup>** (7 items)

-Age under 11 months

-drafted for this survey by members of the NPLH evaluation team

OR

**C2. –Brief Infant Toddler Social Emotional Assessment (BITSEA)<sup>10</sup>** (7 items)

-Ages 11-36 months

-Externalizing subscale only

OR

**C3. Strengths and Difficulties Questionnaire (SDQ)<sup>11</sup>** (5 items)

-Ages 3 and up

---

<sup>8</sup> Survey developed by the FRIENDS National Resource Center for Community-Based Child Abuse Prevention in partnership with the University of Kansas Institute for Educational Research & Public Service.

<sup>9</sup> Covers constructs from both Bates (Infant Characteristic Questionnaire) and Rothbart (Infant Behavior Questionnaire), 4 out of 7 items correspond to Rothbart’s Negative Emotionality construct for which infant behavior predicts later negative affect in toddlerhood, and 2 of the 3 remaining items reflect Bates’ Fussy/Difficult/Demanding factor from his widely-used measure that we were going to use but for which the wording was too complicated.

<sup>10</sup> Briggs-Gowan, M. J., & Carter, A. S. (2008). Social-Emotional Screening Status in Early Childhood Predicts Elementary School Outcomes. *Pediatrics*, 121(5), 957-962.

Briggs-Gowan, M. J., & Carter, A. S. (2007). Applying the infant-toddler social & emotional assessment. *Infant Mental Health Journal*, 28(6), 564-583.

Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology*, 29(2), 143-155.

<sup>11</sup> YouthinMind. (2014a). Scoring the Strengths & Difficulties Questionnaire for age 2-4. Accessed Sept., 15, 2014 at <http://www.sdqinfo.com/py/sdqinfo/c0.py>; YouthinMind. (2014b). Scoring the Strengths & Difficulties Questionnaire for age 4-17. Accessed Sept., 15, 2014 at <http://www.sdqinfo.com/py/sdqinfo/c0.py>.

-Externalizing subscale only

d. **Child disability**<sup>12</sup> (2 items)

-*Rationale*: Children's disability is a risk factor for child welfare involvement and is significantly and positively associated in prior studies with a likelihood of getting a family meeting, and, thus, an important variable for the Propensity Score Matching.

**3a. FGDM Participant Fidelity**<sup>13</sup> (*to all meeting participants; some questions before the meeting and the majority at the conclusion via paper and pencil*) (41 items plus request for contact information)

-*Rationale*: The purpose of this instrument is to measure fidelity of the various meeting models being implemented in the three sites. It will be used to understand the consistency between family meeting practices and the FGDM guidelines. It will also be used to measure similarities and differences in family meeting participants' perceptions of what occurred before and during the meeting. Participants are asked to complete the first section of the survey before the family meeting begins. This section asks participants to answer questions about the preparation process and demographics. The second section of this survey is completed at the conclusion of the family meeting. A number of sources were used in the creation of this instrument.

**3b. Facilitator-Coordinator FGDM Fidelity Survey** (*post-meeting, electronically*) (32 items)

-*Rationale*: The Coordinator survey builds on the FGDM fidelity instrument for the other participants and the FGDM Guidelines, established by American Humane. Many of the questions are reworded to allow for analysis among the various respondents.

**4. Case-Specific Questionnaire**<sup>14</sup> (*at case closure to caseworkers, electronically*) (9 items)

-*Rationale*: The purpose of this instrument is to obtain case worker assessment of service needs and use by the caregiver. It will be used to understand outcomes in relationship to service needs and receipt. Service need data will supplement and validate or enhance administrative data on service needs. It will also be used to assess caseworker perception

---

<sup>12</sup> Adapted from Canadian Census:

([http://www.cdc.gov/nchs/ppt/citygroup/meeting9/citygroup9\\_Session3\\_2\\_Stobert.pdf](http://www.cdc.gov/nchs/ppt/citygroup/meeting9/citygroup9_Session3_2_Stobert.pdf)) and Children with Special Health Care Needs Screener: (<http://cahmi.org/ViewDocument.aspx?DocumentID=115>)

<sup>13</sup> Pennell, J. (2003). Achievement of objectives: pre and during conference. Raleigh, NC: North Carolina State University, Department of Social Work.

Rautkis, M. Family Group Survey. University of Pittsburgh, Pennsylvania Child Welfare Training Program.

Casey Family Programs Research Services, Lakota Oyate Wakanyeja Owicakiyapi, and Sicangu Child and Family Services. (2011). Participant Satisfaction Survey. See <http://www.casey.org/resources/publications/pdf/Evaluation-FGDM-Native.pdf>

Triwest Group. (2011). Family Team Decision Making Quality and Fidelity Index. Family Connections Demonstration Project.

University of Vermont (2011). After the Meeting Evaluation.

Darlington, Y., Healy, K., Yellowlees, J., & Bosly, F. (2012). Parents' perceptions of their participation in mandated family group meetings. *Children and Youth Services Review*, 34, 331-337.

<sup>14</sup> Adapted from QIC-DR Case-Specific Questionnaire developed by Gary Siegel and Tony Loman at IARSTL.

of whether service referral or provision for a caregiver stem directly from a family meeting or not.

**5. Follow-Up Survey** (*1x post-meeting to all families, via paper and pencil*)

*-Rational:* In order to assess family satisfaction with child welfare, we wanted a single tool that could be used with both our treatment and control groups. The items are based on the Colorado Family Exit Survey and the Participant Satisfaction Survey. Wording of many questions has been changed so that this satisfaction measure could be used for families whether or not they had a meeting. Additionally, items were also reworded to increase ease of use by allowing a set of uniform responses across most items. Various versions of the survey include additional components: follow-up fidelity (a re-asking of the post-meeting fidelity survey questions) and the caregiver post-test (a re-asking of Protective Factors, child behavior, mental health, substance abuse, and mental health). 3 versions are drafted based on target audience:

**5a. Intervention Caregiver Follow-Up** (*1x to caregivers in the intervention group via paper and pencil*) (78 items)

- Caregiver Post-test
- Fidelity follow-up
- Satisfaction

**5b. Control Caregiver Follow-Up** (*1x to caregivers in the control group via paper and pencil*) (63 items)

- Caregiver post-test
- Satisfaction

**5c. Meeting Participant Follow-Up** (*1x meeting participants who self-selected to be contacted post-meeting for follow-up information via paper and pencil*) (36 items)

- Fidelity follow-up
- Satisfaction

## Appendix D. General Staff Survey Cross-Site Results

It is believed that the successful implementation of an intervention in child welfare often depends on agency ownership, confidence in intervention effectiveness, and the ability to implement with fidelity. One such intervention, the Family Group Conference (FGC), emphasizes the rights of children and families to make decisions regarding placement and well-being. Because child welfare worker attitudes can impact referral rates to and participation in FGCs, researchers sought the following information in order to understand child welfare professionals' perspectives about the effectiveness of this intervention. These perceptions are important because, beyond referral and participation, they also reflect organizational support of FGCs and implementation fidelity.

The following data were collected from child welfare agency staff in three sites (four jurisdictions) participating in the No Place Like Home (NPLH) three-year evaluation of the use of family meetings in child welfare for in-home services. Staff completed an online survey composed of questions pertaining to case skills, FGC knowledge and attitudes, organizational culture and climate, service availability, and child safety vs. family preservation orientation. The sample of staff surveyed consisted of any staff who had a role in the NPLH evaluation, their supervisors, and FGC coordinators and trainers. In all, 301 staff members responded to the survey (58 percent from Dallas and Tarrant Counties in Texas; 33 percent from Larimer County, Colorado; and 9 percent from Rapid City, South Dakota).

### Staff Tenure and Duties

The vast majority of survey respondents (73 percent) were caseworkers, with supervisors and coordinators each representing an additional 12 percent of the sample. The remainder of the

sample responded “other” as their primary job responsibility (primarily consisting of program directors and trainers). Respondents spent a median of 5 years working in child welfare, though tenure varied based on job title and location (Median years in child welfare: Caseworkers – 3 years, supervisors – 12 years, coordinators – 12 years; Texas – 5 years, Colorado – 5 years, South Dakota – 3 years). Similarly, respondents spent a median of 3 years in their current position (Median years in current position: Caseworkers – 2 years, supervisors – 5 years, coordinators – 3 years; Texas – 3 years, Colorado – 2 years, South Dakota – 3 years). For those individuals who carried an active caseload, the average caseload size was 9.5 families; Colorado and South Dakota averaged slightly lower caseloads than workers in Texas. Caseworkers typically had “some” experience with family meetings, while supervisors and coordinators held “a lot” of experience; no differences were noted across sites.

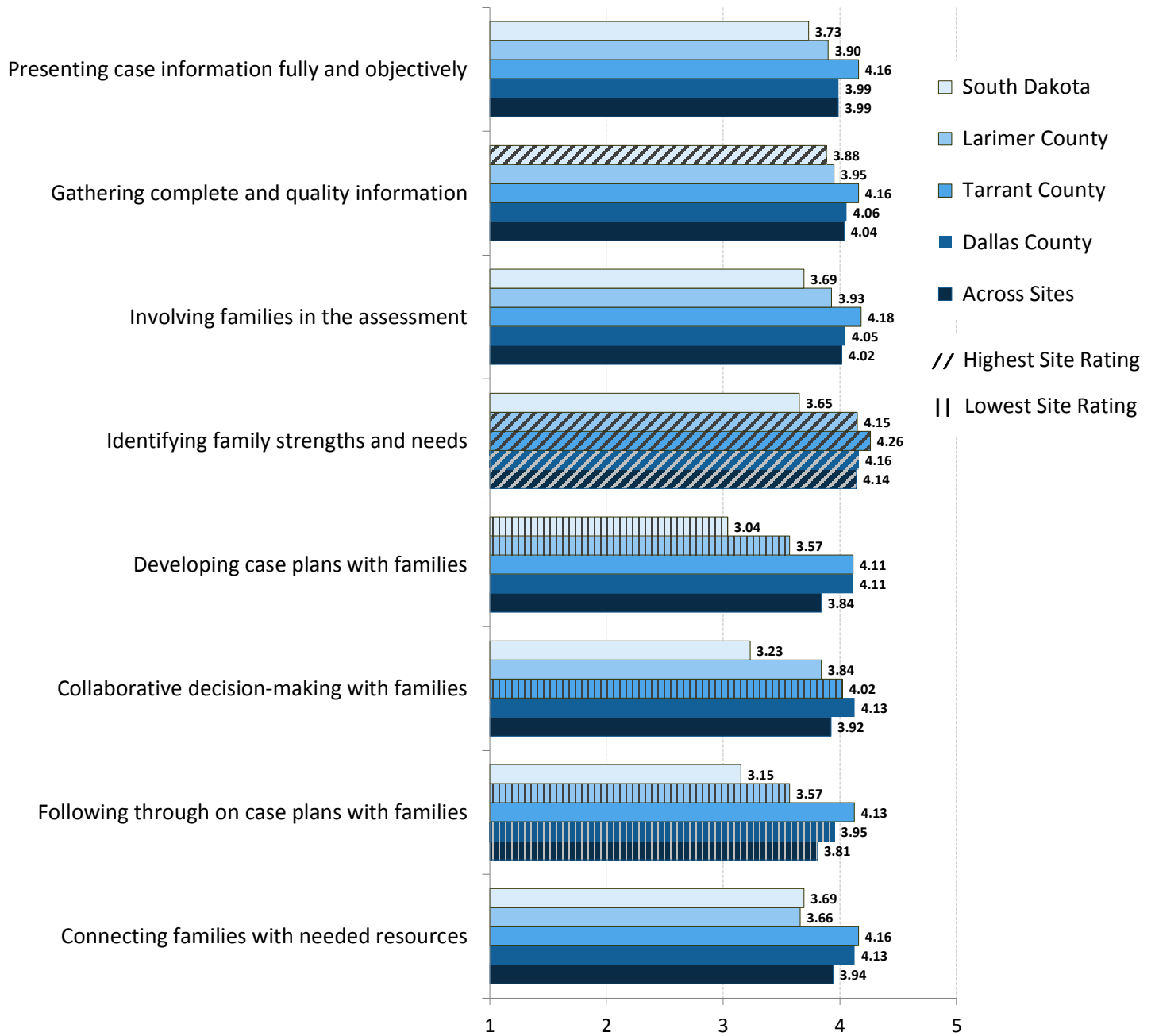
#### Case Skills

Each child welfare worker was asked to rate their skill level in different areas of casework practice. Figure 1 displays the average rating, separated by county and combined, for each area, with the highest- and lowest-rated areas highlighted. Most case skill areas received a rating of approximately 4 out of 5 (between a moderate and advanced level of skill). However, the skill areas receiving the highest and lowest ratings tended to differ somewhat by site. Larimer County (Colorado), Tarrant County (Texas), and Dallas County (Texas) rated *identifying family strengths and needs* as the highest skill area, whereas South Dakota (Rapid City) rated *gathering complete and quality information* the highest. Differences existed along the lowest-ranked skill area as well; Dallas and Larimer ranked *following through on case plans with families* the lowest, Tarrant ranked *collaborative decision-making with families* the lowest, and South Dakota and Larimer ranked *developing case plans with families* the lowest (Larimer County had a tie for



two lowest-ranked skill areas). Across sites, workers ranked identifying family strengths as a high-skill area, whereas follow through on was ranked lowest.

**Figure 1. Average Skill Level in Each Area by County**  
(1=Basic, 3=Moderate, 5=Advanced)



## Job Satisfaction

Child welfare staff were asked two questions regarding overall job satisfaction and whether or not the use of family meetings has made them more or less likely to remain working in the field of child welfare. Across all four sites, workers responded, on average, that they were *moderately to very satisfied* with their current job. By position, caseworkers averaged a rating of 3.29 (slightly above *moderately satisfied*), supervisors averaged 3.84 (slightly below *very satisfied*), and coordinators averaged 4.20 (between *very* and *completely satisfied*). Similarly, workers responded that they were between *moderately* and *very likely* to remain in child welfare because of the use of family meetings; however, workers at one site (Dallas) responded between only *slightly* and *moderately*. As with job satisfaction, caseworkers rated their likelihood of remaining in child welfare lower than supervisors, whose rating was lower than coordinators.

## Family Meeting and FGDM Knowledge and Attitudes

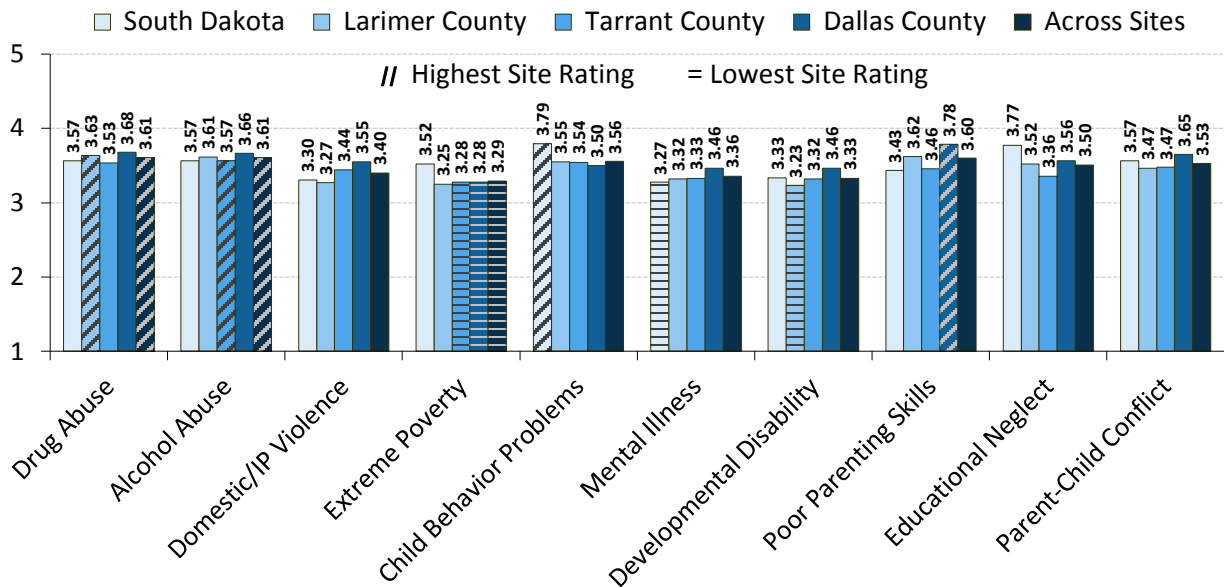
Figure 2 presents the level of agreement with 11 statements about children, families, and child welfare agencies, which are derived from the tenets of FGCs and family group decision making (FGDM). For all 11 statements, workers at each site ranked their level of agreement somewhere between *agree* and *strongly agree* with the exception of the following statement, *families know how to construct thorough plans for resolving their issues*, which received a rating between *slightly agree* and *agree*. This statement received the lowest level of agreement from all four sites. As before, the sites differed slightly regarding the statement with which they agreed the most; Larimer and Dallas ranked the following statement highest: *Children have a right to maintain their cultural identity throughout their lives*. By contrast, South Dakota and Tarrant ranked this statement the highest: *All families are entitled to be respected by CPS*. Caseworkers, supervisors, and coordinators tended to answer these questions similarly.

**Figure 2. Level of Agreement: FGDM Knowledge and Attitudes by County**  
 (1=Strongly Disagree to 6=Strongly Agree)

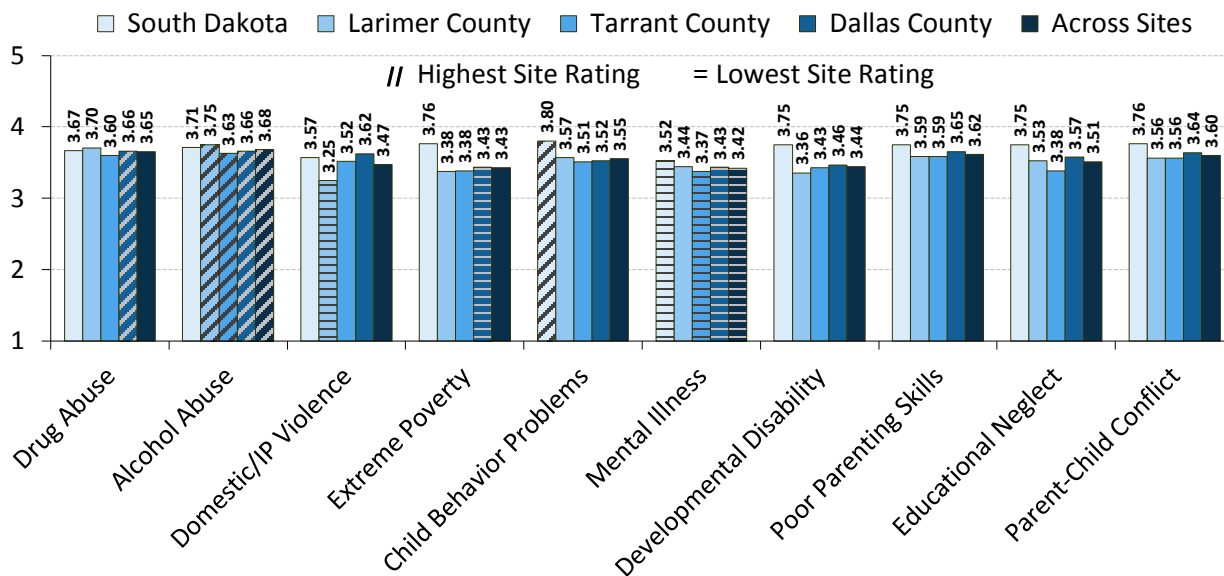


Figures 3 and 4 illustrate the answers provided by staff when asked to rank the effectiveness of family team meetings (including FSRTs and FUMs) or FGCs when working with families dealing with 10 different issues.<sup>15</sup>

**Figure 3. Effectiveness of Family Team Meetings in Working with Families with Issues of...** (1=Not at All Effective to 5=Completely Effective)



**Figure 4. Effectiveness of Family Group Conferences in Working with Families with Issues of...** (1=Not at All Effective to 5=Completely Effective)



<sup>15</sup> IP = Intimate Partner. Child Behavior Problems represent “extreme” child behavior problems, and Poor Parenting Skills represent “extremely” poor parenting skills.

For all 10 issues, across both types of meeting, workers ranked effectiveness between *moderately* and *very effective* (with *completely effective* being the highest possible ranking). Quite a bit of consistency existed across sites and across meeting types regarding the issues for which meetings were viewed as most effective. Both family team meetings and FGCs received the highest ranking of effectiveness, across sites, for dealing with *alcohol abuse* in families. (For family team meetings, *drug abuse* tied *alcohol abuse* for the highest ranking.) *Alcohol abuse* or *drug abuse* issues in families received the highest ranking of effectiveness for both meeting types in Larimer, Dallas, and Tarrant with only one exception: Dallas ranked family team meetings (FTM) the highest in effectiveness for working with families exhibiting *extremely poor parenting skills*. South Dakota differed from the three other sites in ranking *extreme child behavior problems* as the family issue with which FTMs and FGCs were most effective.

Slightly more variation in responses existed across meeting types and location regarding the issues for which meetings were least effective. While Larimer ranked the effectiveness of FTMs lowest for families dealing with *developmental disability*, South Dakota ranked FTMs lowest for dealing with *mental illness*, and both Tarrant and Dallas ranked FTMs lowest in dealing with *extreme poverty*. Regarding FGCs, three of the four sites (South Dakota, Tarrant, and Dallas) ranked *mental illness* as the issue for which meetings were least effective. Larimer, by contrast, ranked FGCs as least effective in dealing with *domestic/intimate partner violence*. (Note: For Dallas, *extreme poverty* tied *mental illness* in having the lowest ranking of FGC effectiveness.)

Overall, FTMs were ranked least effective in dealing with *extreme poverty*, whereas FGCs ranked lowest in dealing with *mental illness*. Rankings across caseworkers, supervisors, and coordinators tended to be similar, though supervisors and coordinators consistently ranked effectiveness higher.

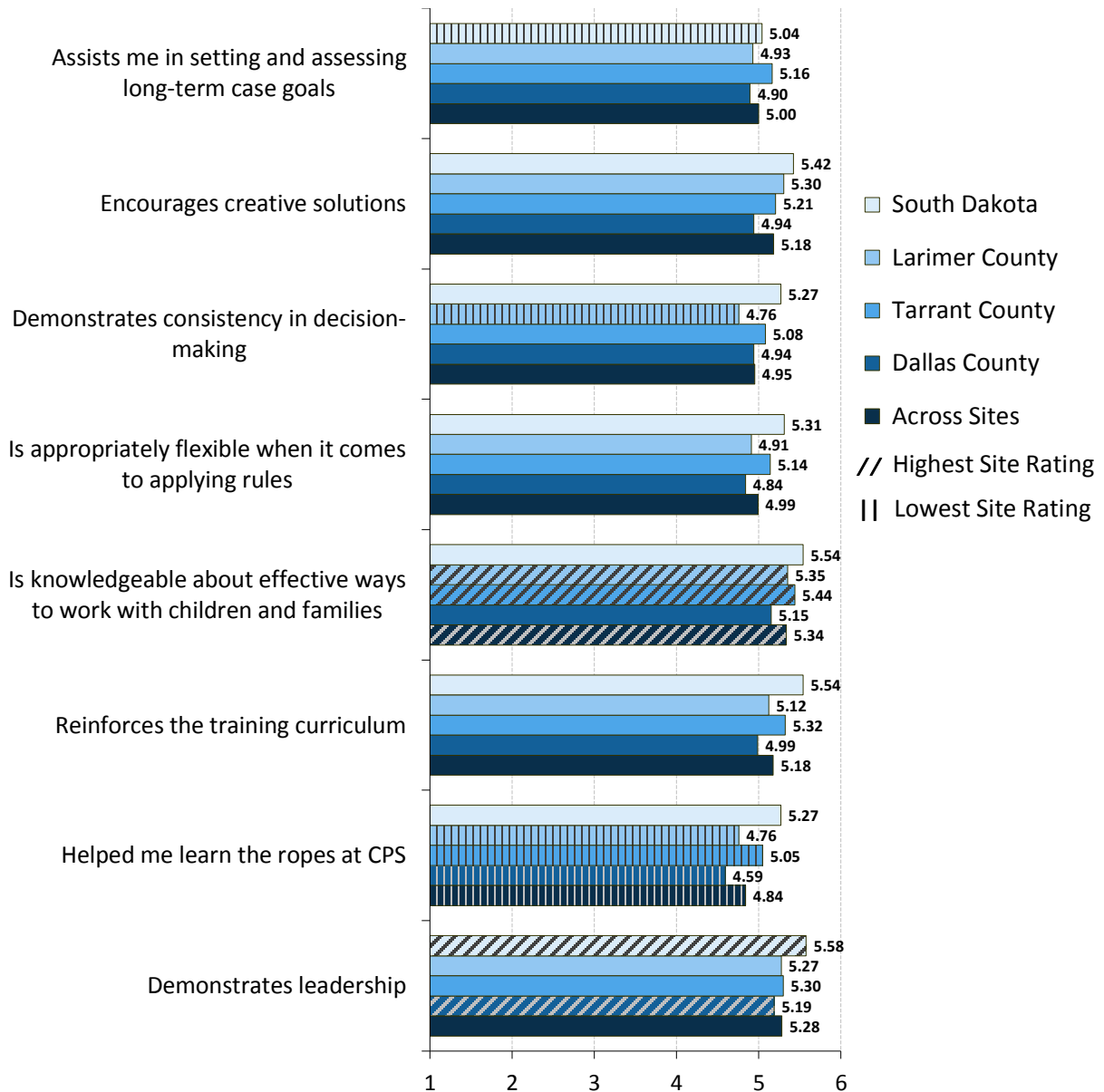
Beyond meeting effectiveness, child welfare staff were asked about family meeting utility and the accompanying increase (or decrease) in workload from said meetings. Dallas and Tarrant found family meetings to be between *moderately* and *very useful*, whereas Larimer and South Dakota both ranked the family meetings as *very useful*. Similarly, Dallas and Tarrant noted veritably no increase in day-to-day responsibilities as a result of the use of family meetings, whereas Larimer and South Dakota both noted a moderate increase in their respective workloads. Both supervisors and coordinators ranked the utility of family meetings much higher than caseworkers, though no difference was observed across job types in terms of increased workload.

### Organizational Climate and Culture

For the organizational climate and culture of the child welfare agencies, each staff member was asked to agree or disagree with a series of statements pertaining to their direct supervisor, agency leadership, and workers in their respective units. Regarding their supervisors (as shown in Figure 5), most sites (South Dakota being the exception) exhibited the least agreement with the statement, *my supervisor helped me learn the ropes at CPS*. Child welfare staff at South Dakota, conversely, exhibited the lowest level of agreement with: *my supervisor assists me in setting and assessing long-term case goals*. The greatest level of agreement was split across two statements: South Dakota and Dallas tended to agree strongly with, *my supervisor demonstrates leadership*, whereas Larimer and Tarrant tended to agree strongly with, *my supervisor is knowledgeable about effective ways to work with children and families*. While caseworkers, supervisors, and coordinators all ranked helping workers “learn the ropes” with a low level of agreement, some differences emerged across these job types in the highest level of agreement. Caseworkers tended to agree more strongly that their supervisors were knowledgeable, whereas

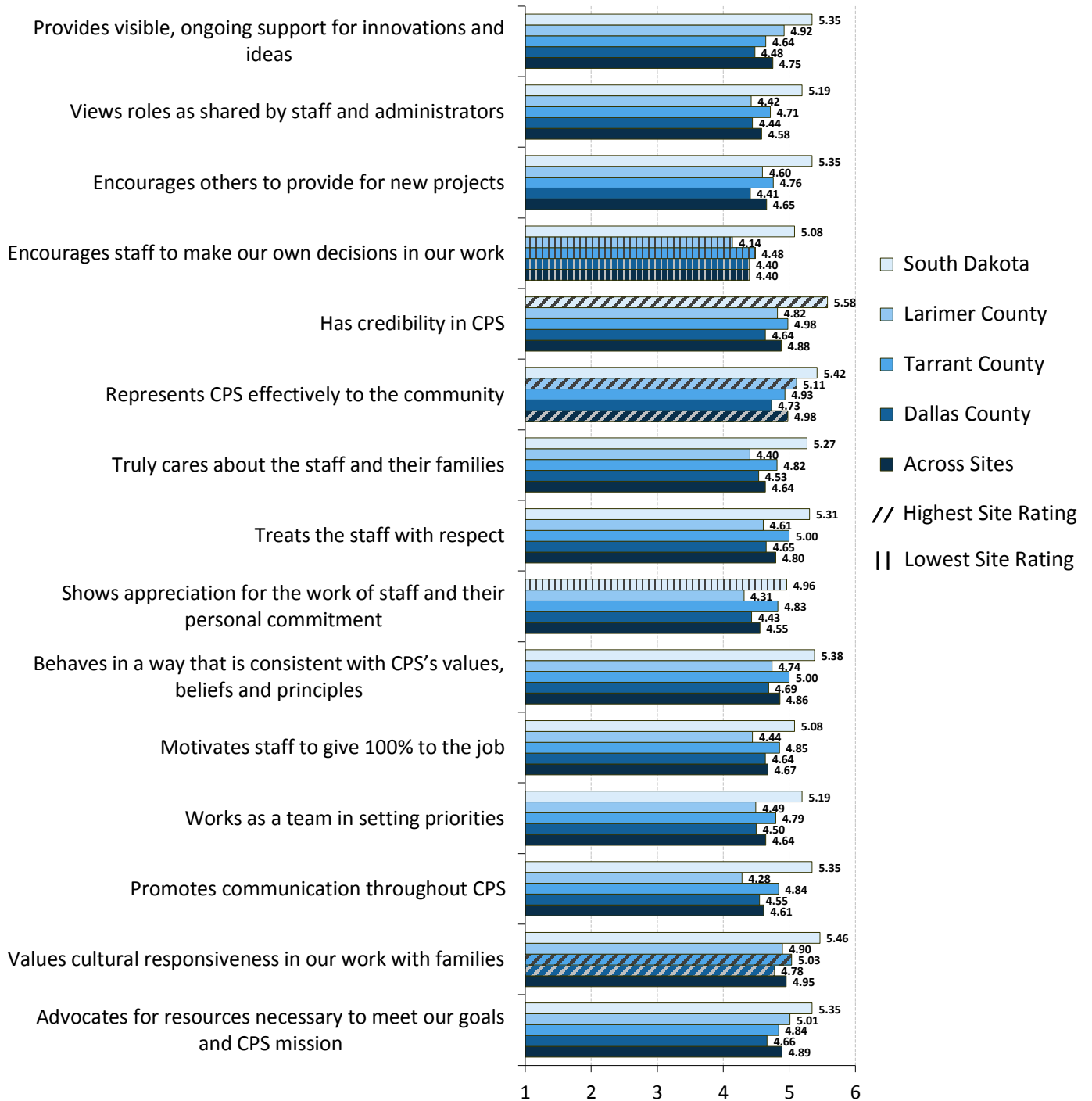
supervisors tended to agree with *my supervisor encourages creative solutions*, and coordinators tended to agree that their supervisors demonstrate leadership.

**Figure 5. Level of Agreement (by County): My Supervisor...**  
(1=Strongly Disagree to 6=Strongly Agree)



Child welfare staff were additionally asked about their level of agreement with 15 different statements regarding agency leadership (see Figure 6).

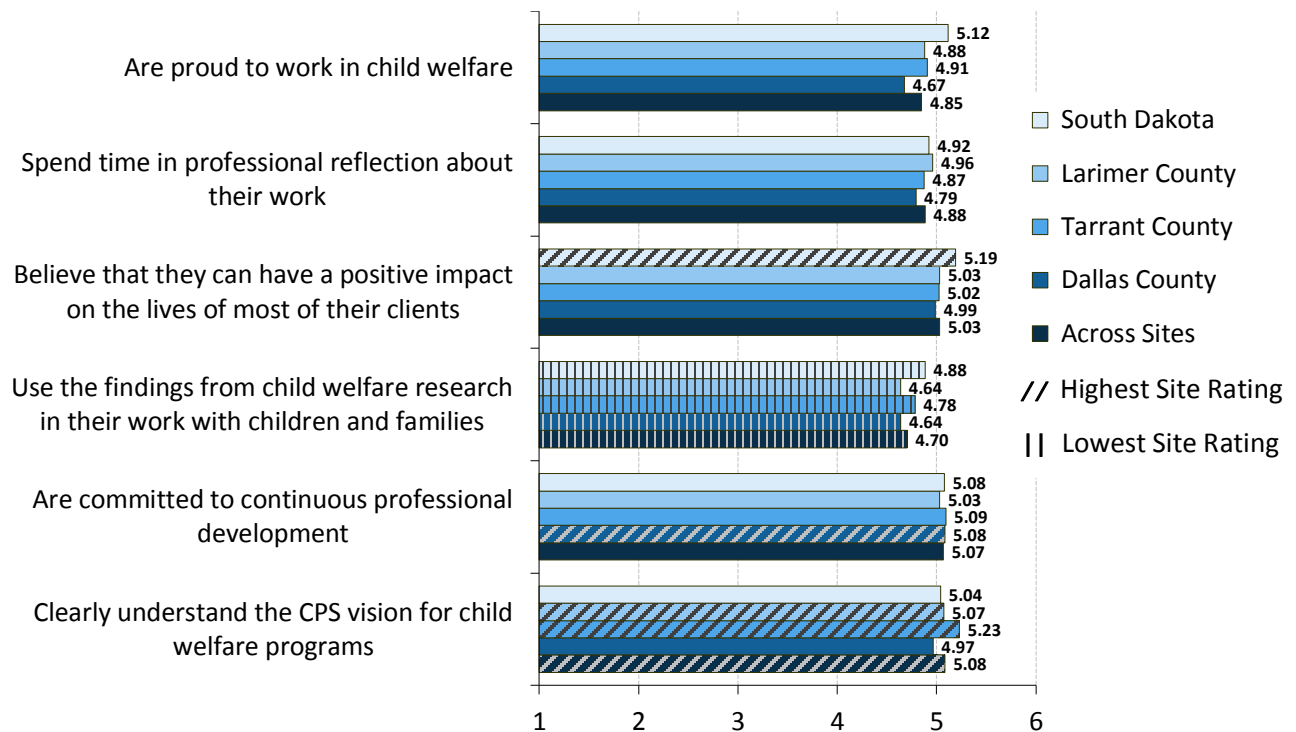
**Figure 6. Level of Agreement (by County): Leadership...**  
(1=Strongly Disagree to 6=Strongly Agree)





For three of the four sites, the level of agreement ranged between *slightly agree* and *agree* for almost all of the statements. (Responses from South Dakota tended to range between *agree* and *strongly agree*). Dallas and Tarrant agreed the strongest with the statement, *leadership values cultural responsiveness in our work with families*, while Larimer agreed the strongest with *leadership represents CPS effectively to the community* and South Dakota agreed the strongest with *leadership has credibility in CPS*. Dallas, Tarrant, and Larimer exhibited the lowest level of agreement with the same statement: *Leadership encourages staff to make our own decisions in our work*. South Dakota agreed the least with the statement that *leadership shows appreciation for the work of staff and their personal commitment*. No major differences were observed across job types.

**Figure 7. Level of Agreement (by County): Workers in My Unit...**  
(1=Strongly Disagree to 6=Strongly Agree)



Lastly, with regard to organizational climate and culture, staff were asked to agree or disagree with a series of statements about workers in their respective units. Figure 7 displays the

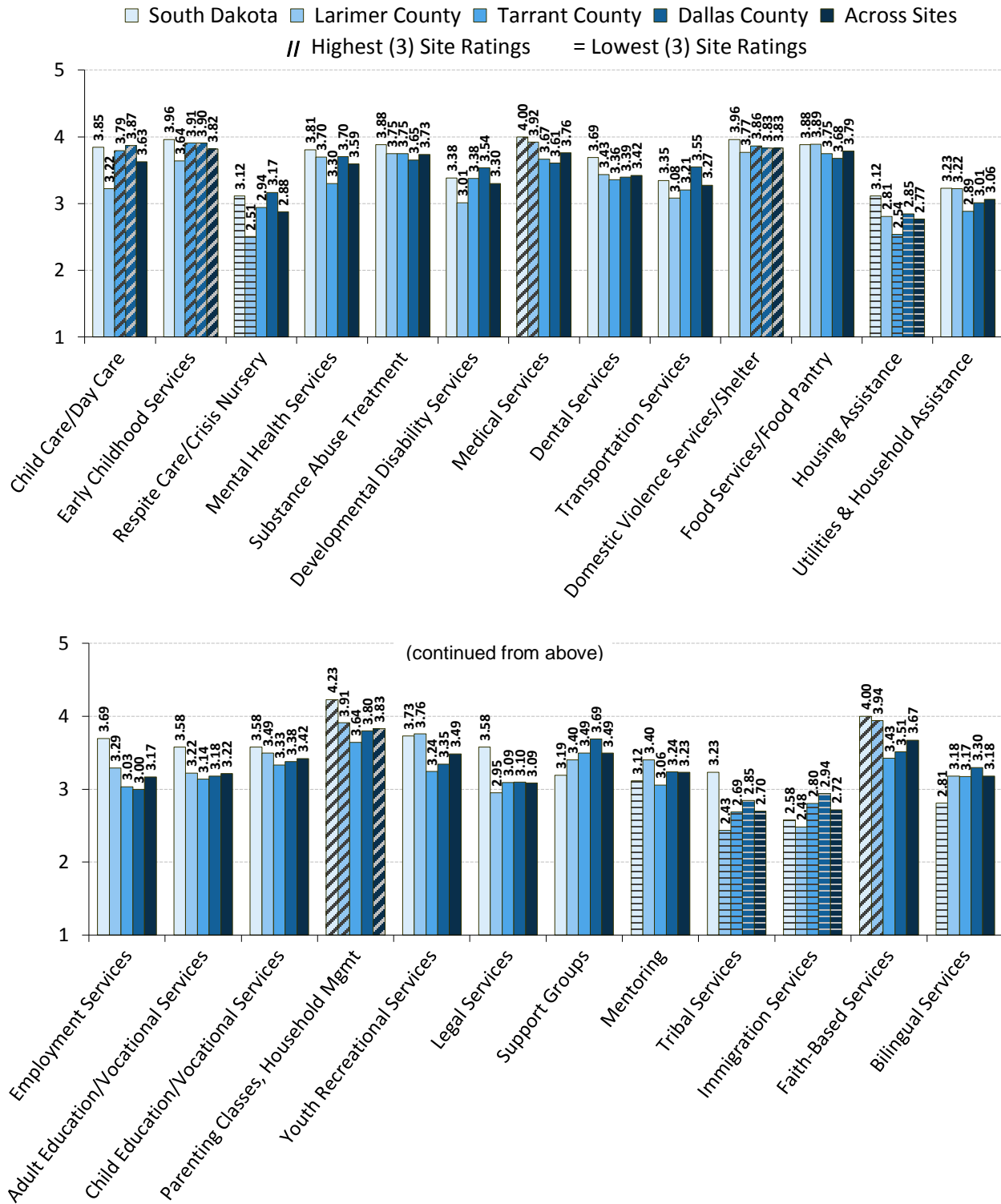
results for this series of questions. All sites selected *agree* on average with the six questions, though one question, on average, stood out as having the lowest level of agreement across the sites: *Workers in my unit use the findings from child welfare research in their work with children and families*. Conversely, some difference existed across sites regarding the question for which sites demonstrated the highest level of agreement. Larimer and Tarrant agreed most strongly with *workers in my unit clearly understand the CPS vision for child welfare programs*, whereas South Dakota agreed the strongest with *workers in my unit believe they can have a positive impact on the lives of most of their clients*, and Dallas agreed the strongest with *workers in my unit are committed to continuous professional development*. A high level of consistency existed across job types.

### Services

Two questions were asked concerning workers' ability to find services in the community and how easy it is to work with service providers. Across sites and job types, staff responded with a level of agreement between *slightly agree* and *agree* that they are able to locate services in the community and the service providers are easy to work with.

Figure 8 displays the confidence level of child welfare staff that local providers are able to meet any of 25 different potential family needs. (As opposed to prior figures, Figure 8 displays the highest *three* rankings and the lowest *three* rankings for each site.) Despite being located in three different regions across four counties, staff exhibited a surprising amount of agreement around their confidence in the ability of local providers to meet needs. *Housing assistance*, *tribal services*, and *immigration services* were commonly ranked lowest in terms of staff confidence. *Respite care/crisis nursery* ranked among the lowest in South Dakota and Larimer County. South Dakota also ranked *mentoring* and *bilingual services* among the lowest in staff confidence levels.

**Figure 8. Ability of Local Community Providers (Contract or Non-Contract) to Meet the Following Needs by County**  
 (1=Not at All Confident to 6=Completely Confident)



Regarding the highest levels of confidence, staff in Dallas and Tarrant Counties ranked their highest level of confidence in the ability of local providers to meet family needs concerning *child care/day care, early childhood services, and domestic violence services/shelter*. South Dakota and Larimer County ranked three different service categories among their highest level of confidence in the ability of local providers to meet family needs: *medical services, parenting classes/household management, and faith-based services*. In relation to services provided in the community, staff were asked one final question about how responsive said services are to culturally diverse groups. Across sites and job types, workers responded that community-provided services are *moderately to very responsive* to different cultural groups with one exception: supervisors ranked services as only *slightly to moderately responsive*.

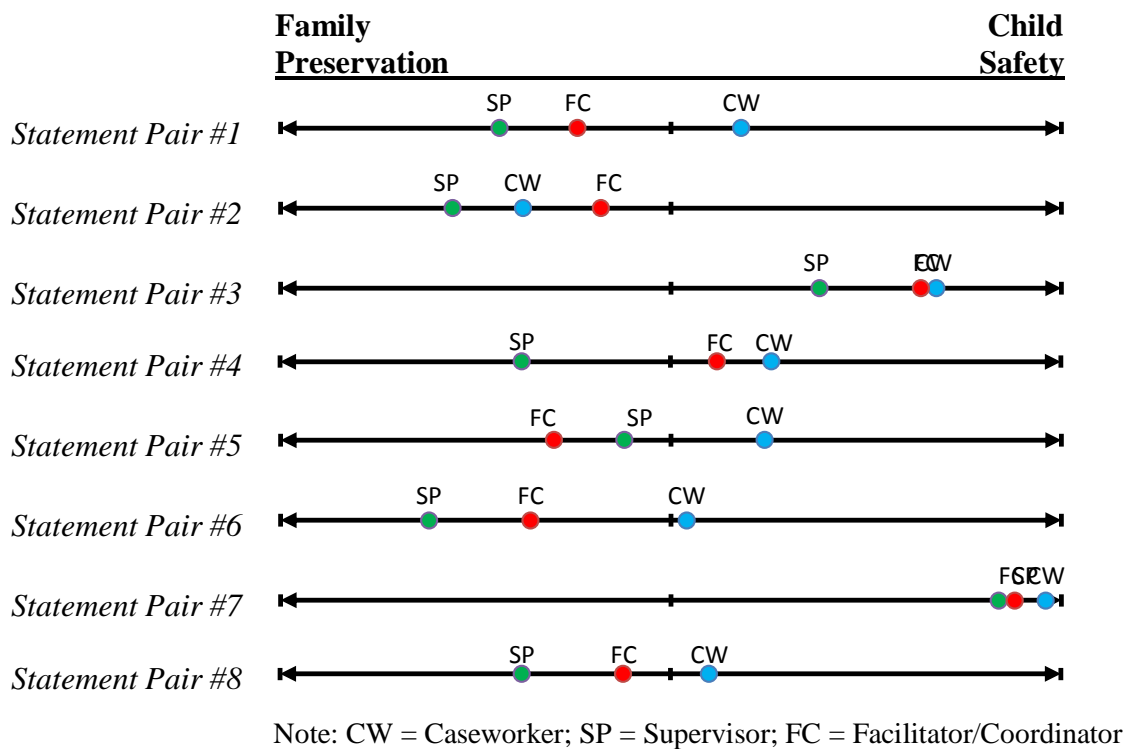
#### Child Safety or Family Preservation Continuum

Each child welfare staff member was asked to read a set of paired statements and choose the statement that best reflected his or her general work focus and beliefs. In each pair, one statement had an orientation toward child safety (e.g., The client is the child and all other work is secondary.) and the paired statement had an orientation toward family preservation/reunification (e.g., Work should be focused on keeping the family together.). Staff members were not told how each statement was oriented, and the statements were paired randomly. **Figure 9** displays the strength of orientation toward child safety versus family preservation by job type.

Looking at Statement Pair #1, for example, the responses from supervisors and facilitators are, on average, oriented toward family preservation, whereas the average response from caseworkers is oriented toward child safety (for the same pair of statements). For several of the paired statements given to each staff member (i.e., pairs #2, #3, and #7), the orientation was consistent across positions within the agency. Wherever the orientations diverged based on

position in child welfare, however, the orientation diverged in the same way: Caseworkers selected the statement with a child safety orientation, whereas supervisors and/or facilitators selected the statement with a family preservation orientation (see, for example, pair #1 or #5). While Figure 9 displays the cross-site results, similar patterns were observed for each site.

**Figure 9. Cross-Site Child Safety versus Family Preservation Orientation by Position (Job) within Child Welfare (N = 285)**



The information contained in this appendix was designed to help agency leadership and staff members better understand agency strengths and possible areas for improvement across a number of domains. It can be used to inform, for example, issues for which FTMs and FGCs are most effective. These results should not be the only consideration for decision-making within a child welfare agency, and are not representative of child welfare staff as a whole.

## Appendix E. Larimer County Propensity Score Matching (PSM)

Propensity score matching (PSM) was used to examine FGDM outcomes in Larimer County. The initial sample consisted of 870 individuals (representing 503 families). A total of 12 cases were excluded for having missing information on baseline covariates of interest, leaving a sample of 858 individuals (498 families) available for analysis.

The ‘treatment’ group consisted of 466 individuals (54.3% of the sample) in families who received at least one FGC or FUM meeting during a six- to nine-month period after the case was opened to in-home services. The comparison group consisted of individuals in families who did not receive an FGC or FUM in the six- to nine-month period after the case was opened. Case and comparison families were matched on the basis of propensity scores, which collapse a set of background covariates into a single summary measure (the propensity score), representing an estimate of the probability of receiving treatment. PSM involves several analysis steps, namely, (1) selecting background covariates to be included in the propensity score model; (2) estimating propensity scores and using the scores for matching; and (3) evaluating the propensity score model by examining balance diagnostics. These steps are described below.

### *Selection of variables to be included in PSM models*

Propensity scores collapse a set of observed background covariates into a single summary measure (the propensity score), representing an estimate of the probability of receiving treatment. Four individual (child)-level covariates and six family-level covariates were selected. The individual-level covariates were **gender, age, abuse allegations, neglect allegations,** and **initial placement**. The family-level covariates were **caregiver history of child maltreatment, domestic violence in the household, caregiver substance abuse, prior involvement with CPS, number of children in the case,** and **prior FAR involvement**. These variables were

selected on the basis that they were measured at baseline (time of referral) and could potentially affect outcomes (potential confounders) or both treatment assignment and outcomes (true confounders). As described in Appendix F. Larimer Child Characteristics Prior to PSM, comparisons of the intervention and comparison groups prior to matching revealed several statistically significant differences between the groups prior to matching. Specifically, rates of previous CPS involvement, caregiver history of childhood abuse or neglect, substance abuse, and initial placement were significantly higher in the intervention group than in the comparison group. Additionally, the intervention group was more likely to include boys and larger families (more children per family) compared with the comparison group. In contrast, rates of prior involvement in FAR were significantly higher in the comparison group than in the intervention group. Thus, **gender, initial placement, caregiver history of child maltreatment, substance abuse, prior involvement with CPS, number of children in the case, and prior FAR involvement** were specifically included in the PSM as potential true confounders.

#### *Propensity score modeling and matching*

Propensity scores were estimated using a fixed effect logit model, with a dummy variable representing family to account for the fact that cases were “clustered” within families (fixed effect model). This model was used to estimate the probability of receiving an FGC meeting (being in the “treatment” group), based on the 10 covariates listed above, and was estimated in the *R* statistical environment using the *lme4* *R* package.

Following the estimation of the propensity scores, the *MatchIt* *R* package in the *R* statistical environment was used to form matched sets of treatment and comparison cases. Specifically, a genetic matching algorithm was used to minimize a measure of the maximum observed discrepancy between the matched treatment and comparison covariates at every iteration of

optimization; that is, systematic differences in baseline covariates between treatment and comparison cases. Additionally, *matching with replacement* was allowed; that is, if one comparison case matched more than one treatment case, the matched dataset included the multiple matched control observations to ensure optimal matching. The matched data were weighted to reflect the multiple matches. Based on these specifications, 466 children who received an FGC or FUM meeting were matched to 287 “comparisons” who did not receive an FGC or FUM meeting. A total of 105 comparison children could not be matched to an individual in the treatment group, and were not used in further analyses. A summary of the characteristics of the matched treatment and comparison cases is provided in Table 14 of this report.

### *Balance diagnostics*

To evaluate the extent to which covariate imbalances were minimized in the PSM analyses, treatment and comparison cases were compared on the means of covariates, variance ratios, standardized mean differences, and the higher-order moments and interactions between covariates. Overall, these indicators suggested that treatment and comparison cases were more similar after matching than before matching. A selection of the results, specifically the means of the treatment and comparison cases and the standardized mean differences (SMD), are shown in Table 1, below.

**Table 1. Balance Statistics Before and After Matching**

	Means Treated	Means Comparison	Standardized Mean Difference
<b>Balance statistics, before matching</b>			
Gender	.55	.48	.14
Abuse allegations	.23	.24	-.02
Neglect allegations	.79	.76	.08
Risk of substance abuse	.50	.41	.19
Risk of domestic violence	.56	.55	.02
Caregiver history child maltreatment	.34	.20	.29
Prior CPS involvement	.78	.71	.16

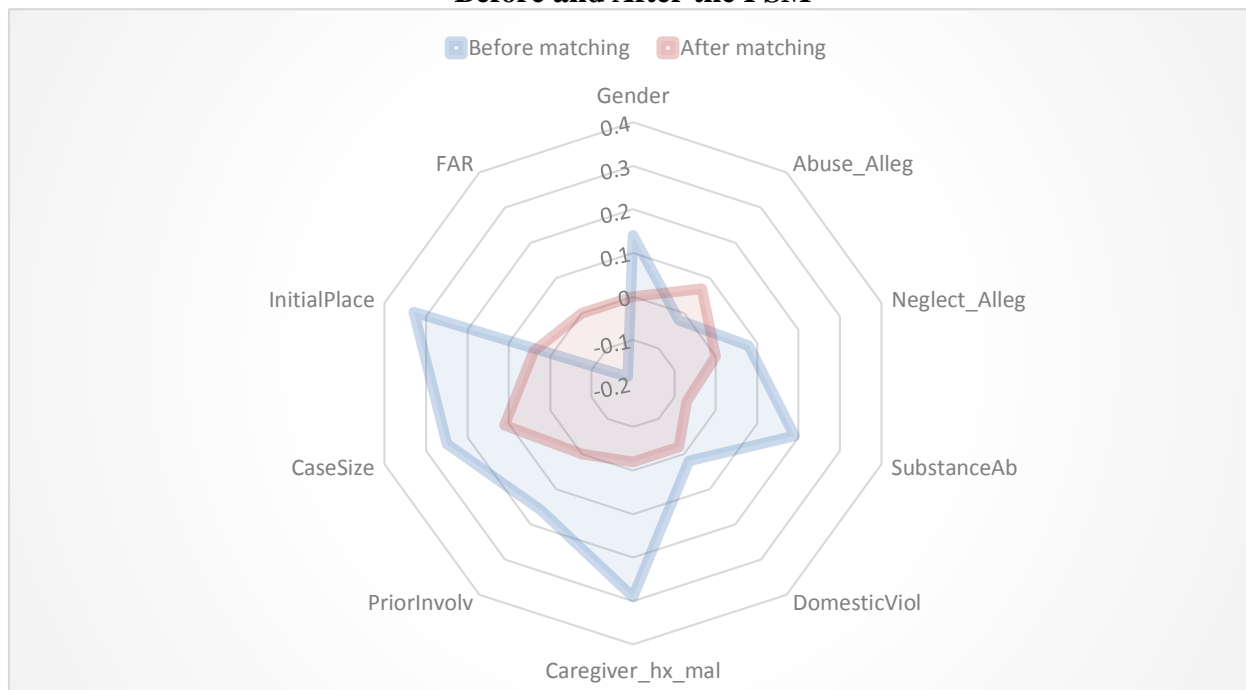


Number of children in case	2.45	2.10	.25
Initial placement	.23	.09	.33
FAR	.55	.64	-.18
<b>Balance statistics, after matching</b>			
Gender	.55	.55	.00
Abuse allegations	.23	.20	.07
Neglect allegations	.79	.79	.00
Risk of substance abuse	.50	.54	-.07
Risk of domestic violence	.56	.57	-.02
Caregiver history child maltreatment	.34	.35	-.02
Prior CPS involvement	.78	.78	.00
Number of children in case	2.45	2.29	.11
Initial placement	.23	.21	.04
FAR	.55	.55	.00

Note: Gender is coded 0 = Female, 1 = Male; Abuse allegations, neglect allegations, risk of substance abuse, risk of domestic violence, caregiver history of child maltreatment, Prior CPS involvement, Initial placement, and Prior FAR involvement are coded 0 = No, 1 = Yes. Number of children in case is a continuous variable that ranged from 1 to 7.

After matching, the mean scores on all covariates were identical or very similar for treatment and comparison cases. The SMD is a measure of effect size defined as the ratio of mean to standard deviation of the difference of two random values respectively from the intervention and comparison groups. Conventionally, an SMD less than 0.1 indicates a negligible difference in the mean or prevalence of a covariate between groups. The improvement in SMD scores can be seen in Figure 1. The blue area indicates SMD scores on each variable prior to matching; in most cases, SMD scores were greater than 0 (though less than 0 for FAR, indicating that the comparison children were more likely to have previous FAR involvement than intervention children). The red area indicates SMD scores on each variable after matching; for almost all variables, SMD scores were 0 or close to 0. Thus, there was considerable shrinkage in SMD scores after matching.

**Figure 1. Standardized Mean Differences between Treatment and Comparison Cases Before and After the PSM**



**Note: Abuse\_Alleg = Abuse allegations; Neglect\_Alleg = Neglect allegations; SubstanceAb = Substance abuse; DomesticViol = Domestic violence; Caregiver\_hx\_mal = Caregiver history of childhood maltreatment; PriorInvolv = Prior history of CPS involvement; CaseSize = Number of children in case; InitialPlace = Child received prior placement; FAR = Prior involvement in FAR.**

### *Summary*

As a result of the PSM analyses, a matched sample of 466 individuals was identified who received a FUM or FGC meeting and 287 individuals who did not receive a FUM or FGC meeting. This matched sample was used in the outcome analyses for Larimer County, reported in Section VII.A of the report.

## Appendix F. Larimer County Child Characteristics Prior to PSM

Characteristics of children involved in the Larimer County Study were drawn from SACWIS administrative data. Sample characteristics of the PSM-matched treatment ( $n = 466$  individuals) and comparison ( $n = 287$  individuals) groups are reported in Section VII.A of this report. Prior to matching, the sample consisted of 475 individuals in the treatment group and 395 individuals in the comparison group. Sample characteristics for these groups, as well as for the combined sample ( $n = 870$  individuals), are reported in this Appendix. Comparisons between the treatment and comparison groups were made using chi-square tests for continuous and ordinal variables, and independent  $t$ -tests for continuous variables.

### *Demographic characteristics*

The sample overall included slightly more males (52.3%) compared with females (47.7%). Children in the treatment group were significantly more likely to be male (55.6%) than female (44.4%) compared with the comparison group (48.4% male, 51.6% female).

The mean age of children in the sample overall was 6.45 years (SD: 4.81). The difference between the mean ages of children in the treatment and comparison groups was not statistically significant, although children in the comparison group were, on average, slightly younger (mean: 6.23 years; SD: 4.93) than children in the treatment group (mean: 6.73 years; SD: 4.66). The majority of the sample was between 2 to 5 years of age (30.8%).

In both the treatment and comparison group, just over one-quarter of children were Hispanic (27.6% in the sample overall). Within the treatment group, other minority racial/ethnic groups included African American (.8%), and non-Hispanic mixed race (3.4%). Within the comparison group, other minority racial/ethnic groups included African American (.3%), Asian (1.3%) and non-Hispanic mixed race (3%).

The treatment group of 475 children were part of 265 families, whereas the comparison group of consisted of 395 children in 238 families. The average number of children in treatment families (mean: 2.47; SD: 1.39) was significantly greater than the average number of children in comparison families (mean: 2.10; SD: 1.05).

#### *Maltreatment allegations and risk factors*

In both the intervention and the comparison group, neglect allegations were almost 4 times more common than abuse allegations. Compared with the comparison group, the treatment group had significantly higher rates of caregiver history of child maltreatment (34.1% in the treatment group vs. 20.4% in the comparison group) and substance abuse (50.4% in the treatment group vs. 41.1% in the comparison group). The rate of domestic violence was also higher in the treatment group (56.4%) than in the comparison group (55.4%); however, this difference was not statistically significant.

#### *Prior CPS involvement*

Prior CPS involvement was assessed in several ways. First, previous reports of suspected maltreatment were examined. Around three-quarters of the sample overall had previous reports of suspected maltreatment. However, the treatment group was significantly more likely to have a prior report of maltreatment, particularly at least one prior report of maltreatment that had been accepted for further investigation, compared with the comparison group. Specifically, children in the comparison group were more likely to have no prior reports than the treatment group (29.1% vs. 21.7% for the comparison and treatment groups, respectively), whereas children in the treatment group were more likely to have at least one prior report that had been accepted for further investigation compared to the comparison group (69.5% vs. 60.5% for the treatment and

comparison groups, respectively). Around 10% of children in both groups had no previous accepted reports of maltreatment, but at least one prior report that had been screened out.

Second, we considered whether children had previously been involved in a Family Assessment Response (FAR). A significantly greater proportion of children in the comparison group had previously been involved in FAR (63.5%) compared with the treatment group (54.3%).

Third, we considered whether children had previously been removed from the home within 90 days of a report, whether they were in placement at the time of referral, and what type of placement they were in at the time of referral. Just under 5% of children had previously been removed from the home (4.6% in both the treatment and comparison groups). Children in the intervention group were significantly more likely to be in placement at the time of intake; specifically, rates of placement in the treatment group (23.2%) were more than double the rates of placement in the comparison group (8.9%). The most common type of placement within both groups was kinship care (17.7% of children in the treatment group; 6.3% of children in the comparison group), followed by foster homes (5.3% of children in the treatment group; 2.8% of children in the comparison group). One child in the comparison group (.3%) was in a group home at the time of referral, and five children (2 in the treatment group, .4%, and 3 in the comparison group, .8%), were in hospital.

**Table 1: Child Characteristics, Larimer County, Before PSM**

Categorical/ordinal variables		Treatment (n = 466)		Comparison (n = 287)		$\chi^2$ (df)	p	Total (n = 753)	
		n	%	n	%			n	%
<i>Demographic characteristics</i>									
Gender	Male	264	55.6	191	48.4	4.51 (1)	.04	455	52.3
	Female	211	44.4	204	51.6			415	47.7
Age	0-1 years	110	23.2	97	24.6	6.39 (3)	.09	207	23.8
	2-5 years	132	27.8	136	34.4			268	30.8
	6-10 years	128	26.9	87	22.0			215	24.7

	11+ years	105	22.1	75	19.0			180	20.7
Ethnicity	African American	4	.8	1	.3	8.09(5)	.15	5	.6
	Asian	0	0	5	1.3			5	.6
	Caucasian	327	68.8	261	66.1			588	67.6
	Hispanic	126	26.5	114	28.9			240	27.6
	2 or more non-Hispanic	16	3.4	12	3			28	3.2
	Unknown	2	.4	2	.5			4	.5
<i>Maltreatment allegations and risk factors</i>									
Abuse allegations	Yes	110	23.2	93	23.5	.02 (1)	.89	203	23.3
	No	365	76.8	302	76.5			667	76.7
Neglect allegations	Yes	375	78.9	298	75.4	1.51 (1)	.22	673	77.4
	No	100	21.1	97	24.6			197	22.6
Caregiver childhood history A/N	Yes	159	34.1	80	20.4	19.91 (1)	< .01	239	27.9
	No	301	65.9	312	79.6			619	72.1
Risk of substance abuse	Yes	235	50.4	161	41.1	7.50 (1)	.01	396	46.2
	No	231	49.6	231	58.9			462	53.8
Risk of domestic violence	Yes	263	56.4	217	55.4	.10 (1)	.75	480	55.9
	No	203	43.6	175	44.6			378	44.1
<i>Prior CPS involvement</i>									
Priors - ordinal	No priors	103	21.7	115	29.1	7.94 (2)	.01	218	25.1
	No accepted priors, but at least 1 prior that was screened out	42	8.8	41	10.4			83	9.5
	At least 1 prior that was accepted	330	69.5	239	60.5			569	65.4
Priors - categorical	Yes	372	78.3	280	70.9	6.34 (1)	.01	652	74.9
	No	103	21.7	115	29.1			218	25.1
FAR	Yes	258	54.3	251	63.5	7.57 (1)	< .01	509	58.5
	No	217	45.7	144	36.5			361	41.5
Prior removal	Yes	22	4.6	18	4.6	.00 (3)	.96	40	4.6
	No	453	95.4	377	95.4			830	95.4
Initial placement	Yes	110	23.2	35	8.9	31.7 (1)	< .01	145	16.7
	No	365	76.8	360	91.1			725	83.3
Initial placement type	None	364	76.6	355	89.9	31.6 (4)	< .01	719	82.6
	Foster home	25	5.3	11	2.8			36	4.1
	Group home	0	0	1	.3			1	.1
	Hospital	2	.4	3	.8			5	.6
	Kinship care	84	17.7	25	6.3			109	12.5
		<b>Treatment</b>		<b>Control</b>				<b>Total</b>	
<b>Continuous variables</b>		<b>n</b>	<b>Mean (SD)</b>	<b>n</b>	<b>Mean (SD)</b>	<b>t (df)</b>	<b>p</b>	<b>n</b>	<b>Mean (SD)</b>

*Demographic characteristics*

Age	475	6.23 (4.93)	395	6.73 (4.66)	1.54 (868)	.12	870	6.45 (4.81)
Number of children	475	2.47 (1.39)	395	2.10 (1.05)	-4.53 (860)	< .01	870	2.30 (1.26)

Note: Caregiver childhood history A/N: Caregiver history of childhood abuse or neglect. Prior - ordinal: Prior involvement with CPS (0 = no priors; 1 = no accepted priors but  $\geq$  prior that was screened out; 2  $\geq$  1 prior that was accepted), Priors - categorical: Prior involvement with CPS (0 = no prior involvement; 1 = prior involvement); FAR: Prior involvement in Family Assessment Response (FAR); Prior removal: Child removed within 90 days of a previous report of child maltreatment. Initial placement: Child had placement at time of referral.

In summary, there were several statistically significant differences between the treatment and comparison groups. Rates of previous CPS involvement, caregiver history of childhood abuse or neglect, substance abuse, and initial placement were significantly higher in the intervention group than in the comparison group. Additionally, the treatment group was more likely to include boys and larger families (more children per family) compared with the comparison group. In contrast, rates of prior involvement in FAR were significantly higher in the comparison group than in the intervention group. Due to the possibility that these factors may have influenced selection of children into the comparison or intervention group, they were included in the PSM process, as described in Appendix F. Larimer County Propensity Score Matching.

## Appendix G. The Implementation of Family Group Decision Making in South Dakota during the No Place Like Home project

### **Introduction**

No Place Like Home (NPLH) was a project funded by a 2011 Family Connections grant, one of seven three-year grants administered that year by the U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau, to implement and evaluate Family Group Decision Making (FGDM) in child welfare for families receiving in-home services to prevent placement of children in foster care. The project was a collaboration between The Kempe Center for the Prevention and Treatment of Child Abuse and Neglect, Casey Family Programs, and three public child welfare agencies—Rapid City, South Dakota; Larimer County, Colorado; and Dallas and Tarrant Counties in Texas. The project started in October 2011, providing training and preparing for the evaluation in year 1 and implementing the evaluation in years 2 and 3. Following a fourth year funded by a no-cost extension to complete final analyses, the final full evaluation report will be released in December 2015.

The primary goal of the No Place Like Home (NPLH) project was to assess the effectiveness of FGDM in preventing out-of-home placement for child welfare-involved children and youth through implementation, program enhancement, and rigorous evaluation of FGDM. An additional goal was to assess impact of FGDM on racial disproportionality with regards to placement with a site-specific focus on African Americans in Texas, Latinos in Colorado, and American Indian children in South Dakota. Towards achieving these ends, the Kempe Center and Casey Family Programs, in partnership with the sites, worked to:

1. Analyze implementation of multiple FGDM models to provide knowledge on fit, flow and sustainability



2. Conduct an outcome evaluation to test the effectiveness of FGDM in preventing children from entering or re-entering foster care
3. Determine the costs for FGDM implementation
4. Disseminate implementation, cost and outcome findings and products

## **Training**

In addition to the evaluation efforts, The Kempe Center provided the three child welfare agencies with training and technical assistance opportunities customized to each site's needs. There were also several opportunities for shared learning among the sites as well as with the other federal grantees in the 2011 Family Connections grant cluster.

## **The Evaluation**

The Kempe Center and Casey Family Programs conducted an evaluation of FGDM in these three child welfare agencies that were specifically selected because of their long-term implementation of FGDM. The existing national and international FGDM research has predominately focused on the effects of FGDM on the safety, permanency, and well-being of children in foster care. Less is known about the effectiveness of FGDM on children and families receiving in-home services, the best configuration of FGDM services to meet the needs of this specific population, and FGDM's impact on the over-representation of racial and ethnic groups within a child welfare system. Therefore, it was the intent of the NPLH evaluation to focus on:

- The effectiveness of FGDM on children and families receiving in-home services.
- How FGDM can meet the needs of children and families receiving in-home services.
- The effectiveness of FGDM in supporting culturally diverse populations.

For the purpose of this report, the term Family Group Conferencing (FGC) will be used when referring to the particular FGDM process that was the focus of the NPLH project evaluation in South Dakota. South Dakota, like many other states, implements various family meeting models.

The family group conference (or FGC) that was implemented through this project is closely aligned with the core elements of FGDM, as established by the National Center on FGDM.

## **Description of This Report**

This report focuses solely on South Dakota's implementation of FGC as a part of the NPLH Project. It describes the history of FGC in South Dakota, the methodology used for the implementation evaluation, and provides an overview of FGC practice during the project timeframe, including the challenges, successes and other system changes/factors that impacted the implementation of FGC.

## **A Brief History of FGDM in South Dakota**

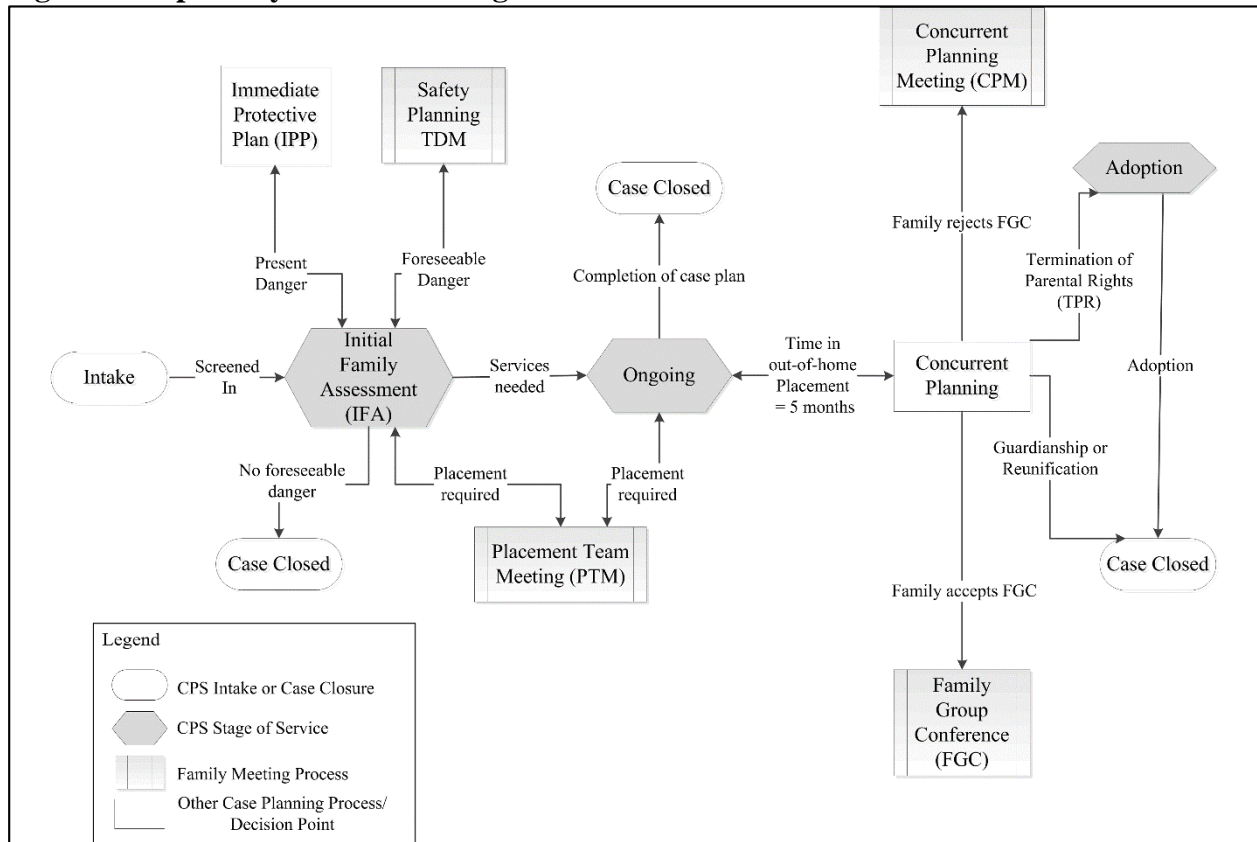
This section discusses the history of FGC and other family engagement processes that have been implemented to address child protection and welfare concerns, and attempts to illustrate the use of FGC in various communities around South Dakota, with a focus on the Rapid City region, and the Pine Ridge and Rosebud Reservations. While the No Place Like Home grant focused on the implementation and evaluation of FGC in Rapid City, it is important to understand the history of FGC and the deeper context of this work in South Dakota that preceded 2011.

In 2005, South Dakota Department of Social Services (SDDSS) began implementing family group conferencing in the Rapid City region. This practice change was made by local agency leadership and not in response to legislative mandates or availability of additional funding. They received some training and consultation from FGC practitioners from the Pine Ridge and Rosebud Reservations. In 2006, the SDDSS in the Rapid City region expanded its family

engagement work by implementing Team Decision Making models, resulting in a continuum of decision making processes that could be used throughout cases.

As Allan and Maher (2014) have noted, the number of family meeting models offered by Rapid City SDDSS has expanded to include four different types that vary by timing, circumstance and the degree to which they are family- or agency-led. Figure 1 illustrates the four family meeting types which are utilized during the Initial Family Assessment (IFA) (e.g. the investigation or assessment) and Ongoing stages of service. Listed in order of the degree to which they are family-led, from least to most, these meetings include: Concurrent Planning Meetings (CPMs), Placement Team Meetings (PTMs), Safety Planning Team Decision Making Meetings (TDMs), and FGCs. What this diagram shows is the complexity, interplay and intersection of these various family meeting types and how cases may flow through this child welfare system.

**Figure 1. Rapid City Case Flow Diagram**



*Source:* Interviews with agency staff, December 8-9, 2011, as reproduced in Allan and Maher (2014).

While there were multiple meeting types being implemented by SDDSS, the NPLH project focused on the FGC model that emerged from New Zealand practice and policy, the first country to codify this decision making process. In Rapid City, four of the six core elements were embedded in the implementation of FGC, including: 1) an independent Coordinator to bring the family group and agency personnel together; 2) time and attention to find and prepare the family group; 3) time for the family group to meet privately to create an initial plan to the agency’s concerns; and 4) preference is given to the plan developed by the family if it meets agency concerns. According to Rapid City’s policy, FGC is the preferred model for permanency and concurrent planning, and FGCs can be used at any point in a case, including prior to a placement,

but preferably no later than five months into placement at which time a concurrent plan is required to be identified and documented in the Child's Case Plan Evaluation.

## **Expansion of FGC in South Dakota**

South Dakota's involvement in the NPLH project, contributed to the statewide expansion of FGC in child welfare beginning in 2012. Since then, SDDSS secured other funding sources, independent of the NPLH grant, to intensively train all SDDSS staff on FGC. Post-training, SDDSS has also worked to sustain FGC practice by building its training capacity through supporting training of several experienced FGC coordinators so that they are equipped to provide FGC overview training on an ongoing and as needed basis.

South Dakota clearly benefitted from Rapid City's early efforts in FGC by using the experience and knowledge gained to inform implementation and practice efforts across the state. Rapid City's Regional Manager, FGC Supervisor and Coordinators provided, consultation and technical assistance to other regions, sharing their foresights, learnings and available materials, including agency guidelines and forms. As a result, in many ways, regions did not have to reinvent the wheel, which has been especially beneficial in rural regions where there are fewer staff and resources to assist with FGC implementation.

### ***Rosebud and Pine Ridge Reservations***

In addition to the public child welfare agency's efforts to implement FGC, it is important to recognize the history and work of a number of tribal child welfare agencies in implementing family-led, culturally responsive decision making processes, like FGC. In 2004, Sicangu Child and Family Services (SCFS) on the Rosebud Reservation and Oglala Sioux Tribe Child Protection Services (OST-CPS), formerly known as Lakota Oyate Wakanyeja Owicakiyapi

(LOWO), on the Pine Ridge Reservation began offering FGDM to families, with support and collaboration from Casey Family Programs. Staying true to its indigenous roots, the FGDM model implemented at SCFS and LOWO child welfare agencies has been adapted to reflect the traditional and cultural practices of the Lakota. Community members have been trained as FGDM facilitators, which encourages and supports community engagement and ownership, and supports the sustainability of FGC, regardless of formal system involvement. The overarching purpose of FGDM for SCFS and LOWO is to decrease disproportionality of American Indian children in the South Dakota child welfare system through prevention, and keeping children within their family/kin network and community, while reducing placement with non-native families (Marcynyszyn, Small Bear, Geary, Conti, Pecora, Day, and Wilson, 2012).

### **The Evaluation of FGC in SDDSS: Description of Methodology**

The proposed South Dakota evaluation design for the NPLH project included an analysis of FGC process, outcomes, and costs. The target population in South Dakota was initially families who were assessed as needing in-home services and where the primary maltreatment type was neglect. The design was a quasi-experimental “intent-to-treat” method whereby every case meeting this eligibility requirement would be assigned at random, upon completion of the initial assessment to either a control group or an intervention group. The quasi-experimental design aspect meant that only some families in the intervention group would receive FGC meetings; however, all entrants would be defined as part of the intent-to-treat population. Differences in both characteristics and process measures for the sub-group receiving FGC meetings and the sub-group that does not would be analyzed as part of the study. These group assignments would be made over a nine-month period beginning six months after funding. Families previously assigned to the intervention or control group reentering the system after case closure would not

be eligible for participation. Of concern was the understanding that sample sizes would likely be rather small (estimated at approximately 6 to 10 FGC referrals per month) and could potentially limit the types of outcome analyses that could be conducted in the site. In fact, the referrals fell far short of targets which resulted in substantial revisions to the evaluation design, , including the expansion of the target population to families experiencing an out-of-home child placement where reunification was the goal within the next six months.

## **Process Evaluation**

In addition to the outcome evaluation, process evaluation activities were conducted over the life of the grant to gain greater understanding around the implementation of FGC in Rapid City, South Dakota. Due to the drastically low referrals to the target population and continued low referrals following expansion to dual target populations, the outcome evaluation was unable to be executed as designed. As a result, the focus of the evaluation shifted focus to a more comprehensive process/implementation evaluation than was originally planned, which included a need to understand the low referral rates. In particular, focus groups were held three times over the life of the project: 2011, 2013 and 2014. The 2011 focus groups convened caseworkers, supervisors, FGC coordinators, and family/kin. In 2013, caseworkers and supervisors participated in two focus groups. And, in 2014 evaluation staff met with caseworkers and supervisors, in two focus groups, while the regional manager, FGC supervisor and coordinators were interviewed individually. The findings from the interviews and focus groups provided much context to the understanding of barriers to FGC practice observed in Rapid City during the NPLH project.

## **Process Evaluation Findings – Challenges and Barriers to holding FGCs**

A number of barriers to both referring and holding FGCs, particularly with the in-home services population were uncovered through process evaluation activities in Rapid City over the life of the project. They are described in more detail below.

### ***Low Referrals***

The first year of the NPLH project was used to provide training and coaching on FGC, to document the case flow and develop the evaluation design and procedures. In the second year, in-home service families were eligible to be referred for FGCs and tracked for the evaluation. However, in the first four months of the enrollment period, as a result of Rapid City SDDSS not opening any cases to in-home services, no NPLH-eligible FGC referrals were made. The initial evaluation design stipulated that all in-home cases would be eligible for FGC referral and would be randomized in order to create a treatment group and a control group. When this lack of referrals was discovered, and after consultation with Rapid City SDDSS project leadership, the following decisions were made to widen the scope of FGC referrals: 1) the randomizer was suspended to include all in-home FGC referrals in a treatment group (i.e., not assign anyone to the control group), and 2) an additional study population, consisting of cases where a child was currently in placement but was anticipated to return home within six months and would be referred for an FGC to aid in reunification planning, was added. It was hoped that these changes would increase FGC referrals and provide the evaluation with the numbers sufficient to conduct an outcome evaluation as well as generate a greater degree of process data related to FGC implementation (particularly around practice challenges and systemic barriers) and fidelity. After an 18-month enrollment period, there were 21 total referrals, and of those, only nine resulted in an FGC being held, a substantial shortfall from the approximately 100-150 referrals that were anticipated prior to data collection.



In attempting to understand the reasons for the low numbers of in-home services in Rapid City, information was gathered, as aforementioned, via focus groups and interviews with project leadership. There were several factors identified that contributed to both the lack of in-home service cases and the lack of FGC referrals, generally, which are discussed below.

### ***Parental Consent***

According to the Kempe Center (2013, p. 20), “FGC is based on the principle that children have the right to their family group coming together to plan for them and therefore, those with connections and relationships to the children are entitled to participate. Thus, agency policies and protocols need to be structured in a way that allows the implementation of this core FGC principle.” Nationally, legislation was developed to involve and include extended family and kin when a child has been removed from home. The Fostering Connections Legislation of 2008 (P.L. 110 - 351) has provisions that require child welfare agencies to notify maternal and paternal relatives within 30 days of the child’s removal from the parent, of their options to participate in the care and placement of the child, to ensure that these relatives are able to participate in decision making and planning for the child’s care, through efforts such as family group decision making. In South Dakota, according to focus group participants, in order to contact and engage extended family and kin in the FGC process, parental consent must be obtained first. This practice has posed numerous barriers: 1) decreasing referrals for FGCs; 2) empowering parents/caregivers to not authorize the FGCs to occur; and 3) limiting the widening of the extended family circle when FGCs are held.

### *Lack of In-Home Cases*

Initially, as stated previously, the NPLH project focused solely on evaluating in-home cases that received FGCs. When evaluators became aware of the lack of in-home cases, focus groups were held to understand this phenomenon. During a focus group, a caseworker who had worked at the Rapid City office for many years commented on the decline of in-home cases, noting it had been some time since she recalled having an in-home case. She went on to say that many of her coworkers agree that, while they cannot pinpoint the cause, the number of in-home cases has decreased over the last several years. While the caseworkers could identify the phenomenon but not a reason for the occurrence, the evaluation team sought perspectives from other respondents. Additional reasons for the number of in-home cases declining in recent years were identified as:

1. The start of the NPLH project coincided with a change in practice through the launch of the Structured Team Response (STR) meeting process in December, 2012. The STR process includes the Safety Plan Determination instrument that may be relatively rigid in what it directed casework staff to do, which may have impacted in-home referrals. This new model has shifted the agency's focus to one that is more focused on safety than risk. It may have screened out more lower-risk cases than were screened in prior to the use of this instrument, which may include those that would have typically gone to in-home services. In April, 2013, the Safety Plan Determination instrument was revised to include conditions for return, which were not included in the instrument previously.
2. In instances where a child has been determined to be endangered, the agency will seek to implement a Present Danger Plan to manage the danger of the child in lieu of law enforcement taking legal protective custody while the Initial Family Assessment is being completed. One criteria that needs to be met to implement a Present Danger Plan is that

the parent is willing and able to consent to a plan. If they are not willing and able to consent than law enforcement will take temporary legal custody of the child, placing the child in an alternative care home such as kinship or foster care, thus bypassing ongoing in-home service provision. In Rapid City, this situation occurs several times a month.

3. There are a high number of tribal transfers when SDDSS transfers a family's case to the tribe in which they are affiliated in situations where the court has become involved (e.g. when SDDSS has taken emergency custody of a child/youth). SDDSS reports that these transfers happen within a couple of days of court involvement which triggers tribal notification, per ICWA guidelines. This may also impact the number of in-home cases, as some of those transferred cases may have ended up as SDDSS in-home cases had jurisdiction not transferred to the tribal authority.

### ***FGC Referral Process***

Prior to the NPLH project, Rapid City SDDSS did not refer cases in the IFA stage of service for FGCs; rather, this decision making process was reserved for placement cases. Thus, early on in the project, there was no referral policy in place for those IFA circumstances. Instead, it appears that workers initiated FGC referrals based on if they, and/or their supervisors, thought FGC should occur for those assessments on a case-by-case basis. This put caseworkers and supervisors in the role of gatekeepers as to which families were referred for FGC, This stands in contrast to the protocol for FGC for placement cases; in those instances when children came into care, caseworkers had a checklist of things to do that included FGC referral thus automating the referral process. For IFA cases, there was no checklist or other prompt to refer, which appears to have resulted in a much greater degree of individual discretion at the worker level in making referrals. It was unclear whether caseworkers agreed with this new policy of conducting FGCs

during the assessment stage of service. This was corroborated by focus group respondents who stated that they didn't think there were referrals being made for FGC for in-home cases, and not all supervisors thought they should do an FGC for in-home cases.

In order to increase referrals, it became a requirement for workers to refer all eligible in-home cases to the FGC supervisor, who would then make the determination for which meeting process to pursue (FGC or other decision making meeting). Rapid City leadership reported that the referral requirement may not have increased the number of FGC referrals, specifically, but it did increase their meeting numbers overall.

Another issue, which caused notable delays in FGC referrals, became evident during the last six months of the NPLH project. FGC referrals were initiated upon the completion of the IFA, which were supposed to be completed within a 45 day timeframe. However, some IFAs were not being completed on time, with some taking between 60 and 120 days to complete.

#### ***Other Decision Making and Planning Meetings Available***

As noted earlier in this report, in addition to FGC, there are several other family meeting types for decision making and planning that are utilized in SDDSS, including team decision making meetings, placement team meetings, and concurrent planning team meetings. This continuum of meetings may also be a contributing factor to low FGC referrals because if an FGC isn't held, there are other options to engage the parent and possibly others in the family system. As discussed below, some of these other options have more lenient standards and protocols, and thus, may be easier to conduct.

While the focus groups did not focus on the successes or barriers of each decision making meeting offered by SDDSS, a number of themes emerged. In general, non-FGC meetings

typically require less time between referral and the meeting occurring, as well as less time required for meeting participation. While the FGC requires a release of information from parents, most other meeting types (with the exception of the TDM) do not. In addition, privileging the family voice and perspective in decision making is a hallmark of the “family-led” FGC, with the other meetings being more “agency-led.” This means that the parent and other family members may be included in these other meetings, but SDDSS has a larger say and the ultimate decision making authority in those meetings, rather than the family leading the development of a plan with the facilitator helping the family and agency reach a negotiated agreement in an FGC. Related to this, one focus group member described the following: “When looking at termination, the case must have either an FGC or Case Planning Meeting. If FGC doesn’t happen for whatever reason, then we must have a CPM.” Staff report that CPMs occur more frequently than FGCs. However, Rapid City’s policy states: “FGC is the preferred process to identify the concurrent plan. Only if an FGC is not an option, would a CPM be needed.” It is unclear the reasons for this discrepancy between policy and practice but issues related to staff preference and parental consent may well be pertinent here.

### ***Family/Kin Engagement***

Rapid City staff reported that it was a challenge to identify and engage family/kin members to participate in the FGC process. They listed the following contributing factors:

- Among the reservations and tribal communities throughout South Dakota, there is a lack of trust in SDDSS because there have been lawsuits and media that have alleged violations of the Indian Child Welfare Act on the part of SDDSS. Some Native American families state that they do not want to work with social services or have anything to do

with them. As stated in a focus group: “Sometimes we’ll schedule a meeting where a handful of people agree to come but then only one shows up.”

- Rapid City is a very transient area especially for those who are going back and forth to reservations, which may result in difficulty in locating family/kin supports. As a result, there may not be enough family to hold an FGC.
- In cases where family/kin members are located and those individuals are dealing with issues related to poverty and/or addiction, participation in the FGC process can be impacted negatively.
- In a focus group, staff reported that FGCs have been halted midway and were converted into another type of meeting. An example given was that, at an FGC, a case supervisor decided that the family was not “behaving well enough,” and therefore they were going to stop the FGC to have a different type of meeting where the agency had more decision-making authority. FGCs being changed, mid-meeting, were noted more than once by staff, though it is unclear how often this occurred. It is possible that this was a reflection of staff perceptions about FGC. Early in the project, an online survey was administered to staff participating in the NPLH project, which included questions pertaining to FGC knowledge and values. Staff response regarding their agreement about whether families know how to construct thorough plans for resolving their issues hovered around “slightly agree”.
- A strength that the agency can build on moving forward in the implementation of FGC is derived from the same survey noted above. Staff responses for the majority of statements indicating the level of agreement about various about FGC values ranked between “Agree” and “Strongly agree.” The highest overall rating staff gave was to the following

statement: *All families are entitled to be respected by CPS* (a rating of “strongly agree”).

This demonstrates that for the most part, staff beliefs are aligned with FGC values, which is a vital component for successful FGC implementation and practice.

## **Other contextual factors**

In addition to the various changes made around safety assessments and the system/practice barriers that have impacted the NPLH project, other contextual factors, including some changes to FGC practice, have occurred over the project period that may have impacted the NPLH project. Some of those additional contextual systems factors are described in more detail below. It is unclear whether these changes are a direct result of the NPLH project or other system reforms that were being implemented at the time.

### ***FGC Supervision Efforts***

The FGC supervisor reported some changes in how she supervises the FGC coordinators, based on her increased awareness of practice issues, challenges and their specific developmental needs as coordinators. The changes she made included: (1) tracking cases more fully, including having the coordinators track outcomes with a focus on asking “How can we show that we are having an impact?” (2) focusing her efforts on regularly working with the coordinators on their roles and boundaries and encouraging them to follow FGC values; and (3) helping coordinators deal with their feelings regarding other staff’s biases about FGC and/or families.

### ***FGC Coordinator Efforts***

The FGC coordinators noted some changes as well, which included: (1) keeping cases open so that they could go back and follow up. Family members were told that another meeting could be called at any time if there were issues. There was a period where the coordinators were

closing cases after the initial FGC was held. As a result of a gaining a greater understanding of plan implementation and follow-up after the initial conference, the coordinators began keeping referrals open longer, after the initial FGC, to track the need for follow-up meetings; (2) coordinators reported having more tools as a result of training, coaching and consultation, and are putting more efforts into preparation; and (3) as an effort to decrease system-imposed practices during FGCs, they no longer hang flip chart paper on the wall to share information during the information sharing stage of the conference. They realized that this method was artificial and system-influenced and potentially invasive to private family time, as most families don't typically hang flip chart paper around the room when they are making decisions and plans.

### *Staff turn-over*

Throughout the project period, SDDSS contended with staff turnover at the Coordinator position, which can be difficult since the FGC team consisted of a supervisor and 2.5 coordinators. One coordinator remained in her position throughout the project, while the other 1.5 coordinators changed at various times, which impacted workloads and extra time was needed to train new coordinators and reach full productivity.

### *Casework Staff Perceptions*

In focus groups, some staff talked about changes in their perceptions of FGC, noting that: (1) the purpose of the FGC seemed to be clearer to participants; (2) staff seemed more familiar with what can be accomplished in FGCs and that positive relationships can result; and (3) workers were doing a better job of explaining impending dangers to families, which may result in a more well-informed understanding by families about the concerns they need to address in their decision making and planning.



### ***Media and Legal Context***

Around the beginning of the project, there was a notable amount of unfavorable media at local, state and national levels, and lawsuits filed, which contributed to negative portrayals of child welfare practice in South Dakota. In focus groups, some staff shared that they were concerned about being personally sued, and were hesitant to leave children in their homes because of the risk involved.

### **Conclusion**

South Dakota's child welfare climate appears to be one in which policies are highly complex and have changed frequently in the past few years; the system is risk-averse; and, in Rapid City, there were ongoing workload challenges based on staff turnover throughout the project period. They are not unlike other jurisdictions attempting to implement this approach under similar circumstances. Cultivating fertile soil for family-led decision making to grow and thrive in such a climate takes a great amount of time and perseverance. This is clearly evident in Rapid City, given their dedicated state, regional and front-line leadership who believes in the capacity of families, and their continued efforts to offer FGC over the last decade. Multi-level leadership support, however, is not sufficient for any intervention to take hold and be sustained.

What the process evaluation has uncovered is that, despite the best intentions of leadership and staff at all levels, there are substantial barriers to implementing FGC. These barriers exist and interact in multiple and complex ways in Rapid City's practice landscape. If SDDSS leadership is committed to family-driven decision making, then solutions, generated and supported by various stakeholders, to maximize the implementation of FGC are needed. The NPLH team proposes a few ideas for SDDSS' consideration in this regard:

1. **Streamline the number of “family engagement or decision making processes.”** As Figure 1 demonstrates and focus group participants noted, there are numerous meetings that SDDSS can hold for decision making purposes. All of these meetings, other than the FGC: retain the system stakeholders as the primary decision-makers; privilege service providers’ perspectives over those of the family group; and minimize, unintentionally or not, the extended formal and informal family network in the lives of children who are involved with public child welfare. Being able to distinguish between the different meetings with various names, triggers, purposes and processes is a substantial system barrier and challenge, not only for agency staff but for parents and family members who find themselves involved with the child welfare system.

When child welfare systems, like SDDSS, offer such a wide continuum of “family meetings,” it is not surprising that the meetings that are more fully embraced and implemented with the greatest frequency are the ones that require the least amount of change—both programmatically and philosophically—for system professionals. Many of these other “family meeting” types may be presenting a false illusion, by name and perhaps by intent, that the family group is an active partner in decision making, when that is not the case in many instances. Basically, the more options for the child welfare to stay the same, the more likely the system—as influenced by the workforce—will embrace traditional forms of decision making.

Therefore, SDDSS may wish to consider streamlining the number of different meetings they offer, perhaps one for emergency decision making or imminent placement decisions and another type of meeting that permits the agency to gather the family group to partner in decision making.

2. **Hold FGCs earlier in the life of a case.** A core component of this practice model is giving enough time and resources to finding and preparing the family group to take part in the decision-making process. This results in the convening the widest family circle and promotes family leadership and accountability. However, given the complexity of child welfare systems and the many timeframes in which decisions are required, it may be a near to impossible task for FGC coordinators to balance adequate preparation, including widening the family circle, with system timeframe mandates. The alternative—too little time between referral and conference— may result in the decisions at the FGC being more of a rubber stamp on previous decisions made by system providers. It is reasonable to conclude that the earliest decisions in a case influence the trajectory of that case. If the wider family group is not an active participant early on in the life of a case, then their ability to influence outcomes is likely marginalized.
  
3. **Institute Statewide policies that align with the Notice of Relatives provision of the Federal Fostering Connections to Success and Increasing Adoptions Act of 2008** that mandates that “within 30 days after the removal of a child from the custody of the parent or parents of the child, “the State shall exercise due diligence to identify and provide notice to all adult grandparents and other adult relatives of the child (including any other adult relatives suggested by the parents), subject to exceptions due to family or domestic violence.” State interpretations of due diligence vary, but three questions developed by the National Institute for Permanent Family Connectedness (2014) are helpful to consider: 1) What has been done to identify adult grandparents and other relatives; 2)

What has been done to locate (current addresses) those identified; and 3) What has been done to provide notice to those we have identified.

Coordinators and other system professionals have noted that the implementation of FGC in South Dakota has been negatively impacted by the need to secure parental consent to both approach others to widen the circle, thereby reducing the family constellation to participate and to hold the FGC.

While it is always good practice to engage parents and seek their buy-in to the FGC process, given the Fostering Connections legislation and the FGDM Best Practice Guidelines, SDDSS may wish to revisit the parental consent policy or procedure to align it with the provisions of the Fostering Connections Act. While it is recommended that FGDM Coordinators seek first the participation of parents or caregivers, requiring consent should not be a barrier to engaging the wider family network and moving forward with the conferencing process. An expansive interpretation of this Notice of Relatives provision could assist SDDSS to overcome these barriers that have decreased the likelihood of FGCs occurring.

4. **Engage community partners, tribal leaders, and advocacy groups** in the installation of FGC. Historically, in review of other communities' implementation efforts, groups who have been historically been marginalized by the child welfare system—grandparents and other relatives, tribal representatives and leaders, parents, and other advocates—tend to support the concept of FGC as a decision-making construct. These groups may be allies external to the system that SDDSS could engage to re-ignite family-led decision making as a cornerstone to the system policies and structures and help hold the system accountable.

5. **Institute a regular continuous quality improvement process** whereby management and others review SDDSS data regarding the flow of cases into in-home services. It was surprising to the NPLH team that the number of in-home cases was lower than anticipated when the project was scoped. Gaining an understanding as to why that phenomena has occurred will be instrumental to SDDSS in the implementation of in-home services.

This project provided SDDSS, the Kempe Center and Casey Family Programs with an opportunity to better understand the implementation of FGC in South Dakota. We are grateful to SDDSS for their implementation efforts and for their staff and leadership dedication to this practice and the evaluation. We are hopeful that some of what is contained in this report propels FGC forward in serving South Dakota's families.

## References

Allan, H. & Maher, E. (2014). *All in the family: Variations in the use of family meetings in child welfare*. *Child Welfare*, 92, (6). 97-110.

Kempe Center for the Prevention and Treatment of Child Abuse & Neglect. (2013). *Guidelines for Family Group Decision Making in Child Welfare*. Aurora, CO: author.

Marcynyszyn, L., Small Bear, P., Geary, E., Conti, R., Pecora, P., Day, P., and Wilson, S. (2012). Family Group Decision Making with Lakota Families in Two Tribal Communities: Tools to Facilitate FGDM Implementation and Evaluation. *Child Welfare*, 91, (3), 113-125.

National Institute for Permanent Family Connectedness. (2014). Defining Due Diligence:

Identifying Relatives for Foster Youth. Retrieved August 7, 2015 from

[http://familyfinding.org/resourcesandpublications/due\\_diligence.html](http://familyfinding.org/resourcesandpublications/due_diligence.html)

## Appendix H. Results of Bayesian Model Averaging

**Table 3**  
Results of Bayesian Model Averaging, With Four Best Models

Characteristic/attitude	$p(\beta \neq 0)$	EV	SD	Model 1	Model 2	Model 3	Model 4
Primary job:	0.5						
Supervisor/Program							
director/Trainer		0.003	0.053				
Coordinator/facilitator		0.005	0.079				
Current work area FGDM	29.8	0.270	0.470				
Caseload carried	61.7	-0.516	0.474	-0.892	-0.840	-0.810	-0.915
Years in position	1.2	0.000	0.006				
Years in child welfare	1.1	0.000	0.003				
Experience with family meetings:	0.0						
Some		0.000	0				
A lot		0.000	0				
Racial/ethnic minority	39.3	0.263	0.378		0.733		0.606
High job satisfaction	1.0	0.000	0.033				
Family preservation proclivity	1.7	0.009	0.107				
Family ability to plan	1.0	0.001	0.041				
Workload change due to family meetings:	20.5						
Decrease		0.224	0.478				
Increase		-0.035	0.176				
Services composite	42.2	1.860	2.401		4.646	4.533	
Can find services	17.8	0.126	0.312				
Confidence in services for needs	52.0	0.328	0.356	0.692			0.649
Supervisor competence	2.9	0.006	0.045				
Supervisor leadership	1.0	0.000	0.018				
Vision, professionalism, and commitment	1.7	0.003	0.036				
State:	0.0						
State 2		0.000	0				
State 3		0.000	0				
				<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
BIC				-1121.10	-1121.07	-1120.32	-1119.78
Model posterior probability (%)				8.5	8.4	5.8	4.4

Note.  $p(\beta \neq 0)$  = probability (%) the predictor is in the correct model. EV = model-probability-weighted average  $\beta$ . SD = model-probability-weighted standard deviation. BIC = Bayesian information criterion. FGDM = Family Group Decision Making.