

Friends of the Children

Social Innovation Fund (SIF) Program Evaluation Final Report

January 20, 2021

Submitted to:

Susan Walsh, PhD
Friends of the Children

Submitted by:

ICF Incorporated, LLC
9300 Lee Highway
Fairfax, VA 22031





Table of Contents

1. Introduction	4
A. Program Background and Problem Definition	5
B. Overview of Prior Research	7
C. Overview of the Implementation and Outcomes Study	8
D. Research Questions.....	9
1. <i>Outcomes questions and findings</i>	9
2. <i>Implementation questions and findings</i>	10
E. Contribution of the Study.....	10
1. <i>Level of Evidence Generated by the Study</i>	11
2. <i>Strengths and Limitations of the Study</i>	11
3. <i>Connection of this Study to Future Research</i>	12
2. Study Approach and Methods	12
A. Implementation Study Design	12
1. <i>Fidelity to Program Design</i>	13
B. Outcome Study Design	13
C. Sampling, Measures, and Data Collection.....	15
1. <i>Sampling</i>	15
2. <i>Measures and Instruments</i>	17
3. <i>Data Collection Activities</i>	19
3. Statistical Analysis of Outcomes	20
A. Analysis Approach	20
B. Formation of Matched Groups.....	21
C. Treatment of Missing Data.....	22
D. Analysis Model/Type	22
1. <i>Tests for Statistical Significance</i>	23
2. <i>Adjustment for Multiple Comparisons</i>	23
3. <i>Assessment of Effect Sizes</i>	23
4. Findings, Lessons Learned, and Next Steps	24
A. Implementation Findings	27
1. <i>Hours Spent with Youth</i>	27
2. <i>Contacts with Caregivers</i>	30
3. <i>Contacts with Schools</i>	32
4. <i>Fidelity of Professional Development Capacity Building</i>	34
B. Summary of Outcome Findings	36
C. Lessons Learned, Study Limitations, and Next Steps	37
1. <i>Lessons Learned for FOTC Program Replication and Expansion</i>	37
2. <i>Study Limitations</i>	39
3. <i>Lessons Learned working with Federal Initiatives</i>	39
Study Logistics	40
A. Protection of Human Subjects.....	40
B. Budget and Timeline	40
C. Evaluation and Program Staff Involvement	40



Concluding Remarks.....	40
Appendix A: Friends of the Children Logic Model	42
Appendix B: Data Elements Requested versus Received by Site.....	43
Appendix C: Measures of Program Fidelity by Month and Site.....	44
Appendix D: Selecting Youth from Child Welfare Agencies	47
Appendix E: Baseline Equivalence for WA DCYF Groups	48
Appendix F: Selecting Youth from Partner Schools	50
Appendix G: Baseline Equivalence for Pooled School Groups - all	51
Appendix H: Baseline Equivalence for Pooled School Groups to Maximize Quality of Comparison Group	



Feasibility and Outcomes Studies

EXECUTIVE SUMMARY

Friends of the Children (FOTC) is a Portland, Oregon-based nonprofit, whose sites across the country serve high-risk youth ages 4–18+. Starting in 2016 with the Social Innovation Fund (SIF) award, the National office of Friends of the Children (hereafter “FOTC National”) funded seven sites. Two sites, Seattle, WA and Boston, MA, had implemented the FOTC model previously and wanted to serve more children (expansion sites). Five sites were new to implementing the model (replication sites). These sites were Austin, TX; Charlotte, NC; Central Oregon, OR; San Francisco, CA; and Los Angeles, CA. Through a competitive process, FOTC National selected ICF, a strategic consulting firm based in Fairfax, Virginia, along with Dr. Carla Herrera, Senior Advisor, to conduct an evaluation from September 2016 through October 2020.

FOTC intentionally identifies and enrolls children facing the most significant barriers to future success—barriers like systemic poverty, structural racism, childhood trauma, foster care, underfunded schools, and homelessness. Program youth experience an average of four Adverse Childhood Experiences (ACEs) by the time they turn five. The program uses paid professional mentors (Friends) who sustain an intensive mentor-child relationship over 12+ years of the child’s growth and development, beginning at 4-6 years of age. Specifically, FOTC uses an evidence-informed, replicable model to promote child well-being and help children avoid negative outcomes. Friends create individual youth plans designed to align mentoring activities with children’s needs, areas of specific interest, and intermediate and long-term outcomes. One-on-one long-term relationships are used to develop core assets¹ to achieve intermediate outcomes (school success, social and emotional development, improved health, making good choices, and plans and skills for the future) and long-term outcomes (high school graduation, post-secondary education/employment, avoiding the juvenile justice system, and early pregnancy).

Seattle and San Francisco enrolled youth in 2017, 2018, 2019, and 2020. The other five new sites enrolled children in 2018, 2019, and 2020. Across the seven sites, FOTC enrolled 542 children over four years. This is 113% of the target enrollment of 480. During the study period, 49 youth were dismissed from the FOTC program, mostly because they moved out of the service area.

This study added to FOTC’s growing evidence base, which includes 15 years of findings from its annual third-party evaluation of the FOTC-Portland affiliate, which has found that: (1) 83% of program graduates earn a high school diploma or GED; (2) 93% have avoided the juvenile justice system; and (3) 98% have avoided early parenting. Prior research from our ongoing randomized controlled trial (RCT) study has also found that caregivers² report FOTC children exhibit significantly more behavioral strengths, more positive school behavior, and less externalizing problem behavior than control group children.

This evaluation was designed to produce a preliminary level of evidence for several reasons. First, a key component of the evaluation was the implementation study, which FOTC had never conducted. To be

¹ These nine core assets are growth mindset, positive relationship building, find your spark, problem solving, self-determination, self-management, perseverance/grit, hope, and belonging.

² The term caregiver may apply to a parent or another adult who has custody of the youth.



successful in any expansion or scaling effort, FOTC needed to thoroughly examine how the program functioned, and the extent to which its implementation was in line with the FOTC model, what is referred to as “fidelity.” Implementation study activities were only conducted with the seven sites where the program was implemented—there was no comparison group—making a more rigorous design untenable. Another reason for targeting a preliminary level of evidence was the exploratory nature of several aspects of the data collection and analysis evaluation plan. This study advances the evidence base by describing the intended content, quantity, and structure of the FOTC intervention, and assessing the extent to which Friends and children received the intended service. FOTC also now has insight into how best to scale, replicate, and sustain the program, including shedding light on contextual factors that may influence how FOTC works.

This evaluation report covers four main program evaluation activities completed by ICF. The first was an implementation study conducted in seven sites. The second was a caregiver survey conducted in 2019 and 2020 with the caregivers of youth enrolled in FOTC during all but the last year of the study (i.e., caregivers whose child was enrolled in 2020 were not asked to complete a survey). The third activity was a school study, which involved analyzing school administrative data to answer exploratory research questions related to school behavior, academic proficiency on standardized math and reading assessments, and school absenteeism. ICF looked at trend data by site for treatment only youth as well as comparing treatment and comparison youth when possible. The fourth activity was a child welfare study to answer exploratory research questions related to time in foster care, time to permanence, and the number of removals, placements, and re-entries to the child welfare system for FOTC-enrolled youth and a comparison group of youth.

The implementation study centered on monthly hour-long phone calls with the executive director and/or program director at each site. Approximately twice each year, ICF met with Friends, supervisors, or other program staff instead of leadership. Observational visits were conducted at two sites. Notes from monthly site calls and observations served as the raw data from which themes—such as facilitators and challenges to implementation—were identified for the implementation study. ICF also obtained reports from FOTC’s Efforts to Outcomes (ETO) system. The current report addresses three research questions for the implementation study:

1. To what extent did Friends implement the FOTC program model with program participants?
2. Which features of the FOTC settings appear to be associated with the successful implementation of the program model with children?
3. Was the Friends’ professional development capacity building (e.g., Friend training and supervision) implemented with fidelity?

In general, sites implemented the model with fidelity, although they struggled to spend the targeted 16 hours of time with each youth participant every month during the study period. Boston (an existing site), and Austin (a new site) came closest to meeting this goal, on average, each month. All sites exceeded the targeted number of two contacts with schools and caregivers each month. Barriers and challenges to meeting these goals, and the training and supervision provided to Friends, are addressed in this report.



ICF fielded a caregiver survey in 2019 (n=77) and 2020 (N=263) to explore potential program influences on children’s socio-emotional learning, school behavior, family stability and parenting efficacy, as well as caregiver’s assessment of program quality. The caregiver survey was designed to answer the four research questions listed below. Findings from both the 2019 and 2020 surveys indicate perceived gains in many areas, and generally strong positive support for the FOTC program.

1. Does participation in FOTC improve parent’s perceptions of gains in children’s socio-emotional learning and their own parenting efficacy?
2. What strategies are effective in surveying caregivers of children participating in the program?
3. Does participation in FOTC improve parent’s perceptions of family stability?
4. Does participation in FOTC improve positive relationship building and self-management?

The school study involved FOTC-enrolled youth selected through schools and a comparison group. The study took place from 2016- 2020 and covered up to four years of school administrative data depending on when the site began enrollment and how many years of data was provided by schools. Thus, the school study focused on outcomes for students in kindergarten through third grade. As explained in more detail later in this report, ICF had planned to have, and received, comparison group data in two sites—Seattle and rural Central Oregon. We unexpectedly obtained a large dataset from Seattle Public Schools (SPS) and used these data to construct a pool of youth from which to create a comparison group matched using propensity score matching (PSM). While outside the scope of our evaluation plan, ICF wanted to explore the usefulness of this dataset to strengthen the school data analyses in Central Oregon and other SIF sites, despite the obvious limitations of a comparison group comprised almost exclusively of youth from a single school district³ and treatment youth from disparate states. Using a subset of treatment youth with perfect matches in the PSM model (80 treatment youth and 69 comparison youth), there were no significant differences between the treatment and comparison group youth across our three research questions during the early years of their participation in the program:

1. Does participation in FOTC improve children’s school behavior as measured by disciplinary incidents?
2. Does participation in FOTC improve grade-level academic performance as measured by reading and math proficiency on standardized tests?
3. Does participation in FOTC improve school attendance as measured by absenteeism?

The final component of our study was a child welfare study. This study was conducted in one of the two sites, Seattle, that provided FOTC services to children in the child welfare system (there was an insufficient sample size to conduct analyses for the rural site in Central Oregon). The Washington Department of Children, Youth, and Families (WA DCYF) provided ICF with the comparison group of youth. In this study there were 36 treatment youth and 27 youth in the comparison group. ICF used data from the Adoption and Foster Care Analysis and Reporting System (AFCARS) and the Child and Adolescent Needs and Strengths (CANS) assessment to compare outcomes for the treatment and

³ In the large, pooled comparison group of 198 youth, 195 were from SPS and three were from the Oregon school districts who provided comparison youth as part of ICF’s original evaluation plan. In the smaller group of 80 youth, in which only youth with perfect matches were retained, 79 were from SPS and one was from the Oregon school districts.



comparison groups. The current report addresses six research questions for youth selected from the child welfare system and a comparison group of youth.

1. Does participation in FOTC result in a decreased length of stay in foster care?
2. Does participation in FOTC result in a reduced number of placements?
3. Does participation in FOTC result in a faster time to permanency?
4. Does FOTC participation increase the likelihood of achieving permanency? [New question not in SEP]
5. Does participation in FOTC reduce number of re-entries to care? [New question not in SEP]
6. Does participation in FOTC reduce number of removals? [New question not in SEP]

There was one statistically significant difference; treatment youth had a decreased length of stay in foster care (an average of 399 days) compared to the comparison group (an average of 576 days).

There have been no substantive changes to the evaluation timeline, budget, program, or evaluation team since submitting the SIF Evaluation Plan revised December 2017. As an extension of the results included in this report, FOTC has contracted with ICF to provide evaluation services for Los Angeles (an original SIF site) and New York (a non-SIF site) until August 31, 2021 (“the 2021 2Gen evaluation”). Key program activities include administering a caregiver survey in spring 2021, exploring the feasibility of analyzing administrative data on child welfare outcomes, and providing subject matter expertise and evaluation support to sites enrolling families in 2-Generational (2Gen) program services. Findings from the 2021 2Gen evaluation will inform the design of a more rigorous 2Gen evaluation with the goal of having the model accepted as an evidence-based practice for child welfare prevention.

FOTC has a five-year scaling goal to serve children in 25 communities by 2025. The network currently includes services in 22 locations and the six newest sites launched are implementing the 2Gen program model (bringing the total to eleven 2Gen locations). Programmatically, FOTC’s next steps are to strengthen its capacity for training and supporting its teams about the 2Gen program model, while continuing to evaluate how to achieve the greatest positive impact with families. FOTC will partner with caregivers and children to ensure their voices are instrumental in driving the strategic 2Gen program development.

1. INTRODUCTION

This evaluation report covers four main program evaluation activities completed by ICF. The first was an implementation study conducted in seven sites. The second was a caregiver survey conducted in 2019 and 2020 with the caregivers of youth enrolled in FOTC during all but the last year of the study (i.e., caregivers whose child was enrolled in 2020 were not asked to complete a survey). The third activity was a school study covering grades kindergarten through 3rd grade- the early years of the FOTC program, which involved analyzing school administrative data to answer exploratory research questions related to school behavior, academic proficiency on standardized math and reading assessments, and school absenteeism. ICF looked at trend data by site for treatment only youth as well as comparing treatment and comparison youth when possible. The fourth activity was a child welfare study to answer exploratory research questions related to time in foster care, time to permanence, and the number of removals, placements, and re-entries to the child welfare system for FOTC-enrolled youth and a comparison group of youth.



FOTC's funder, AmeriCorps (formerly the Corporation for National and Community Service),⁴ provided reporting guidance for implementation, feasibility, and impact studies. The structure of this report follows those reporting guidelines, with small modifications made to accommodate the four studies within the larger program evaluation. For reporting, ICF has classified the school study and caregiver survey as *implementation* activities and the child welfare study as the *outcome* study. This evaluation did not include an impact study.

Several changes were made to the evaluation since SEP approval. Enhancements included:

1. Expanding the caregiver survey from a pilot study to a census approach in which all eligible caregivers were invited to complete a survey in 2019 and 2020. This change was made to obtain quantitative data from caregivers after focus groups conducted with caregivers revealed the benefits of the FOTC program to the caregivers of enrolled youth—not just the youths themselves.
2. Conducting an exploratory study in which pooled data from five sites were combined, and comparison youth were identified using propensity score matching (PSM), to examine school administrative data. This change was made to increase the sample size and explore the usefulness of using PSM to create a suitable comparison group.

There were a few planned activities that did not occur. These were:

1. The small number of children selected from the child welfare system in rural Central Oregon precluded ICF from reporting on child welfare outcomes for these children. Specifically, ICF's data sharing agreement with the state stipulated that reporting could not be done on a sample size less than 10 for fear that re-identification would be possible. This resulted in ICF having fewer children in the child welfare outcomes study.
2. ICF was able to obtain school administrative data from six of the seven sites. The school district in Boston denied ICF's Institutional Review Board (IRB) application, stating they did not see the benefit of the study beyond the one school the local site was currently working with. This resulted in ICF having fewer children in the (pooled) school study and the Boston site not having longitudinal analyses specific to the youth enrolled in their site.
3. ICF conducted two site visits in April 2019 and had planned two additional site visits in spring 2020. Due to site and school closures, as well as travel restrictions during the COVID-19 pandemic, these visits could not occur. However, this did not substantively impact the implementation study because those activities were wrapping up by the time of the planned visits and the project team's consensus is that little new information would be gained from the in-person visits.

A. PROGRAM BACKGROUND AND PROBLEM DEFINITION

FOTC is a Portland, Oregon-based international nonprofit that was serving approximately 950 youth ages 5–18+ in 2016, prior to this evaluation. FOTC intentionally identifies and enrolls children facing the most significant barriers to future success—barriers like systemic poverty, structural racism, childhood trauma, foster care, underfunded schools, and homelessness. Program youth experience an average of four Adverse Childhood Experiences (ACEs) by the time they turn five.

⁴ Corporation for National and Community Service. (2016) *Social Innovation Fund Evaluation Reporting Guidance: Feasibility, Implementation and Impact Study Reports*. Washington, D.C.



FOTC uses an evidence-informed, replicable model for helping children tap into their innate resilience to avoid the negative behaviors and life outcomes that can stem from early adverse experiences (the FOTC program model is found in Appendix A). Unlike many traditional mentoring programs, FOTC hires paid professional mentors and uses a longitudinal service approach. FOTC identifies 4–6 year old children from schools, foster care, and community partner agencies, pairing them with highly trained, skilled mentors (called “Friends”) who sustain an intensive Friend-child relationship over 12+ years of a child’s growth and development. On average, each child will have two to three Friends during their time in the program: one Friend who specializes in the early years of development, followed by a Friend trained to work with teens. The Friends are paid, full-time employees with a roster of 8–12 children each. Friends primarily meet with children one-on-one, spending an average of 3-4 hours per week (14-16 hours/month) with each child.

Friends provide developmentally appropriate experiential teaching and modeling of healthy behaviors, directly coaching children on social-emotional skill development areas, referred to as “core assets” in the FOTC program model. The nine core assets are growth mindset, positive relationship building, find your spark, problem-solving, self-determination, self-management, perseverance/grit, hope, and belonging. Friends also assist youth in achieving individualized “road map” goals in the areas of school success, making good choices, planning skills for the future, prosocial development, and healthy habits.

This one-on-one approach is based on research showing that the single most important factor in fostering resiliency in children is a caring and consistent relationship with an adult.⁵ The model has been successful in addressing persistent challenges that young people face by striving to close the opportunity gap and breaking the cycle of intergenerational poverty for children who have experienced abuse or neglect and/or multiple, compounding risk factors related to education, social/emotional wellness, social capital, and community or family supports.

According to the National Center for Children in Poverty, 21% of children in the United States live in poverty.⁶ Exposure to multiple poverty-related risks increases the odds that children who are socioeconomically disadvantaged will demonstrate less social and emotional competence, lower executive functioning skills, and more behavior problems.⁷ Moreover, poverty-related risks have been linked to negative social outcomes, such as high dropout rates, teenage parenthood, and delinquent behavior.⁸

In year 1 of the Social Innovation Fund (SIF) project, FOTC replicated their model in one new community, San Francisco, and expanded the capacity of an existing FOTC site, Seattle, WA. Beginning in year 2 (2017–2018), FOTC recruited children for their program from four additional new sites: Austin, Texas; Central Oregon; Charlotte, North Carolina; and Los Angeles, California; and scaled an existing site in Boston, Massachusetts. By the end of the 4-year SIF project, it was expected that: (1) the FOTC network

⁵ Shonkoff, J. P., & Meisels, S. J. (2000). *Handbook of early childhood intervention*. New York: Cambridge Universities Press.

⁶ Jiang, Y., Granja, M. R., & Koball, H. (2017). *Basic facts about low-income children: Children under 18 years, 2015*. New York: National Center for Children in Poverty. Retrieved from http://www.nccp.org/publications/pdf/text_1170.pdf

⁷ Webster-Stratton, C., & Reid, J. (2008). Strengthening social and emotional competence in young children who are socioeconomically disadvantaged. In W. H. Brown, S. L. Odom, & S. R. McConnell (Eds.), *Social competence of young children: Risk, disability, and intervention* (pp. 185–203). Baltimore: Brooks Publishing.

⁸ Merrell, K. W., & Gueldner, B. A. (2010). One size does not fit all: Adapting social and emotional learning for use in our multicultural world. In K. W. Merrell & B. A. Gueldner (Eds.), *Social and emotional learning in the classroom: Promoting mental health and academic success* (pp. 83–102). New York: Guilford Press.



would have grown by approximately 50%, serving a total of 1,430 children nationwide; (2) the 480⁹ additional children selected for the program as a result of SIF would be on track to achieve the FOTC long-term outcomes of high school graduation, avoidance of teen parenthood, and avoidance of juvenile justice system involvement; and (3) a successful implementation evaluation of the FOTC program would have occurred. Both of those goals have been exceeded.

At the close of this study, there were 493 youth still enrolled in FOTC across the seven participating sites. In the school study, ICF analyzed data for 419 youth (221 treatment youth and 198 comparison youth). In the child welfare study, ICF analyzed data for 63 youth (36 treatment youth and 27 comparison youth).

B. OVERVIEW OF PRIOR RESEARCH

From inception, FOTC has used research and programmatic data to develop and refine its model to meet the needs of the children it serves. FOTC has invested in evaluation strategies at the local and national levels to measure the ability of the program model to effect long-term change for children in the highest risk circumstances. Data related to each youth's academic success and social-emotional development are entered into the ETO program database. The ability to track and monitor academic data for each child in the program provides Friends, program managers, and executive leaders with real-time information to guide programmatic decisions which ensure that youth remain on track for high school graduation and post-secondary enrollment.

For the past 15 years, NPC Research has conducted an annual third-party outcome evaluation of the Portland FOTC chapter and has also conducted evaluations at the Seattle FOTC chapter for the past seven years. Their recent research indicates that the majority of FOTC participants succeed in achieving the program's three long-term outcomes (Kissick et al., 2016):

- 83% of program graduates earned a high school diploma or GED
- 93% of program graduates avoided the juvenile justice system, and
- 98% of program graduates avoided early parenting.

Since 2007, FOTC has also participated in an ongoing, third-party RCT led by the Oregon Social Learning Center and more recently the University of Washington (The Child Study). The study (N = 278; n = 156, intervention group; n = 122, control group) includes a racially and ethnically diverse sample (46% African American, 18% Latino, 18% multi-racial, and 14% white) of boys (n = 130) and girls (n = 147) considered to be at "high risk" for future problems. Study participants were selected from FOTC's Boston, New York, Portland, and Seattle programs. The research team collected and analyzed data from youth, caregivers, teachers, Friends, and schools. The study reported few or limited significant outcomes in the first four years of the study. According to a 2017 publication,¹⁰ compared to children in the control group at the five-year time point, caregivers of FOTC children are reporting that FOTC children have significantly:

- more behavioral strengths, as measured by increased family involvement, interpersonal

⁹ Note that the SEP indicates that 604 additional youth would be served through SIF. This should have been reduced to 480 when the grant period was shortened from five to four years.

¹⁰ Eddy, J. M., Martinez, et. al. (2017). A randomized controlled trial of a long-term professional mentoring program for children at risk: Outcomes across the first 5 years. *Prevention Science: The Official Journal of the Society for Prevention Research*. doi:10.1007/s11121-017-0795-z



- strengths, and school functioning,
- more positive school behavior, and
- less externalizing problem behavior.

The results are encouraging because they found that the FOTC program shifts caregiver perceptions of a child to be more positive. Having caregivers that have positive attributions about and confidence in a child encourages further success. Such encouragement can be a protective factor and lead to later positive outcomes. In 2020, the University of Washington secured a new five-year NIH grant award to complete the RCT, examining outcomes on study participants when they are 19 and 21 years old.

In addition to positive outcomes for children, an independent cost-benefit analysis conducted by the Harvard Business School Association of Oregon (HBSAO) on FOTC's Portland program found that the program achieved significant long-term cost savings.¹¹ The study estimated that for each child that FOTC-Portland assisted, the community saved \$971,000—roughly \$843,000 more than program costs to implement FOTC. For every 100 FOTC-Portland program graduates, society gained 24 more high school or college graduates; 59 fewer teen parents; and 30 fewer people entering prison at a young age and correspondingly fewer crime victims. Based on HBSAO's analysis, FOTC saved various units of government almost \$1 million per participant over the 12+year span of the program. Notably, savings to the child welfare system were not considered in this study because at the time FOTC had yet to begin intentionally selecting children from the foster care system.

As noted, the SIF allowed FOTC National to expand its services and enroll children selected from the child welfare system in two states—Washington and Oregon—and to enroll youth in another site (Los Angeles) that enrolled the children of adults who had themselves been involved in the child welfare system. ICF was able to conduct a small, exploratory outcomes study to assess the usefulness of child welfare data in Seattle to measure key outcomes of interest. The lessons learned from this approach added immense value to FOTC for expanding into the foster care population.

C. OVERVIEW OF THE IMPLEMENTATION AND OUTCOMES STUDY

ICF conducted a 4-year evaluation of the FOTC program using a mixed-methods design in seven sites. As noted in the introduction, the caregiver survey and school study are categorized as part of the implementation study; the child welfare study is categorized as an outcome study. ICF collected primary data from caregivers. School administrative data were obtained from six of the seven evaluation sites. ICF negotiated a process with several school districts to keep data de-identified (described in more detail in a later section).

The implementation study drew from programmatic data found in FOTC's ETO database, data obtained via monthly calls with executive directors, program directors, and other site staff, and direct observations of program activities in select sites. ICF also collected data from the caregivers of FOTC-enrolled youth twice during evaluation—in 2019 and 2020. In 2019, ICF used a combination of email, text, mail, and phone invitations and/or reminders to field an online caregiver survey. These same methods were used in 2020; additionally, ICF mailed nonresponders a print copy of the survey and distributed survey packets to sites for in-person distribution. The school study had two parts—one,

¹¹ Hamilton, B., Cruver, C., et. al. (2012). *Breaking the cycle of poverty: Social return on investment*. Portland, OR: Harvard Business School Association of Oregon.



analyzing school behavior, academic proficiency on standardized math and reading assessments, and school absenteeism for FOTC-enrolled youth over time within a site. The other, looking at these same measures comparing treatment and comparison youth in select sites. As explained in a later section, ICF used data from Seattle Public Schools to find matched comparison youth to treatment youth in Seattle and other sites and conducted exploratory analyses.

The child welfare study answered exploratory research questions related to time in foster care, time to permanence, and the number of removals, placements, and re-entries to the child welfare system for FOTC-enrolled youth and a comparison group of youth. Child welfare data were provided to ICF directly from the WA DCYF. ICF analyzed data to compare the treatment and comparison groups from baseline to one, two, and three years after baseline. Exhibit 1 provides an overview of the school and child welfare data obtained from each site.

Exhibit 1. Overview of School and Child Welfare Data ICF Obtained

Site	School Data	Child Welfare Data
Austin	Treatment	N/A
Boston	None	N/A
Charlotte	Treatment	N/A
Central Oregon	Treatment and Comparison	Planned, not obtained due to small number of youth enrolled
Los Angeles	Treatment	N/A
San Francisco	Treatment	N/A
Seattle	Treatment and Comparison	Treatment and Comparison

D. RESEARCH QUESTIONS

1. Outcomes questions and findings

a) Confirmatory

This child welfare study did not include any confirmatory research questions.

b) Exploratory

The child welfare study addressed six exploratory research questions, listed below. The first three were included in the SEP; questions 3, 4 and 5 were added. Data were not available to answer a fourth question originally included in the SEP (Does participation in FOTC result in a reduced number of school changes?). There was one statistically significant finding (research question 1). Treatment youth averaged 399 days and comparison youth averaged 576 days ($p < .05$).

1. Does participation in FOTC result in a decreased length of stay in foster care?
2. Does participation in FOTC result in a reduced number of placements?
3. Does participation in FOTC result in a faster time to permanency?
4. Does FOTC participation increase the likelihood of achieving permanency? [New question not in the SEP]
5. Does participation in FOTC reduce number of re-entries to care? [New question not in the SEP]
6. Does participation in FOTC reduce number of removals? [New question not in the SEP]



2. Implementation questions and findings

The goal of the SIF implementation study was to explore how and why the FOTC model works by examining its replication and expansion in new sites or to new populations. Specifically, ICF sought to answer the research questions listed below. Sites generally implemented the model with fidelity, with all sites, on average achieving the targeted number of school and caregiver contacts each month. Sites struggled the most achieving the targeted number of hours with youth.

1. To what extent did Friends implement the FOTC program model with program participants?
2. Which features of the FOTC settings appear to be associated with the successful implementation of the program model with children?
3. Was the Friends' professional development capacity building (e.g., Friend training and supervision) implemented with fidelity?

The caregiver survey was initially designed to answer the four research questions listed below. Findings from both the 2019 and 2020 surveys indicate perceived gains in many areas, and generally strong positive support for the FOTC program.

5. Does participation in FOTC improve parent's perceptions of gains in children's socio-emotional learning and their own parenting efficacy?
6. What strategies are effective in surveying caregivers of children participating in the program?
7. Does participation in FOTC improve parent's perceptions of family stability?
8. Does participation in FOTC improve positive relationship building and self-management?

The school study addressed three research questions. Data were not available to answer a fourth question originally included in the SEP (Does participation in FOTC lower rates of school misbehavior such as referrals to the principal's office?). Using pooled data from four sites covering at most baseline and three years post-baseline, when youth were in grades kindergarten through 3rd grade, and a comparison group created almost entirely from a single school district (Seattle Public Schools) using propensity score matching, there were no significant changes in these measures.

1. Does participation in FOTC improve children's school behavior as measured by disciplinary incidents?
2. Does participation in FOTC improve grade-level academic performance as measured by reading and math proficiency on standardized tests?
3. Does participation in FOTC improve school attendance as measured by absenteeism?

E. CONTRIBUTION OF THE STUDY

This evaluation added to the program's existing body of evidence by examining facilitators and challenges to implementing the FOTC program in new sites and a new population—youth enrolled through the child welfare system or through referring partner agencies. Findings from the evaluation will inform the development of our new quality improvement (QI) system for training, supporting and supervising program staff. This QI system will strengthen service delivery at existing sites and ensure that new sites are set up for success during launch and on-boarding. Moreover, the evaluation will add to the body of evidence about mentoring programs in general, including exploring:



- The efficacy of recruitment and retention of high-risk children/families from the foster care system,
- Challenges and successful strategies in working with the child welfare system,
- FOTC intervention fidelity, scale-up, and replication, including whether it is possible to implement the FOTC model at scale in multiple sites,
- The quality of support provided to the Friends, including best practices on training and supporting professional mentors, and
- Program influences on children’s socio-emotional learning, school behavior, family stability and parenting efficacy, as well as program effectiveness from the caregiver’s perspective.

1. Level of Evidence Generated by the Study

As noted, this study was designed to provide a preliminary level of evidence. In four sites, ICF examined participant youth’s outcomes over time using school administrative data. The caregiver survey was given only to the parents/caregivers of youth involved in the program. There were two evaluation activities that included a comparison group—pooled school administrative data from five sites and the child welfare study; however, due to small sample sizes and incomplete data, ICF was not able to achieve a higher level of evidence (i.e., moderate or strong). These limitations are discussed in the next section.

2. Strengths and Limitations of the Study

The primary strength of the evaluation was the implementation study, which FOTC had never conducted. To be successful in any expansion or scaling effort, FOTC needed to thoroughly examine how the program functioned, and the extent to which its implementation was in line with the FOTC model, what is referred to as “fidelity.” This study advances the evidence base by describing the intended content, quantity, and structure of the FOTC intervention, and assessing the extent to which Friends and children received the intended service. The implementation study also provided insights about how best to scale, replicate, and sustain the program, including shedding light on contextual factors that may influence how FOTC works in different jurisdictions.

Another strength was the connections ICF was able to make at school districts and state child welfare agencies, and the resultant data sharing agreements and approved IRB applications. It is hoped that these relationships will facilitate obtaining data for future internal program evaluations or external third-party evaluations for FOTC National and its network sites.

While not without its own limitations, ICF created a comparison group using the large dataset provided by Seattle Public Schools. As noted, this was outside the scope of our evaluation plan, which specified comparison groups for school administrative data only in Seattle and rural Central Oregon. This strengthened our evaluation by providing a larger comparison group, which increased the probability of detecting statistically significant findings. The tradeoff was a comparison group comprised almost exclusively of youth from a single school district and treatment youth from disparate states. While matched using PSM on key demographics (gender, race/ethnicity, the school’s Title 1 status, grade level, school year, English language learner status, and special education status) and outcomes (categorical measures of youths’ disciplinary incidents, reading academic performance, and absentee rates), the conceptualization and operationalization of some of these variables was likely not uniform across school districts. Specifically, districts likely differed in how they defined a “disciplinary incident,” “English learner status,” “special education status,” and perhaps even absenteeism. Perhaps more important, comparison group students were not assessed for trauma or adverse childhood experiences, so these



factors could not be included in the propensity score model. Other study limitations are related to the sample size, the representativeness of the sample, and the strength of the comparison group. For example, across all seven sites, there were 431 youth whose data ICF wanted to include in the school administrative data study. Of these, ICF obtained data for 221 (51.3%). One site, Boston, did not provide IRB approval and therefore would not provide ICF with any school administrative data. At another site, ICF obtained data from 24% of the eligible youth. At another site, ICF obtained data for 97% of eligible youth. Another limitation is that ICF was unable to receive all data elements it requested. This hindered our ability to create strong comparison groups and to answer all research questions. Appendix B shows the school administrative data requested and received from each site.

3. Connection of this Study to Future Research

FOTC's ongoing third-party evaluations have demonstrated that the FOTC program model can change the life trajectory of children whose families have experienced systemic barriers and who have complex trauma histories. SIF enabled FOTC to expand their services in two existing sites and replicate their services in five new sites. Evaluation findings expanded FOTC's understanding of how partnering with child welfare and community partners for child selection, impacts aspects of program fidelity.

FOTC has continued to expand to new sites implementing the 2Gen model and enroll children connected to the child welfare system. Findings from the implementation study, in this report and communicated to FOTC and participating sites throughout the study, have provided FOTC with valuable insights about how and why the FOTC model works, including under what circumstances and for whom. The caregiver survey, and gathering school administrative data, have likewise provided FOTC with valuable information on the best process for collecting data from parents and guardians, and the challenges in obtaining administrative data from schools and child welfare agencies.

2. STUDY APPROACH AND METHODS

A. IMPLEMENTATION STUDY DESIGN

ICF's implementation study used a mixed-methods design that drew from programmatic data found in FOTC's ETO database, data obtained via monthly calls with executive directors, program directors, and other site staff, direct observations of program activities in select sites, a survey administered to the caregivers of children enrolled in FOTC, and school administrative data for FOTC-enrolled youth.

ICF had access to several extant data sources from which to collect programmatic information, including subgrantee applications, the ETO database, and FOTC National Friend training materials. To supplement these sources, we collected information from key program staff through phone interviews with the seven sites starting December 2017 for all sites except Los Angeles, which started July 2018. These monthly calls, which were held through November 2019, were generally conducted with the executive director and program director. Three times during the evaluation, ICF met with other site staff—twice with site Friends and once with Friend supervisors. ICF conducted two site visits—one of which had expanded to select youth through the child welfare system. While on-site, ICF interviewed program leaders, Friends, and supervisors and observed 1:1 Friend-child and group activities. ICF conducted seven interviews during these site visits.

Monthly site calls were facilitated by a team member with an advanced degree (master's degree or higher) and notes were taken by a research assistant. These notes were stored on a shared project team



site and reviewed to extract common themes and unique experiences—such as facilitators and challenges to implementation. ICF’s research focused on the key components of the FOTC logic model that indicated fidelity and dosage (contact with youth, caregivers, and schools). ICF charted these metrics by site by month throughout the study period, compared trends among and across sites, and looked for factors in the monthly call notes that explained the trends we noted. ICF distilled the recommendations from interviewed staff, and ICF’s own observations, into a set of lessons learned that FOTC might reference in supporting the launch or expansion of new sites. These are listed in bullet form starting on page 36. Program fidelity is discussed in the next section.

1. Fidelity to Program Design

The central role of a Friend, defined by the FOTC program model, is that each Friend spends a minimum of 16 intentional hours per month with each child, has a minimum of two contacts each month with the youth’s caregiver, and has a minimum of two contacts each month with their youth’s school. Caregiver and school contacts are designed to enable Friends to establish a presence in the child’s home, school, neighborhood, and community so that he or she can help advocate for their youth and become a trusted resource for the family.

We therefore focused the measurement of fidelity using the aspects of adherence and exposure,¹² which are key program performance indicators from FOTC’s logic model and captured in the program’s communal Efforts to Outcomes (ETO) case management platform:

1. average monthly hours Friends spent with youth (per child),
2. average monthly number of Friend contacts with caregivers (per child), and
3. average monthly number of Friend contacts with schools (per child).¹³

ICF obtained ETO data from July 2017 through March 2020 to assess fidelity to these program goals. As shown in Exhibits 6, 7, and 8 and Appendix C, there was variability both within and across sites in terms of youth’s exposure to the program as measured by ETO data obtained by ICF. Sites routinely exceeded, on average, the minimum number of school and caregiver contacts. While school contacts understandably dipped in the summer months, Friends exceeded two contacts in almost all other months. Caregiver contacts routinely exceeded the minimum (averaging 3.9 to 7.3 each month across the sites).

A third aspect of fidelity ICF examined was program responsiveness. FOTC participants are the youth enrolled in the program, and their caregivers—who in most cases had to consent for their child to be in the study and were asked to complete a survey. Of the 431 FOTC-enrolled youth eligible for the school study, ICF analyzed data for 221 (51.3%). The response rate for the caregiver survey was 29% in 2019 (71 of 244) and 58.2% in 2020 (263 of 452).

B. OUTCOME STUDY DESIGN

The SEP outlined a plan whereby the state child welfare agency would create an initial pool of potential participants based on FOTC criteria, listed in Appendix D. The treatment group would consist of children

¹² Dane, A.V. and Schneider, B.H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review* Vol. 18, Issue 1, Pp 23-45. [https://doi.org/10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3)

¹³ A fourth indicator, average number of monthly outings, was also reviewed but was not included in this analysis because it also tracked the amount of time Friends spent with youth, but less precisely than the average monthly hours spent with youth.



from the pool who were selected to participate in FOTC. The comparison group would consist of children in the selection pool who were not selected to participate in FOTC. Therefore, children in the treatment and comparison groups would meet the same basic eligibility criteria. ICF then planned to calculate a propensity score for each study participant.² As described by Rosenbaum and Rubin,¹⁴ a propensity score is “the conditional probability of assignment to a particular treatment given a vector of observed covariates,” in this case, the probability of a child participating in the FOTC program. Stratifying on propensity scores would have provided an adjustment for selection bias to help support strong inferences about program effects and establish a meaningful comparison group whereby the two groups would be balanced with respect to the selection variables.¹⁵

A propensity score indicates the likelihood that the study participant is included in the treatment group. By using logistic regression to predict membership in the treatment group, ICF would have identified a very close match (or multiple matches) for each child. We planned to match on demographic variables, placement-related variables (e.g., number of placements, length of time in foster care, removals from home or entries into the foster care system), risk factors maintained by state and local child welfare agencies as part of federal requirements through AFCARS such as the type of abuse or neglect that precipitated foster care system involvement, and parental risk factors affecting system involvement (e.g., drug or alcohol abuse, incarceration, inadequate housing, relinquishment). Once matching was complete, ICF would have verified the comparability of the matched pairs by comparing the matched pairs.

WA DCYF instead provided ICF with a list of comparison youth they matched to treatment youth. Treatment and comparison youth did not differ at baseline on the 22 variables tested; thus, the strength of this quasi-experimental design is that ICF was able to compare treatment youth to a group of comparison youth (Appendix E presents these baseline equivalence tests). Note that age was not included as a variable; however, when the FOTC site requested a comparison group, they requested a youth who were five or six years of age and in kindergarten to match the FOTC-enrolled youth.

There are several threats to internal and external validity that can jeopardize research designs and findings.¹⁶ Two factors jeopardized the external validity of the current study and therefore limit our ability to generalize the findings to a larger population. First, ICF could only include youth from one site (Seattle) because the other site (Central Oregon) was unable to enroll a sufficient number of youth due to administrative challenges with the child welfare agency. Second, the sample size in Seattle was small.

The primary threat to internal validity was experimental mortality. ICF was initially provided with 47 treatment and 33 comparison youth; however, the final analytic sample was 36 treatment and 27 comparison youth (ICF removed 11 treatment youth and six comparison youth). Reasons for removal included leaving the program (six treatment youth), missing all AFCARS variables (two treatment and

¹⁴ Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.

¹⁵ Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399–424.

¹⁶For discussions of internal and external validity see Miller, D.C. and Salkind, N.J. (2002). *Handbook of Research Design and Social Measurement*, 6th ed. Sage Publications: Thousand Oaks, CA and Frankfort-Nachimas, C. and Nachimas, D. (2000). *Research Methods in the Social Sciences*, 6th ed. Worth Publishers: New York, NY.



one comparison youth), a non-sequential jump in removals¹⁷ (one treatment youth), and a discharge from foster care prior to their FOTC enrollment date or because they lacked post-baseline data (two treatment and five comparison youth). The last dataset ICF received for the outcome study was AFCARS data ending March 2020.

C. SAMPLING, MEASURES, AND DATA COLLECTION

As noted in the introduction, this study included an implementation study—itself composed of many component pieces including a caregiver survey in 2019 and 2020 and a school study looking at change over time within a site for FOTC-enrolled youth and change over time for treatment and comparison youth. The outcomes child welfare study answered exploratory research questions related to time in foster care, time to permanence, and the number of removals, placements, and re-entries to care for FOTC-enrolled youth and a comparison group of youth. Within each section, we have addressed each of these four study components. None of the data were weighted.

1. Sampling

Recruitment into the FOTC program was conducted by each site—ICF was not involved in recruiting or selecting youth for the FOTC program. Appendix F provides the steps to enroll youth based on observations conducted in partner schools. Exhibit 2 shows the number of youth enrolled in FOTC during the study period, the number dismissed in the same period (9% overall), and the net number of youth remaining in the program at the close of the study. Across the four cohorts, 8% of youth (n=42) were four years old at enrollment, 24% (n=134) were five, 55% (n=305) were six, 12% (n=68) were seven, and 1% (n=7) were eight.

Exhibit 2. SIF Enrollment and Dismissals per Cohort

SIF Cohort	Initial Enrollment	Dismissed from FOTC*	Final Enrollment
Cohort 1 2016-17	65	11	54
Cohort 2 2017-18	203	24	179
Cohort 3 2018-19	211	14	197
Cohort 4 2019-20	63	0	63
Total	542	49	493

* These youth were dismissed at different times during the study period—not just during their cohort year. Most of the youth were dismissed because they moved out of the service area. As youth were dismissed from FOTC, some sites enrolled youth from their wait list.

For the implementation study, ICF relied primarily on data collected during calls with staff at the seven sites. Most months, we spoke with executive directors and/or program directors because, given their leadership positions at the sites, they were most knowledgeable about current and planned activities at their location in addition to articulating the challenges and facilitators to program implementation. Several months, we interviewed Friends or Friend supervisors to obtain data about training and

¹⁷ AFCARS data is reported in March and September each year. There were instances in the AFCARS data received when the number of removals from home for a child was not sequential (i.e., from 1 to 3 instead of from 1 to 2). For example, a child in the sample was in their first removal from their home in March 2018. In September 2018, this child was in their third removal from their home. During the six months between March and September 2019, we knew this child experienced a second removal, but the dates for that second removal were not provided. Because we could not determine the start and end date of the second removal, we could not accurately calculate the child’s net length of stay in foster care. We therefore could not answer research questions related to the child’s length of stay in foster care, number of placements, and time to permanence since FOTC enrollment.



supervision both within their site and from FOTC National. Calls with six of the seven sites started in December 2017; calls with staff in Los Angeles started in July 2018. These regular calls lasted through October 2019, with final phone interviews conducted in spring 2020 to capture updates from November 2019 through March 2020, especially related to program challenges resulting from COVID-19. In all, ICF conducted 114 phone interviews—12 with Los Angeles staff, 14 with Seattle, 17 with Austin, Charlotte, Central Oregon, and San Francisco, and 20 with Boston. Monthly calls were not conducted every month due to scheduling conflicts, holiday breaks, and two planned site visits that replaced the monthly call.

In addition to the monthly calls, ICF conducted one visit at two sites. In Seattle, ICF conducted interviews averaging 45 minutes each, with the executive director, program director, program manager, and two team leads who provide supervisory support to Friends. In Charlotte, a joint interview was conducted with the executive director and project director, and a separate interview was conducted with the vice president of training and operations. Executive and program directors were interviewed, as with the monthly calls, given their leadership role in the organization. Team leads and the vice president of training and operations were interviewed to glean additional information about training and supervision.

Caregivers were eligible for the 2019 survey if their child (or children—some sites allowed siblings) enrolled in the program prior to 2019. This allowed caregivers to have approximately one year in the program before the survey. Of the 268 youth enrolled at the time of the survey,¹⁸ the caregivers of 244 (91%) returned a consent form indicating their willingness to participate in the study. Of these, surveys were completed by 71 (29%). Caregivers were eligible for the 2020 survey if their child (or children) enrolled in the program prior to 2020. Of the 480 youth enrolled at the time of the survey, the caregivers of 436 (91%) returned a consent form indicating their willingness to participate in the study. Of these, surveys were completed or partially completed by 303 (69%).

For the school study, data were only requested for youth enrolled in cohorts 1, 2, and 3 (i.e., youth enrolled in 2019-20 were not included) so that ICF would have at least one year of post-test data to examine. As shown in Exhibit 3, ICF analyzed data for 58% of youth, overall—with coverage ranging from zero (in Boston where we did not obtain IRB approval) to 97% in Charlotte. As noted, 9% of youth were dismissed from FOTC.

Exhibit 3. School Selected Youth Enrollment, Dismissals, and Analytic Sample by Site

Site	Enrolled in FOTC	Dismissed from FOTC	Eligible Sample	Analytic Sample
Austin	87	5	82	37 (45%)
Boston	46	4	42	0 (0%)
Central Oregon	35	4	31	25 (81%)
Charlotte	36	1	35	34 (97%)
Los Angeles	65	6	59	14 (24%)
San Francisco	104	14	90	79 (88%)
Seattle	50	5	45	32 (71%)
Total	423	39	384	221 (58%)

¹⁸ For ease of calculation, youth enrollment does not include youth dismissed since they were dismissed at different times during the study period. Thus, the percentage reported is likely an overestimate of the eligible children in the program at the time of the survey.



For the child welfare study, the Seattle site enrolled 51 youth from the child welfare system in cohorts 1, 2, and 3; eight youth (16%) were removed due to leaving FOTC. Of the 43 youth remaining, ICF analyzed data for 36 treatment youth (84%).

2. Measures and Instruments

For monthly meetings with program site staff, ICF generally relied on standard questions addressing staffing, training, child selection, Friend activities, and relationships/partnerships with schools, caregivers, and communities. Several calls were structured interviews with Friends and/or addressed a specific topic (i.e., ETO, training, administrative functions). Additionally, each ICF site liaison included follow-up items based on the prior month's call. For example, a December call would include questions related to a "Friendsgiving" event that site staff held in November. For the two site visits, ICF created consent documents, a site observation checklist, and interview guides. As with the monthly call, there were standard items across these instruments, and unique items based on the specific site.

The data received from the six school districts varied in quantity, completeness, and level of granularity. Data also differed by grade level and across years within the same site. For example, some school districts provided children's raw test scores for reading and math while others provided proficiency levels. When raw scores were provided, ICF used materials provided by the school districts or found online to determine whether children did or did not meet the benchmark for their grade level for these subjects. When detailed proficiency levels were provided (e.g., not yet meeting, approaching, meeting, or exceeding expectations) ICF recoded the data to arrive at a binary variable denoting proficiency below or at/above. Across all sites, the same naming conventions and variable values were used to indicate children's reading and math proficiency so that data could be easily compared and analyzed across and within sites. The same steps were needed to recode and standardize demographic data for youth so that data from different sites could be merged for propensity score matching and analysis. Creating the variables needed for analysis included:

- Calculating absent rates and tardy rates from provided data (i.e., using variables such as the number of days enrolled during the year, attendance rates, the number of tardy instances, etc.).
- Determining whether children's reading and math scores were below, at, or above the school district's definition of benchmark competency.
- Identifying each child's baseline school year and school years representing one year, two years, or three years post-baseline to assess longitudinal performance across outcomes.
- Recoding variables as needed to use in the propensity score matching process.
- Labeling the values of categorical variables based on discussions with data providers and accompanying codebooks.

ICF relied primarily on the AFCARS codebook¹⁹ for the child welfare study. Interestingly, the WA DCYF did not retain AFCARS variable names. For example, the AFCARS variable to indicate "date of first removal" is *Rem1Dt*, but this variable was labeled *FIRSTRMVL* in the Washington data. This and other discrepancies necessitated extensive cleaning and coding prior to data analysis. To calculate *length of stay in child welfare*, ICF computed the net amount of time (in days) a child was in foster care *after* FOTC

¹⁹ https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/afcars-foster-care-file-codebook.pdf

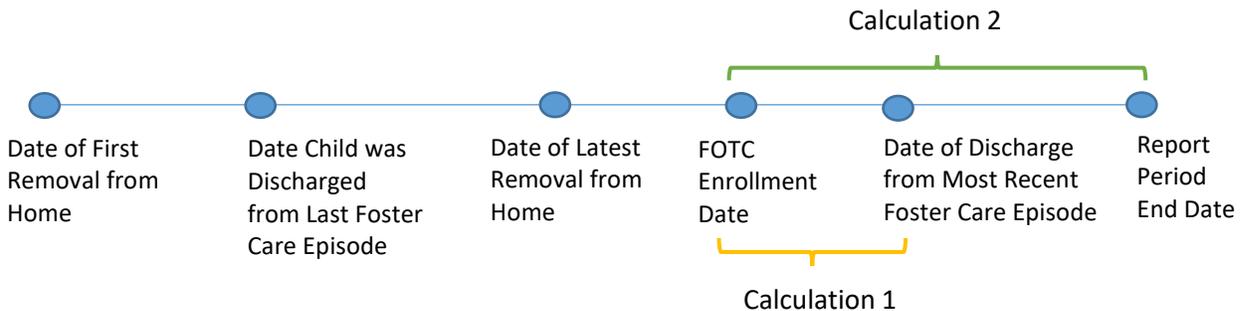


enrollment. Any time a youth spent outside of foster care, such as with his or her family, was not subtracted. Using the AFCARS data, ICF calculated:

1. The number of days between a child’s FOTC enrollment date and the discharge date of their most recent discharge if that date was *after* their FOTC enrollment date and if the child’s last removal date from their home was *before* their FOTC enrollment date.
2. The number of days between a child’s FOTC enrollment date and the most recent report period end date if the child was not discharged during that report period, if the child’s last removal date from their home was *before* their FOTC enrollment date, and if a child’s FOTC enrollment date was *before* the most report period end date.

The number of days from each calculation was summed to determine the net length of stay in foster care. This process is shown in Exhibit 4.

Exhibit 4. Length of Stay Date Calculations



The data received by ICF at six-month intervals included a variable that counted the *number of foster care placements* for the child’s most recent removal from their home—the variable was not cumulative. ICF used longitudinal data from previous reporting periods to calculate the number of total placements (i.e., for the child’s first removal and any subsequent removals). To capture the cumulative number of foster care placements each child had experienced since FOTC enrollment, the summation of all placements for all removals experienced at the report period end date *prior* to FOTC enrollment (i.e., baseline) and that at the report period end date *right after* FOTC enrollment were calculated. The difference between these totals represents the number of *additional* placements a child experienced after enrolling in the FOTC program. For example, at their respective baseline report period, a treatment child in the analytical sample had experienced three placements after their first removal. However, after FOTC enrollment and for the most recent report period, the child had four reported placements after their first removal. Thus, this treatment child had one additional placement since FOTC enrollment. Overall, the *additional* number of placements experienced since FOTC enrollment was calculated for children in both research groups.

To calculate *time to permanence*, ICF started with the net length of stay in foster care explained at the beginning of this section. However, only day counts for children who were discharged *and* reached permanency during the study period were compared. Discharges to permanency include reunification



with parents or principal caretaker(s), living with other relatives, adoption, and guardianship.²⁰ These reasons as well as others provided in the AFCARS data are presented in Exhibit 5.

Exhibit 5. Permanency-related Discharge Reasons

Discharge Reason	Definition ²¹	Related to Permanency
Reunification with Parents or Primary Caretakers	The child was returned to his or her principal caretaker(s)' home.	Yes
Living with Other Relatives	The child went to live with a relative other than the one from whose home he or she was removed.	Yes
Adoption	The child was legally adopted.	Yes
Emancipation	The child reached majority according to the law by virtue of age, marriage, etc.	No
Guardianship	Permanent custody of the child was awarded to an individual. Includes both relative and non-relative guardianships.	Yes
Transfer to Another Agency	Responsibility for the care of the child was awarded to another agency: either in or outside of the State or Tribal service area.	No
Runaway	The child ran away from the foster care placement.	No
Death of Child	The child died while in foster care.	No

3. Data Collection Activities

ICF collected primary data from sites during monthly calls, which ended in October 2019. However, final phone interviews were held in spring 2020 to capture updates from November 2019 through March 2020, especially related to program challenges resulting from COVID-19. Site visits were conducted in April 2019. Caregivers participated in surveys in 2019 (August 12th to October 29th) and 2020 (August 6th to October 16th). ICF did not collect any primary data for the school study or the child welfare study—relying on data provided directly by school districts and child welfare agencies. Youth were four or five years of age when first enrolled in FOTC; thus, primary data collection from these program participants was not attempted.

ICF did not begin receiving school data immediately due to the time required to secure IRB approval and execute DSAs. When the first school data file was received from a site, it contained data from earlier years to coincide with the year youth were enrolled in FOTC so that those data could serve as baseline measures. For example, if a site first provided ICF with data after the Fall 2018 semester, ICF requested data from the 2016-2017 and 2017-2018 school years as well. After these initial data requests, ICF generally received school administrative data to coincide with the end of each semester of the school year for each site.

ICF received bi-annual child welfare data to coincide with AFCARS reporting schedule (each March and September); we requested reports from the child’s birth through the present reporting period. The first dataset from WA DCYF was provided in October 2018 (after September 2018 data was finalized) and the

²⁰ Statewide Data Indicators and National Standards for Child and Family Services Reviews, 79 Fed. Reg. 61241 (2014) (to be codified at 45 C.F.R. pt. 1355). Link to online source: <https://www.federalregister.gov/documents/2014/10/10/2014-24204/statewide-data-indicators-and-national-standards-for-child-and-family-services-reviews>.

²¹ National Data Archive on Child Abuse and Neglect (NDACAN). (2019). AFCARS Foster Care Annual File Code Book. Retrieved from https://www.ndacan.acf.hhs.gov/datasets/pdfs_user_guides/afcars-foster-care-file-codebook.pdf.



last in July 2020 (after March 2020 data was finalized). As noted previously, ICF created a process with sites to ensure only deidentified data were provided to ICF. The process for school districts is outlined below. A similar process was used with child welfare agencies, except we used “family ID” instead of “student ID.”

1. FOTC local sites provided their school district (not ICF) with an input file with data about participating children: name, student ID, year enrolled, etc.
2. School districts removed the student ID and generated a Research ID for each child, and then provided an output file to ICF (not FOTC local site) with the Research ID, pass-through variables (e.g., year enrolled in FOTC, school name, risk and/or protective factor score), and the requested data elements outlined in the DSA.
3. School districts held the key with Research IDs through the end of the project and flagged participating children for subsequent data pulls.
4. For subsequent data pulls, FOTC local sites submitted an updated input file to their school district with information for newly enrolled children, and the Research IDs for previously submitted children.

Even though files did not contain personally identifiable information, data files were transferred to ICF using a secure file transfer protocol. Once received, data files were stored on a secure server, which were protected from unauthorized users via usernames and user-level passwords, including multi-factor authentication.

ICF encountered numerous challenges securing the ten IRB approvals needed for this study. We conducted a webinar for FOTC National and evaluation sites in August 2018 to cover key terms including IRBs, data sharing agreements (DSAs), the Family Educational Rights and Privacy Act (FERPA), consent, and release of information (ROI). During that webinar, and in a subsequent report to FOTC National the following year, ICF detailed these challenges—many of which were unique to a school district or state child welfare agency. Thus, the details of the MOUs and DSAs are not included in this report but they have been shared with FOTC. Since local sites often had MOUs with schools—not school districts—we believe the relationships established through the evaluation process will be beneficial.

3. STATISTICAL ANALYSIS OF OUTCOMES

As noted elsewhere in this report, exploratory outcomes were assessed for youth in Seattle selected through the child welfare agency. This section addresses those outcomes, assessed using treatment and comparison groups matched on 22 variables (see Appendix E).

ICF also created a comparison group to analyze school administrative data for a pooled sample of youth. To ensure the process to create that group is fully transparent, we have included information about its formation in this section. However, the comparison group created for those analyses was *exploratory* and had several limitations—as discussed in the sections that follow. So, although included in this section, the reader should keep in mind the findings should not be interpreted as “outcomes.”

A. ANALYSIS APPROACH

Matched comparison groups were used to analyze data obtained from child welfare agencies and from school districts. The primary threat to internal validity was experimental mortality. ICF was initially provided with 47 treatment and 33 comparison youth for the child welfare study; however, the final



analytic sample was 36 treatment and 27 comparison youth. Reasons for removal included leaving the program (six treatment youth), missing all AFCARS variables (two treatment and one comparison youth), a non-sequential jump in removals (one treatment youth), and a discharge from foster care prior to their FOTC enrollment date or because they lacked post-baseline data (two treatment and five comparison youth). For the school study, 384 treatment youth were eligible; ICF analyzed data for 187 treatment and 198 comparison youth.

Child welfare data were analyzed primarily using *t-tests* and chi-square analyses. The school data models were more complicated and are discussed at length in Section D. All data were analyzed using SAS.

B. FORMATION OF MATCHED GROUPS

The WA DCYF provided a comparison group to ICF for the child welfare study, the baseline equivalence of treatment and comparison youth on the 22 variables provided are presented in Appendix E.

As depicted in Exhibit 1, ICF obtained treatment and comparison data for children enrolled through schools in two sites—Seattle and Central Oregon. In Seattle, ICF had 32 youth in the treatment group and 43 in the comparison group; in Oregon, ICF had 25 youth in each group. These data were analyzed as planned and a summary of these findings is included beginning on page 22. Full reports for each site were provided to FOTC under separate cover.

For the Seattle portion of the study, ICF obtained a database from Seattle Public Schools (SPS) that contained over 100,000 records for SPS enrolled student. The ICF project team decided to use this large dataset to construct a comparison group for sites in which no comparison group had been planned. ICF narrowed the file to 13,085 non-FOTC students who were in PreK through 4th grade during the study period (the 2016-17 to 2019-20 academic years). We used propensity score matching (PSM) to find one or more children that were identical to youth in the treatment group on seven demographic characteristics²² and three of the four outcomes²³ we were interested in examining.²⁴

The PSM model looked for exact matches without replacement. Namely, the model identified comparison youth that exactly matched treatment youth across the ten variables. Once a comparison youth was matched to a treatment youth, that comparison youth was removed and could not be used for future matches. In most instances, the model found one match for a treatment child (a 1:1 match). In other instances, the model found two matches for a treatment child (a 1:2 match). Eleven treatment youth had two matches each, resulting in 11 more children in the comparison group (n=198) than the treatment group (n=187). There were two statistical differences between the research groups across demographic and outcome variables at baseline, as shown in Appendix G (statistical models included

²² Demographic variables were gender, race/ethnicity, the school's Title 1 status, grade level, school year, English language learner status, and special education status.

²³ Outcomes were categorical measures of youths' disciplinary incidents, reading academic performance, and absentee rates. Math academic performance was not used in the PSM model, but we still assessed the treatment and comparison groups for baseline equivalence on this variable.

²⁴ Categorical measures for disciplinary incidents, academic performance, and absenteeism were used in the PSM model to facilitate finding an exact match given the limited number of comparison options. For instance, finding comparison group students that matched on all other predictors, and had a specific value for attendance (like .938), was very difficult. By clustering values, we found matches within a range (e.g., a treatment youth with a .938 attendance rate would get matched to a comparison youth with a .960 rate if they were both in category 4). To assess outcomes, we used continuously distributed variables to increase the variability and increase our ability to detect differences. Effectively, we accepted some amount of error on matches, but attempted to maximize our statistical power by working with the most precise outcome data we could.



controls for variables when baseline equivalence was not achieved). ICF sought to optimize the quality and strength of the matches, so reduced the treatment and comparison groups to only those youth with perfect matches. The baseline equivalence for these 149 youth (80 treatment and 69 comparison) are presented in Appendix H.

ICF's analytic approach had several limitations. These analyses included data from four of seven sites: Austin, Central Oregon, Los Angeles, and Seattle. We never received IRB approval for school data in Boston. Although we had IRB approval for school data in North Carolina, we were not given demographic data so we could not do PSM. Additionally, while we had data from San Francisco Unified School District, we did not have disciplinary incidents (as we did for other sites), so perfect matches in the PSM process could not be found. As noted, ICF used school administrative data from a different state, provided by Seattle Public Schools, to create a comparison group of youth. Even in Oregon, where ICF had comparison group data provided by local school districts, the final comparison group included three youth from Oregon and 22 youth from SPS. While the PSM model found better matches in the Seattle data than in the Central Oregon data, there may be differences in measurements—such as how incidences and absences are assessed and recorded. Additionally, Central Oregon is much more rural than Seattle, Washington.²⁵

In four of the six sites included in the school study, caregiver consent was needed for youth's deidentified data to be provided to ICF. In another site, while consent was not needed, caregivers were excluded from the study if they did not return a FERPA form to the local site. In all sites, school data were provided only for youth enrolled in the school district where the majority of students were located; youth outside the school district, for instance in charter or private schools, were not included. As noted, ICF also could not include youth dismissed from FOTC during the study period. Lastly, there were missing data for some youth in the datasets ICF did receive. For example, academic performance data were not available for every youth for every period and academic year.

Another limitation is the small sample size, especially for analyses conducted three years after baseline. Additionally, while ICF was able to create a comparison group and control for child-level characteristics, we were unable to control for all potentially important contributors (e.g., family-level demographics). For these reasons, our results may not be representative, and we urge caution interpreting the results.

C. TREATMENT OF MISSING DATA

For most analyses listwise deletion of cases was used in statistical models.

D. ANALYSIS MODEL/TYPE

For the school study involving treatment and comparison groups, ICF used a Constrained Longitudinal Data Analysis (cLDA)²⁶ model to examine outcomes for children enrolled in FOTC (the treatment group) to those of a matched group (the comparison group). This method was chosen to account for the correlation among time points; specifically, that a youth's outcomes (e.g., reading performance, math performance, attendance) are related. While we don't test for change over time within a study group,

²⁵ Seattle, Washington's population per square mile in 2010 was 7,250.9 compared to 2,322.0 for Bend, Oregon.
<https://www.census.gov/quickfacts/fact/table/US/PST045219>

²⁶ Liang, K. and Zeger, S.L. (2000). Longitudinal data analysis of continuous and discrete responses for pre-post designs. *Biostatistics*. Vol. 62, No. 1. Pp. 134-148.

the cLDA model accounts for correlations for individuals over time. In other words, the models are hierarchical with time points nested within students where time is level-1 and students are level-2.

The cLDA model was also chosen to account for instances of missing baseline data and to ensure baseline equivalence was achieved between the two research groups across all outcomes prior to analysis. This procedure forces both research groups to share a common mean (i.e., equal value) at baseline and then adjusts post-baseline values for each research group accordingly. Exhibit 6 lists the assumptions for the cLDA model and whether these assumptions were maintained or violated, along with explanatory notes.

Exhibit 6. Model Assumptions

Model Assumptions	Assumption Held or Violated	Notes
Change across time (at level-1) is linear	Assumption held	
Time-invariant predictors (at level-2) are linear	Assumption held	
Level-1 and Level-2 errors are normally distributed and independent from each other	Assumption held, generally	Small sample size made a definitive assessment difficult
Homoscedasticity (equal variances) of Level-1 and Level-2 errors	Assumption held, generally	Small sample size made a definitive assessment difficult
Group means assumed equal at baseline	Violated	Expected since equal group means are generally only reasonable in an RCT design; controlled for baseline differences

1. Tests for Statistical Significance

Four models were created corresponding to the four outcome measures (incident count, reading proficiency, math proficiency, and absenteeism). Each model included all years of available data for the treatment and comparison groups. The cLDA model is:

$$Y_{ij} = \beta_B + \beta_i + X_i\beta_{B+1...I} + \mu_{0j}$$

where β_B represent the grand mean across Treatment and Comparison units at baseline (B), β_i represents the effect of the first time point beyond baseline B, X_i is an indicator of treatment (i.e., treatment = 1, comparison = 0) and $\beta_{B+1...I}$ represents the differential effect of treatment versus comparison units at time point i , for each time point up through the total I . μ_{0j} represents the random effect for each individual, where time points are nested within each unit.

2. Adjustment for Multiple Comparisons

ICF did not make any adjustments for multiple comparisons.

3. Assessment of Effect Sizes

The effects of interest, in particular $\beta_{B+1...I}$, are inferentially assessed using t-tests. Because we knew the parameters of interest in advance, we coded the effect estimates of interest to be tested specifically. This effect estimate is represented by the β in the effect size formula below.

$$d_{change} = \beta \left(\frac{1 - 3}{4(n_C + n_T) - 9} \right) / \sqrt{\frac{((n_C - 1)\sigma_C^2) + ((n_T - 1)\sigma_T^2)}{n_C + n_T - 2}}$$

Where β represents the difference between Tx and Cp at a point in time (e.g. at time 1), n_C is the number of comparison units at baseline, n_T is the number of treatment units at baseline, σ_C^2 represents the variance in baseline scores for comparison units, and σ_T^2 represents the variance in baseline scores for treatment units.

4. FINDINGS, LESSONS LEARNED, AND NEXT STEPS

Before presenting findings related to program fidelity, we summarize findings using school administrative data. ICF produced six separate reports for FOTC National detailing findings for all sites from whom we received school data (all sites except Boston). Seattle and Oregon’s reports include analyses for treatment and comparison youth. Reports for Austin, Los Angeles, North Carolina, and San Francisco included analyses for treatment youth only. Findings from the 2019 and 2020 Caregiver Surveys were also provided to FOTC National under separate cover.

In Central Oregon, ICF used data provided by school districts to assess the differences in disciplinary incidents, reading and math test scores, and absenteeism for youth enrolled in FOTC and a comparison group of youth matched on demographic and outcome variables. Baseline equivalence was achieved for the treatment and comparison groups across six of seven demographic variables and all three outcome variables. Our analyses controlled for grade/school year since there was a statistical difference between study groups at baseline.

Two years after baseline, there was a significant difference between treatment and comparison youth in the number of disciplinary incidents reported, with treatment youth having a mean of 0.4 and comparison youth having a mean of 0; the effect size for this finding was small. There were no significant differences in benchmark reading proficiency between enrolled youth and the comparison group and effect sizes were small. A statistically significant difference in the proportion of treatment and comparison youth testing at or above the district benchmark in math was found two years after baseline; these data were suppressed due to small sample sizes. The effect size was large at 3.75. There were no significant differences in terms of absenteeism, and effect sizes were small.

In Seattle, ICF used data provided by Seattle Public Schools (SPS) to assess the differences in disciplinary incidents, reading and math test scores, and absenteeism for school-enrolled youth in FOTC and a comparison group of youth matched on demographic and outcome variables. Baseline equivalence was achieved for the treatment and comparison groups across seven demographic variables and three outcome variables.

There were too few disciplinary incidents to examine differences between the two study groups. There were no significant differences in absenteeism and effect sizes were small. There were no significant differences in benchmark reading or math proficiency between enrolled youth and the comparison group, although some effect sizes were medium to large and in a direction favorable to the treatment group. Specifically, two years after baseline, a higher proportion of treatment youth compared to comparison youth were reading at or above the district’s benchmark. One year, two years, and three



years after baseline, a higher proportion of treatment youth compared to comparison youth were at or above the district’s benchmark for math.

In the remainder of this section we summarize aggregate/pooled school administrative data using the strongest comparison group available. This subset of data included treatment youth from Austin, Los Angeles, Seattle, and Oregon. As noted, we did not have school data from Boston and we did not have demographic data from North Carolina so we could not include them in the PSM model. Additionally, the San Francisco Unified School District did not provide disciplinary incident counts, so San Francisco youth were dropped from the comparison group for this subset analyses since perfect matches could not be found.

Research Question 1: Does participation in FOTC improve children's school behavior?

Depending on the school district, this was operationalized as disciplinary incidents, in-school, or out of school suspensions. As shown in Exhibit 7, there were no significant differences over time using the pooled data for just those youth we could match with precision.

Exhibit 7. Disciplinary Incidents for Pooled Data using Small Matched Comparison Group

Group	Treatment	Comparison	P-value	Effect Size
Baseline	0.12 (N = 53)	0.12 (N = 64)	-	-
1 Year Post	0.32 (N = 53)	0.13 (N = 64)	0.448	0.18
2 Years Post	0.15 (N = 52)	0.02 (N = 64)	0.068	0.43
3 Years Post	-----	0.00 (N = 22)	0.328	0.00

Note: 3 Years Post data for treatment youth has been suppressed due to a small sample size

Research Question 2: Does participation in FOTC improve grade-level academic performance?

This was operationalized as testing below or at/above proficiency based on standardized test scores in reading and/or math. Math scores were not provided by the Austin, Los Angeles, and North Carolina school districts; thus, given the smaller sample size that would have resulted, we present results for reading only. As shown in Exhibit 8, there were no significant differences over time in reading proficiency at or above a school district’s benchmark—again using the pooled data for just those youth we could match with precision.

Exhibit 8. Reading Performance for Pooled Data using Small Matched Comparison Group

Group	Percent of Youth Testing at or Above Benchmark		P-value	Effect Size
	Treatment	Comparison		
Baseline	41% (13 out of 32)	41% (18 out of 41)	-	-
1 Year Post	39% (16 out of 38)	44% (27 out of 61)	0.622	0.41
2 Years Post	57% (21 out of 38)	68% (23 out of 34)	0.323	0.98



Research Question 3: Does participation in FOTC improve school attendance?

This was operationalized as the percent of days a youth was absent during the total number of days school was in session. As shown in Exhibit 9, there were no significant differences over time in absenteeism on this measure using the pooled data for youth we could match with precision.

Exhibit 9. Absent Rate for Pooled Data using Small Matched Comparison Group

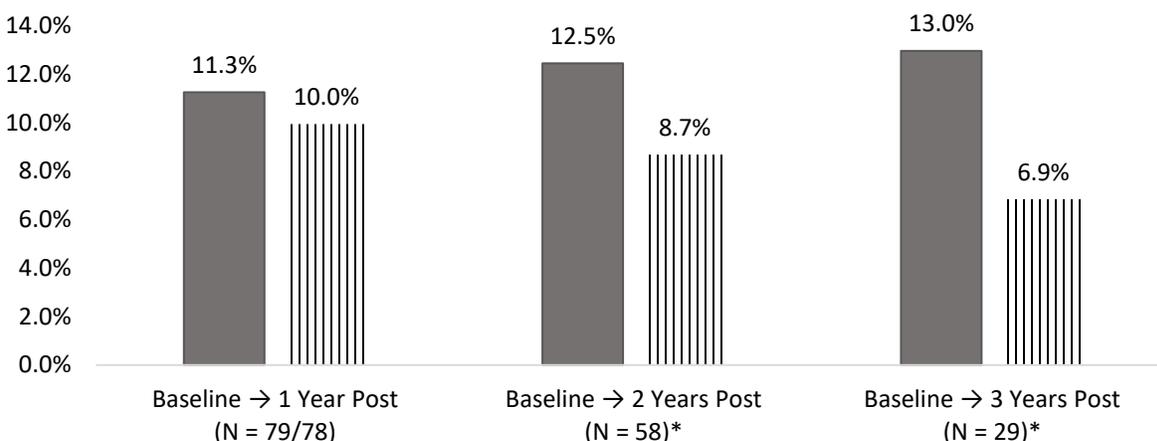
Group	Treatment	Comparison	P-value	Effect Size
Baseline	8% (N = 53)	8% (N = 64)	-	-
1 Year Post	7% (N = 53)	7% (N = 64)	0.746	0.04
2 Years Post	7% (N = 52)	6% (N = 64)	0.286	0.14
3 Years Post	-----	6% (N = 22)	0.652	0.10

Note: 3 Years Post data for treatment youth has been suppressed due to a small sample size

The pooled data masked some differences found when site-level analyses were conducted. Looking at sites for whom comparison data were available, two years after baseline, central Oregon comparison youth had significantly fewer disciplinary incidents than treatment youth. However, as noted, these results should be interpreted with caution that the Central Oregon comparison group consisted largely of students obtained from Seattle Public Schools. In Seattle, disciplinary incidents couldn't be assessed because all youth at baseline had zero, and very few had any in subsequent years. This highlights one challenge looking at school behavior for young children—especially when school behavior is reported only for the most egregious acts resulting in in- or out-of-school suspensions.

Looking just at longitudinal data over time for FOTC-enrolled youth at the other sites, there were no significant differences over time for school behavior (incident count or suspensions) or academic performance at any site. One site, San Francisco, experienced improvements in absenteeism over time; these results are presented in Exhibit 10. As shown, the change from baseline to one year later, from 11% to 10%, was not statistically significant and the effect size was small (*Cohen's d*=0.14). The other changes were statistically significant ($p<.05$) and had medium to large effect sizes (*Cohen's d*=0.40 and 0.65, respectively).

Exhibit 10. Percent Days Absent Over Time – San Francisco FOTC Enrolled Youth



A. IMPLEMENTATION FINDINGS

Recall that the goal of the implementation study was to explore how and why the FOTC model works by examining its replication and expansion in new sites or to new populations. Specifically, ICF sought to answer the following research questions:

1. To what extent did Friends implement the FOTC program model with program participants?
2. Which features of the FOTC settings appear to be associated with the successful implementation of the program model with children?
3. Was the Friends’ professional development capacity building (e.g., Friend training and supervision) implemented with fidelity?

This section examines three measures of program fidelity by site during the SIF expansion.²⁷ Sections 1, 2, and 3 in the following pages present the average monthly hours Friends spent with youth, the average monthly contacts Friends had with caregivers, and the average monthly contacts Friends had with schools, respectively. Beginning on page 32, the third research question is addressed.

1. Hours Spent with Youth

Spending 16 hours of intentional time with each youth participant each month is a core component of the FOTC model, and average monthly hours per youth is the performance indicator most closely followed by the FOTC National office. As shown in Exhibit 10, overall, the expansion sites struggled to meet this target for engagement with youth during the ramp-up period. The Austin site consistently met the target throughout the grant period, and the Boston and Central Oregon sites often met it. The Los Angeles and Charlotte sites only occasionally met the target, and the Seattle and San Francisco sites consistently fell short of the 16 hour target.

²⁷ San Francisco and Seattle started in July 2017; Austin, Boston, Central Oregon, and Charlotte started in May 2018; and Los Angeles started in September 2018. Seattle and Boston were existing FOTC chapters; the other sites were established through the SIF expansion. The Seattle site already served school-enrolled children; the SIF grant allowed them to expand to serving children recruited from the child welfare system.

Exhibit 11. Average Monthly Hours Spent with each Youth Over Project Period, by Site

Site	Hours per Month
<i>FOTC Hours Target</i>	<i>16.0</i>
Austin	15.9
Boston	14.9
Central Oregon	14.1
San Francisco	12.4
Los Angeles	12.2
Seattle	12.2
Charlotte	11.9
<i>Source: ETO Data, July 2017-March 2020</i>	

a) Barriers to Achieving Targeted Hours with Youth

Over the next few pages, we present barriers to achieving the monthly target number of hours for Friends to spend with their youth. The next section discusses factors that facilitated meeting this goal.

Seasonality presented a challenge for sites, both in terms of the school year cycle and the FOTC recruitment cycle. School closures during the December holiday season and summer break consistently resulted in reduced hours. During the first few weeks of the school year, teachers generally didn't allow Friends into the classroom. They preferred to spend time establishing classroom norms and routines before inviting Friends in. This meant that youth were only available to Friends during the handful of hours between the end of the school day and early evening. The spike in hours reported by all sites in October 2018 partly reflects that school routines and key stakeholder relationships were established by this point.

Building trust with caregivers. Similarly, the FOTC recruitment cycle drove periods of higher and lower engagement with children. During the late spring and early summer, sites invested heavily in establishing rapport with families. Sites reported that some caregivers required a great deal of relationship building before they were comfortable enrolling their child or allowing Friends to spend time with them. In cases where the enrollment and caregiver relationship building process was extended, it took longer for Friends to begin youth outings and therefore for the sites to reach the targeted performance indicator.

In Central Oregon, a new site serving a rural area, staff required a substantial ramp-up period to establish solid relationships with families. Five of the youth chosen through school-based selection in the first cohort were involved in the child welfare system. This was a source of delays for serving youth, as the site learned how to best interact with the array of stakeholders in the child's life, including caregivers, foster parents, and child welfare staff.

In Seattle, staff cited large rosters and relatively weak coordination with parents/caregivers as barriers to achieving the target hours spent with each youth. Not being able to coordinate schedules with parents meant not being able to spend the allocated hours with youth. In some instances, communication with parents was so challenging that Friends would arrive to pick up a child from school only to find out they didn't attend that school any longer (in some cases, the school was not notified either).



Staffing challenges sometimes made it difficult to achieve the targeted number of hours with youth. New sites without prior experience sometimes underestimated the amount of time needed to find, recruit, and hire Friends. The Friend interview process is lengthy, involving a first interview, second interview, panel interview, and often an outing with an existing Friend and youth. Sometimes the challenge was finding Friends with the right mix of skills, such as individuals with prior experience working with young children, who also could work with adults, and who also understood the importance of tracking outcomes in a data system.

In Boston, an existing FOTC site that used SIF funds to expand services, challenges hiring Friends depressed average monthly service hours in 2018, though once the first cohort was fully staffed the site quickly reached the target monthly service hours. Specifically, it took six months longer than anticipated to hire two of the three Friends needed to staff the first SIF cohort; slated for May 2018, these hires were made in November 2018. Once the cohort was fully staffed, though, the site quickly reached the target average monthly service hours.

The San Francisco Program Director expressed challenges identifying staff with the right mix of lived experience and capacity. She adapted by hiring people who seemed like a good fit for the program as she found, even if that meant waiting to fill their roster.

The team in Seattle found that children who were enrolled through the partnership with child welfare seemed to have “crises” on an almost weekly basis, more so than children enrolled through the school. This led the supervisor for the Friends working with children enrolled through child welfare to wonder if eight is too high a roster for this population. One mitigating facilitator that Seattle has adopted is the use of experienced Friends as Team Leads, who can be an extra level of support for Friends who are learning to handle their roster and balance required activities.

Data Collection and Reporting. Friends are required to record each interaction with youth, caregivers, and school staff in ETO, as well as data related to roadmap goals and core assets. Leaders across most sites described challenges in securing staff buy in for this administrative task. They shared that Friends were focused on the direct service component of their job and prioritized spending time with youth over reporting.

In Charlotte, the program director determined that Friends were under-reporting the number of hours that they spent with youth in ETO, thus under-representing the site’s average monthly hours. The program director believed this occurred because Friends deprioritized ETO reporting in favor of tasks they considered more important such as spending time with youth, being in schools, and interacting with caregivers. The program director repeatedly emphasized the importance of ETO to staff, culminating in an early January 2020 meeting that she identified as the turning point in securing staff buy-in. During the meeting, she spelled out that the site’s success – and future – depended on being able to demonstrate that the site was reaching its goals. Following the meeting, the site began exceeding target hours before dipping in March due to COVID-19.

Similarly, San Francisco’s leaders emphasized to staff that if an interaction or activity isn’t recorded in ETO, it’s like it didn’t happen. The site PD and Executive Director focused on staff scorecard reports (drawn from ETO performance data) as an accountability mechanism. The reports were reviewed at both individual and group supervision meetings.

Service Area. Some sites struggled to serve youth that were spread across a large area.



For Central Oregon, the long distances travelled to and from outings in a rural area meant Friends at this site spent more time travelling than Friends at other sites, leaving less time for meeting with children and administrative duties. Staff reported difficulty maintaining consistent hours with two youth whose families were experiencing homelessness. Similarly, San Francisco staff faced long commutes to visit youth spread across the service area. There and in other cities such as Boston and Seattle, gentrification is continually expanding sites' service areas by pushing low-income families further out of the city.

c) Facilitators to Achieving Targeted Hours with Youth

Relationships and Partnerships. The primary factor that facilitated sites meeting the targeted number of hours with youth revolved around the number and strength of relationships, particularly with schools and community partners. The importance of establishing relationships was underscored by challenges that arose without them.

The Austin site was able to meet the target average monthly hours early in the grant period and maintain that service level during the evaluation period. The site's strength in developing and growing relationships was a key component of their success. The Austin site used both top-down and bottom-up strategies to solidify community partnerships. Austin's executive director leveraged a strong local network, including an engaged Board of Directors, to continually connect to new community resources and opportunities. For instance, over the summer she leveraged these relationships to arrange multiple camps, an eight-week literacy program, and more group outings, enabling the site to maintain target monthly hours. Offering high quality, appealing summer activities during the first year of the program positioned them well to keep children and families engaged over subsequent summers. Austin's program director previously worked at nonprofits serving families in the same service area and had existing contacts and extensive knowledge of the resource landscape, as well as a deep understanding of the community served.

The Los Angeles site did not enroll youth through school selection, relying instead on a community-based organization referral process. Youth were selected into the program regardless of what school they attended (i.e., youth were not concentrated in one or several schools). Even so, the Los Angeles site pursued a MOU with the Los Angeles Unified School District, which it did not receive during the evaluation period. Without the MOU or relationships with individual schools that typically develop during selection, Friends did not have access to schools. They could only spend time with youth during a brief window in the later afternoon and early evening, making it challenging to reach target hours. Without MOUs in place, Friends were also sometimes prevented from conducting after-school pick-ups.

2. Contacts with Caregivers

The FOTC model prescribes that Friends contact caregivers of participating children at least two times per month. This target is designed to enable Friends to establish a presence in the child's home and become a trusted resource for the family. As shown in Exhibit 11, sites universally implemented this portion of the model with fidelity. Once sites had completed their initial service ramp-up, most averaged four to eight contacts per month. As with monthly service hours, there was a seasonal aspect, with the number of contacts dipping in the summer and in December. In March 2020, when COVID-19 prompted school closures and stay-at-home orders, sites' contacts with caregivers spiked.

Exhibit 12. Average Monthly Contacts with Caregivers Over Project Period, by Site

Site	Contacts per Month
<i>FOTC Contacts Target</i>	2.0
Austin	7.3
Los Angeles	6.3
Boston	6.2
San Francisco	5.2
Central Oregon	5.1
Seattle	4.6
Charlotte	3.9
<i>Source: ETO Data, July 2017-March 2020</i>	

a) Barriers to Achieving Caregiver Contacts

Seasonality. Newly established sites that conducted child selection in spring 2018 (Austin, Central Oregon, Charlotte) had their fewest contacts with caregivers in May and June 2018 but increased the number of contacts sharply over the first summer of operations. These sites were working against the clock to establish contact with caregivers and enroll their children in the program before school ended. School events (particularly end-of-year graduations and celebrations) and school pick-ups are valuable, built-in opportunities to connect with parents. Without the structure of school, it was more difficult to establish initial contact with caregivers over the summer.

It took time to engage caregivers as partners because caregivers were still learning about the format and expectations of the program. In San Francisco, caregivers asked why Friends were reaching out to them, since it's a program for the youth – not understanding the wrap-around nature of the program.

As with many low-income individuals, caregivers' contact information changed frequently. The living situations of participating youth was transient as well; for example, a youth may have been living with their parent one week and a grandparent the next. For the foster care enrolled cohorts in Seattle, Friends reported it was difficult to reach Washington Department of Children, Youth, and Families case managers for updates as well.

b) Facilitators to Achieving Caregiver Contacts

Caregiver rapport. Establishing rapport with caregivers was a crucial first step in enrolling youth into FOTC – a precursor to delivering services. All sites exhibited strong motivation to connect with caregivers who were both gatekeepers to youth and crucial partners in their development. They tried to connect with caregivers where they were—including homes, schools, and other places in the community. In these initial conversations with parents, program directors consistently emphasized the value that FOTC could provide to their child -- a professional, paid mentor for 12 years.

Events. At the end of the summer, all sites hosted back-to-school events where they supplied youth with backpacks and school supplies. Site program directors stated this was an opportunity to demonstrate the program's value.

In September 2018, Central Oregon held a fall event to connect with youth, families, and teachers, which they now hold annually. This popular event helped to jump start relationships with caregivers, leading to a spike in caregiver contacts the following months.



Demonstrating Value. Connecting with caregivers during the enrollment process provided Friends and staff with opportunities to assist caregivers with personal challenges. This in turn built their value proposition to caregivers and often led to increased contact. Acting as a resource for caregivers continued after enrollment too.

In Charlotte, caregivers from the first cohort regularly contacted the program director for support and referrals while turning to Friends for youth-related support and logistical information.

In Seattle, the positive word-of-mouth reputation of the program in the community seemed to make caregivers more receptive to working with the program.

Other sites established a deeper rapport over time. In Boston, staff reported it took time for caregivers to understand that FOTC was closely involved with but not part of the school, and that in fact they could be advocates for the family. It was particularly transformative when FOTC staff helped the caregiver advocate for their child at school (e.g., during a meeting about a 504 plan). Boston Friends also met with caregivers as part of setting roadmap goals for youth. The site found it to be another meaningful touchpoint and opportunity to collaborate on behalf of youth.

Maintaining Connection during COVID-19. The school closures, economic hardship, and other challenges caused by COVID-19 increased the frequency and intensity of contact between Friends and caregivers. Sites reported that caregivers were reaching out for support in educating their children at home. This included structuring the day, troubleshooting technology issues, and balancing work and care for other children. Sites also delivered food regularly and provided emergency cash assistance to participating families. Friends reported that some caregivers were sharing their own needs and challenges in a way that they had not prior to the pandemic.

3. Contacts with Schools

The FOTC program model prescribes that Friends establish themselves as a presence in the major spheres of a child’s life, including school. Further, some youth are behind academically when they enter FOTC and some are learning strategies to manage behaviors that can be disruptive at school. Site leaders and Friends both report that Friends devote significant time to working on literacy and numeracy skills and appropriate classroom behavior. Friends deliver this support during non-school outings but also during frequent classroom visits. Staff also observed that the large amount of time Friends spend in schools strengthens their relationships with teachers and other school staff, such as counselors and support staff.

As shown in Exhibit 12, sites implemented the school contact portion of the model with fidelity, nearly always exceeding the prescribed two contacts per month, with most sites ranging between four and twelve contacts per month. Contacts with schools are not tracked over the summer (June – August) as teachers and other school staff are unavailable.

Exhibit 13. Average Monthly Contacts with Schools Over Project Period, by Site

Site	Contacts per Month
<i>FOTC Contacts Target</i>	2.0
San Francisco	6.7
Central Oregon	6.6



Austin	5.7
Boston	4.6
Charlotte	4.4
Seattle	3.6
Los Angeles	2.9
<i>Source: ETO Data, July 2017-March 2020, omitting summer months June, July, and August</i>	

a) Barriers to School Contacts

Informal relationships and formal MOUs. Working without a formal MOU made school contacts harder, since it often precluded discussions with school staff during youth pick up.

As mentioned, the Los Angeles site was unable to establish an MOU with the Los Angeles Unified School I time burden associated with getting each school on board. Despite this, following an initial ramp-up period, Los Angeles was able to establish and maintain the prescribed two monthly contacts with schools. However, the site averaged the least number of monthly school contacts without this MOU in place with individual schools and/or school districts.

In San Francisco, newly hired Friends conducted child selection rather than program administrators. Because of this method, a new group of Friends had to build relationships with school staff each year, which presented a barrier to deepening the relationship between the site and the school.

b) Facilitators to School Contacts

Investment in establishing, maintaining, and growing partnerships with schools and staff. In general, sites initially approached and cultivated the school principal, who then introduced the FOTC program and staff to teachers and other school staff. The enthusiasm with which principals embraced the program impacted teacher buy-in, though in some cases teachers and other staff (e.g., reading specialists, counselors, behavioral specialists) independently embraced FOTC and became critical sources of support within the school.

Friends in Seattle, an established site, indicated that the positive reputation of the program in schools made it easier to introduce themselves when new teachers joined the staff. Site staff also said they benefited from institutional knowledge of how to work with their long-time school partners.

Across sites, executive directors were heavily involved in the initial relationship building with schools and continued to participate in child selection meetings. Day-to-day relationship management was conducted by Friends in schools multiple times per week. Interviews with Friends and administrators, as well as some direct observation, indicated that school staff generally valued and collaborated with Friends. FOTC staff noted that teachers were glad to have additional support in the classroom, to the extent that Friends sometimes had to clarify boundaries and remind teachers that they could not act as teacher's assistants. These early and maintained relationships were key to sites' success in reaching and largely exceeding the target level of contact with schools.

All sites invested heavily in building relationships with schools, but the Austin site went further, providing teachers with continuous support and appreciation in the form of classroom gift baskets,



donor-supplied, year-round snacks, and teacher gifts. This thoughtful cultivation of school relationships was apparent in their metrics, with Austin having one of the highest numbers of average monthly contacts with schools over the evaluation period

While these appreciation activities were important for building strong relationships in Austin, not all sites had the funding to provide these extra benefits. FOTC would benefit from doing further research on successful strategies for school partnership-building by talking with Central Oregon and San Francisco who had the highest number of school contacts.

Fidelity of Professional Development Capacity Building

The study's third research question explored sites' fidelity to the FOTC model in the areas of training and supervision. Sites reported and demonstrated adherence to the program's model for supervision, with Friends participating in regular one-on-one and group meetings with supervisors. Staff at all sites also participated in mandatory trainings provided by FOTC National, most notably the New Friends Training, which focused on the core philosophy and drivers of the FOTC model. In addition to these required trainings, sites had the latitude to develop and provide supplemental training.

a) Supervision

Across sites, site leadership were consistent in the amount of contact they had with Friends and the tools and structures they used for supervision. Through texting, social media apps, and calls, supervisors communicated with staff multiple times a day (one-on-one and/or group outreach). Friends at one site reported significantly less supervisor contact; they reported they had to tackle problems on their own. In general, Friends described supervisors as sources of referrals and advice and expressed a comfort discussing difficult issues with their supervisors.

In addition to being continuously available to Friends, supervisors held formal weekly or biweekly with individual Friends. Newly hired Friends were more likely to meet with their supervisors every week, which exceeds the expected fidelity standards of one 1:1 meeting bi-weekly. Furthermore, if a youth on their roster was experiencing a crisis, the Friend and supervisor would meet more often. The one-on-one meetings provided an opportunity to review the Friend's scorecard (standardized reports created by the FOTC National office) and address issues with any children in the Friend's roster.

All supervisors stated they used performance data in ETO to help guide supervision meetings. They reviewed and discussed individual and site-level scorecards during individual and group supervision meetings, respectively. This process allowed staff to identify and discuss areas of strength and challenges, informing the adjustment of activities and efforts. Supervisors also reported using ETO for day-to-day management such as verifying Friends' calendars, planning the content of outings, and strategizing about potential resource referrals. Some supervisors noted that printing a scorecard for all meetings was not possible due to time constraints, even if it was ideal.

Local sites also held group supervision meetings, either once a week or every other week. Supervisors printed site-level scorecards to assess trends and provide macro-level advice and support. For example, if the supervisor knew a Friend recently dealt with a situation that another Friend was facing, this would likely be addressed when talking about the group scorecard.



Some sites asked supervisors to maintain a roster of youth, either permanently or as a backup if the site was short of Friends. This approach had pluses and minuses. It kept supervisors connected to the day-to-day realities of serving as a Friend; however, it also squeezed out time for supervision and supporting the whole site. For example, one supervisor had to stop working on a pamphlet of housing resources to serve as a back-up Friend.

It was often difficult to find an individual fit for a supervisory role. Some sites noted that supervision and task lead responsibilities were good opportunities for Friends to grow. Other sites felt the skills required for a supervisor were distinct from that of a Friend and tended to look outside the organization to fill these roles. Supervisors relayed that experiences that prepared them for their current supervision tasks included direct service experience, working in communities that partnered with selection schools, a clinical background and/or familiarity with child development milestones, and prior mentoring experience.

b) Training

All staff hired as Friends participate in the New Friends Training delivered by FOTC National, as well as site-specific trainings. Depending on the when in the enrollment cycle a Friend is hired, he or she could attend the National training immediately after joining or several months later. Fidelity standards require New Friends Training within the first 90 days. Sites varied in when they preferred to send new staff to National training. Some stated the training offers a good orientation for brand-new hires, while others felt that staff benefited more from the training after several months on the job.

Friends reflected that the National office New Friend Training focused on the FOTC model and expressed a desire for experiential learning opportunities that would deepen their exposure and preparation. They described the value of shadowing more experienced Friends and recommended that the National office training include more practical guidance from established Friends. Friends specifically mentioned seeking guidance from experienced Friends on setting roadmap goals with young children, prioritizing activities (e.g., between data entry into ETO and providing services), and leveraging community resources to support youth and caregivers.

At the site level, common training topics include using ETO, cultural competency, and child development. These trainings were generally provided to enhance and augment topics covered in New Friends Training. Exhibit 13 presents a list of training topics covered by sites during the evaluation period. Note that multiple trainings were held to accommodate the cycle of enrolling youth and hiring Friends each year.

Exhibit 14. Training Topics Covered by Local Sites

	Austin	Boston	Central Oregon	Charlotte	Los Angeles	San Francisco	Seattle
Cultural Competency	X	X	X		X	X	
FOTC Office Procedures			X	X			
Efforts to Outcomes (ETO)	X	X	X	X	X	X	X
Literacy Training						X	
Self-Care Training			X	X			



Special Populations including Special Education and Foster Care		X			X		X
Supervision Training							X
Trauma Informed Care	X		X	X		X	X
Youth Development and Adverse Childhood Experiences (ACEs)	X		X		X	X	

Source: ICF Monthly Calls with SIF Local Site Staff

While sites typically paid external consultants for trainings, some chapters such as Austin and Boston were able to tap into their board and executive directors' connections to receive pro bono instruction. Larger sites that had more experienced Friends used a peer-to-peer training structure, where Friends were responsible for developing their own training topics.

Friends frequently stated that their role required them to be knowledgeable about many different topic areas, from child development to social work and family engagement to academics, to holistically serve youth. Recognizing the impossibility of developing expertise in all these fields, they were eager for training that would allow them to navigate these topics competently. Friends across sites were interested in additional professional development to dive deeper into the topics of trauma-informed care, de-escalation, behavior management, family engagement, and common diagnoses (e.g., ADD/ADHD, autism).

B. SUMMARY OF OUTCOME FINDINGS

The child welfare exploratory outcomes study examined change between the treatment (FOTC-enrolled) and comparison youth in terms of the length of stay in foster care, number of placements, time to permanence, and likelihood of achieving permanence. No additional removals occurred after FOTC enrollment for either group so removals, and subsequent re-entries into care, could not be assessed. *T*-Tests and chi-square analyses were used to assess differences between treatment and comparison youth. As shown in Exhibit 14, treatment youth ($M = 398.9, SD = 315.8$), compared to youth in the comparison group ($M = 576, SD = 314.9$) averaged fewer days in foster care ($t(61) = 2.2, p < .05, Cohen's d = 0.56$).

Exhibit 15. Length of Stay in Foster Care

	N	Mean	Std. Error	Std. Dev.	95% CI	
					Upper	Lower
Comparison	27	576.3	60.6	314.9	451.7	700.9
Treatment	36	398.9	52.6	315.8	292.0	505.7

* $p < .05$

Of the 27 youth in the comparison group, 15 (56%) achieved permanency. This was significantly higher than the proportion of treatment youth who achieved permanency (10 of 26, or 28%). This difference was statistically significant ($\chi^2(1, N = 63) = 4.98, p < .05$).

This evaluation sought and achieved a preliminary level of evidence. The primary impediments to achieving a higher evidence level were the sample sizes, the quality of the comparison group, and perhaps most importantly, the duration of the study. School administrative data and caregiver survey



findings in this study are aligned with findings from FOTC's RCT study, launched in 2007. The RCT research team collected and analyzed data from youth, caregivers, teachers, Friends, and schools until youth averaged 11 years old. FOTC learned that in the early years of the program school administrative data did not provide evidence of impact. Rather, caregiver survey findings revealed significant positive impact from the program on youth prosocial development and caregivers' perceptions of their child's behavior.

The work done to establish relationships with school districts and child welfare agencies should smooth the way for future data collection on the children involved in this study as they continue to participate in the program. A bit surprisingly, working with school districts proved more challenging than working with child welfare agencies. This is encouraging given FOTC's desire to advance its evidence-base as a best practice child welfare prevention program.

C. LESSONS LEARNED, STUDY LIMITATIONS, AND NEXT STEPS

During the implementation study, ICF learned a great deal about the facilitators and challenges to program expansion and replication that did not fall into the three main research questions related to program fidelity. These lessons learned are presented below.

1. Lessons Learned for FOTC Program Replication and Expansion

This section collects insights about ways to support new or expanding FOTC sites. Some insights are based on the SIF sites' successful adaptations, and some are in response to challenges observed at the SIF sites.

Site Staffing

- New sites might benefit from short-term, intensive support from the National office on hiring Friends. Most sites reported difficulty recruiting new Friends, and not having Friends on staff can be a barrier to achieving the target for monthly hours spent with youth.
- It can be beneficial to hire executive directors with experience in the nonprofit social services in the geographic area of the new site, to leverage their network, knowledge of the resource landscape, and understanding of the community served.

Staff Training and Support

- Traditional command supervision²⁸ may not work for Friends; instead, a coaching approach is needed.
- Supervisors should ensure Friends know can access scorecards and other important tracking information in ETO so they can review their performance in advance of meeting with their supervisor.
- Modeling professional norms is a large part of a supervisor's job, given the average age of staff and uniqueness of the position.
- Opportunities for further training for sites (new and existing) include trauma-informed care, family engagement, and working with families experiencing homelessness.

²⁸ Command supervision is when leaders define policies and procedures, and subordinates are expected to apply them correctly. It is untenable to think a policy manual could cover every single thing that could come up, so supervisors have to coach Friends to figure out how to handle unique situations according to the FOTC mission, values, and model.



- Providing increased responsibility for onboarding or training to established Friends is a great way of delegating a supervisor's workload. Plus, it's an opportunity for Friends to grow and potentially position themselves for future team lead positions. Established Friends that have been at the organization for a few years might play a greater role in creating trainings, scoping out trainings/community resources, reviewing onboarding checklists, and/or serving as a buddy that new Friends shadow before they have full rosters.
- Program directors appreciated the PD training hosted by the National office and mentioned that it might be helpful for supervisors to attend as well. Supervisors as well wished for a more-hands National training to bolster their capacity.

Data Collection (ETO)

- Set the expectation for thorough ETO reporting from the beginning of employment, even during the Friend hiring process. Hiring managers need to emphasize more the need to balance data entry and ground level work with youth, caregivers, and schools to ensure activities are consistently and accurately reported.
- Make explicit the connection between thorough ETO reporting, how the reports are a tool to improve Friend service to youth, the local site being able to report outcomes, the local site's sustainability, and Friends' employment security.
- Modeling consistent use of ETO is an important role for supervisors, because if they do not demonstrate a commitment to ETO, it is unlikely a Friend will. Supervisors and Friends practicing filling out ETO together can establish good habits.

Family Engagement

- Late summer back-to-school events, during which sites supply youth with backpacks and school supplies, can be a great opportunity to reconnect with youth and caregivers and demonstrate the program's tangible value.
- Involving caregivers in setting annual roadmap goals for a youth participant can be a meaningful opportunity to collaborate with the caregivers in a youth's life.
- Providing more opportunities for youth to participate in community-based activities during the summer, such as camps, literacy programs, and group outings, can help maintain the target monthly hours spent with youth.

Community Partnerships

- Engagement of the local school district needs to happen as soon as (or perhaps even before) a new site opens, because it is the districts that have the power to approve or deny research/study requests, including providing IRB approval and successfully executing data sharing agreements. Sites worked with the schools they were partnering with for child selection, but most had not established relationships with school districts. In sites without these relationships, ICF needed to explain both the FOTC program and our evaluation. The few sites that did have an MOU with their school district executed the agreement only months before ICF's involvement, and were reticent to jeopardize their nascent relationship with school district staff to discuss the evaluation.
- Even sites serving youth enrolled through child welfare system partners should establish formal relationships with the schools and school districts their participants attend to reach their target monthly contacts with schools and be able to visit classrooms regularly.



- Schools are an important partner for the FOTC model, whether youth are recruited via the schools or other sources.

ETO Capabilities Wishlist (from site staff)

- *Caregiver contacts:* There is a section in ETO for indirect contacts with caregivers. A drop-down option is “linkage to resources” but there isn’t a good way to input how that resource was found or what the outcome was. Three sites indicated that efforts might be lost in ETO because a contact that took minimal effort would appear the same way as a one took hours. Sites would like to make sure that a system that tells a story can demonstrate the full story of the work being done.
- *Track outing location:* The ability to easily see an outing location and activities in ETO would be helpful. To be able to sort that youth went to the park, for example, would be a helpful tracking tool.
- *Report source:* It would be helpful if on ETO reports there is more information on where the data is pulled from.
- *Quick access to reports:* Having a section in ETO database for the most utilized reports or the ability to customize a report dashboard more would be helpful.
- *Tracking direct/indirect:* One site would like an ETO setting where they can track direct and indirect time. This site said that it’s unclear how to record group activities and site events, but that can still be important in tracking development.
- *Excessive data entry:* Some sites indicated that Friends are asked to input too much information into ETO at times where it does not make sense. For example, it does not always make sense to track goals before those goals are selected. This is time consuming and frustrating when Friends can’t move on to the next page until everything is completed.
- *Practice mode/dummy entry:* One site indicated that a practice or research mode, where fake youth can be input into ETO, would be useful. This would give new staff a safe place to learn to use ETO without jeopardizing actual records.

2. Study Limitations

Study limitations are related to the sample size, the representativeness of the sample, and the strength of the comparison group. For example, across all seven sites, there were 431 youth whose data ICF wanted to include in the school administrative data study. Of these, ICF obtained data for 221 (51.3%). ICF obtained school administrative data for a low of 24% to a high of 97% of eligible youth at each of the six sites from whom we obtained data.

3. Lessons Learned working with Federal Initiatives

The major lesson learned by FOTC as a grantee of SIF is the importance of investment in both infrastructure capacity-building and implementation evaluation as an organization scales. With the support of adequate grant funding and robust technical assistance from federal staff, FOTC was able to do both activities simultaneously. By making these investments FOTC was able to reflect and adjust in real time on the most appropriate provision of resources to support new and existing sites as they expanded services. They were also afforded the opportunity to plan on what investment would be necessary to continue to scale with fidelity post-SIF. For example, the year two interim implementation evaluation report revealed that sites needed significantly more support with hiring, training and supporting staff. Based on these findings, FOTC hired additional HR staff at their National office and



secured new funding from the Department of Justice to invest in a network-wide quality improvement system. The QI system is strengthening the content of staff training and support as well as the delivery platform, through a new learning management system.

FOTC plans to use the findings from this study to strengthen fidelity of implementation and advance their evidence-base in the following ways: (1) partner with network sites in a review of program fidelity standards based on these findings, e.g. benchmarks for monthly hours of service, # of monthly contacts with caregivers and schools, use of ETO, and MOU requirements with schools; (2) strengthen their training and support of Friends and supervisors through continued development of their new QI system and more defined roles of national and local sites in professional development; (3) develop recommendations for sites about how, when, and whether to invest in administrative data use agreements with schools and child welfare agencies; and (4) strengthen the FOTC 2Gen logic model, program standards, and quasi-experimental or RCT evaluation plan to achieve a higher level of evidence.

STUDY LOGISTICS

A. PROTECTION OF HUMAN SUBJECTS

At the end of ICF's third year conducting the SIF evaluation, ICF submitted a report to FOTC National titled, *SIF Evaluation IRB Process Report September 23, 2019*. The report provided an overview of the IRB process, and presented key takeaways and lessons learned. ICF submitted ten IRB applications during this period (to seven school districts, two child welfare agencies, and ICF). Additionally, ICF submitted numerous IRB modifications and continuing review applications. Though each site's experience was unique, the underlying IRB structure, processes, and priorities were largely consistent across locations. Somewhat surprisingly, ICF encountered more challenges working with school districts than with child welfare agencies; thus, many of our recommendations focused on how best to work with school districts. The full report is provided in Appendix I.

B. BUDGET AND TIMELINE

The SEP timeline indicated ICF would secure IRB and MOU/DSAs mid-way through year two of the evaluation (by February 2018). This proved untenable given the number of approvals required, and the extensive revisions requested by the agencies.

C. EVALUATION AND PROGRAM STAFF INVOLVEMENT

There were no changes in ICF's evaluation team's personnel or roles that negatively affected the study's timeline, budget, or potential quality.

CONCLUDING REMARKS

The SIF evaluation included four components—an implementation study, a caregiver survey, a school study, and a child welfare study—each with its own methods, research questions, findings, and implications. Of the four research questions we answered for children enrolled through the child welfare agency, there was one significant positive finding in that treatment youth in FOTC, compared to youth in the comparison group, averaged fewer days in foster care. There was, however, one significant negative finding in that 56% of comparison youth and 28% of treatment youth achieved permanency. For the school study, using pooled data that maximized the quality of the constructed comparison group, no differences between the two groups were statistically significant.



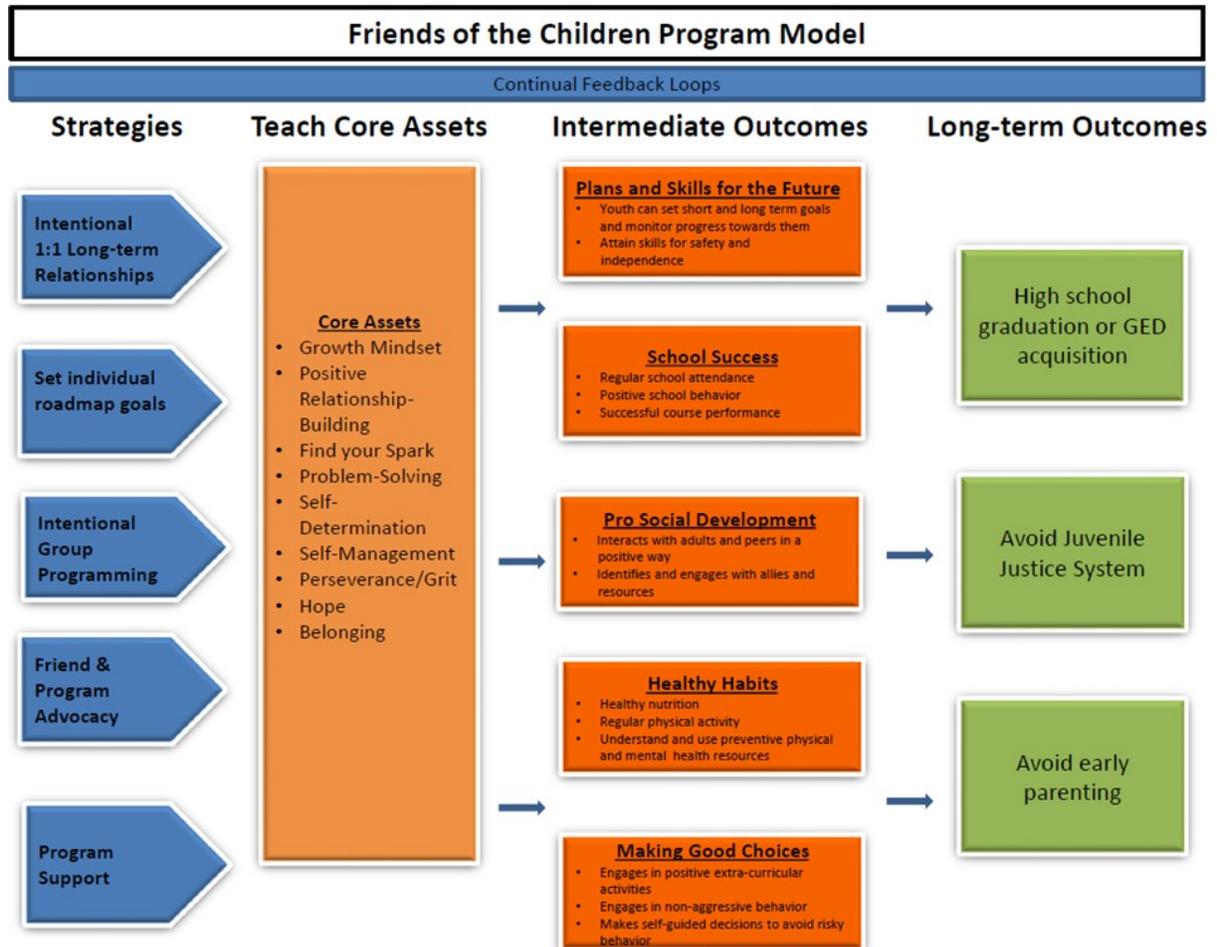
The mixed findings from the child welfare study, and the null findings from the school study, have several explanations noted previously in this report—namely the small sample sizes, the quality of the comparison group, and the duration of the study. Specifically, this evaluation’s findings using school administrative data are aligned with findings from FOTC’s RCT study, launched in 2007, which did not find programmatic impact in the early years of the program.

To assess these early years, it is useful to examine findings from the caregiver survey. As noted, comprehensive reports were provided to FOTC National in 2019 and 2020 under separate cover. We present some key takeaways from the 2020 survey here to demonstrate the program’s influence (as opposed to programmatic impacts). In 2020, caregivers rated their child in five socio-emotional learning areas (self-awareness, self-management, social awareness, relationship skills, and responsible decision making) *before* he/she had a Friend and in *past two months*. The significant difference between caregivers’ responses to the “before FOTC” and the “past two months” questions suggested that caregivers experienced positive changes in their child in all five dimensions of socio-emotional learning, even after controlling for caregiver’s level of education, employment status, involvement in the foster care system as a youth, time living with the child, whether the child lived in a single-parent household, and how long the child had been meeting with his or her Friend.

Other findings were that caregivers perceived that their child liked to go to school more, behaved well in school, made academic progress, and was more hopeful and optimistic. Caregivers reported that FOTC helped them strengthen family relationships, behavior, and parenting; understand how to improve learning in the home and relationships with teachers/school staff; work within systems like child welfare, schools, and hospitals; and connected them to food assistance, educational resources, cultural, art, and sports activities.

Whether FOTC replicates their model in new sites, or scales their model in existing sites, the early years after child enrollment are a time when administrative data (from schools or child welfare agencies) may not be useful to measure change. In these first years, FOTC can focus on avoiding the challenges and building on the facilitators to program fidelity identified during the implementation study, and surveying caregivers to learn about their experiences in the program, both to document how FOTC has helped their children and families and to learn how program services can be improved.

APPENDIX A: FRIENDS OF THE CHILDREN LOGIC MODEL





APPENDIX B: DATA ELEMENTS REQUESTED VERSUS RECEIVED BY SITE

	Austin	Central Oregon	Los Angeles	North Carolina	San Francisco	Seattle
Demographics						
Proxy ID	✓	✓	✓	✓	✓	✓
Current School	✓	✓	✓		✓	✓
Current Grade	✓		✓	✓	✓	✓
Foster	✓		✓		✓	✓
Age	Birth Year				Birth Year	✓
Gender	✓	✓	✓		✓	✓
Race/Ethnicity	✓	✓	✓		✓	✓
ELL Status	✓	✓	✓		✓	✓
Special Education Status	✓	✓	✓		✓	✓
504 Status	✓				✓	✓
Advanced Learning Status	✓		✓			✓
Primary Language					✓	✓
Home Language	✓		✓		✓	✓
Living With Status						✓
Projected Graduation Year						✓
On Track for High School Graduation						
Attendance History						
Attendance Date						
Percent Days Absent		✓				✓
Percent Days Unexcused						✓
Day Tardy Count			✓	✓		✓
Attendance Days Possible	✓	✓	✓	✓	✓	
Absent Days	✓		✓			✓
Attended Days		✓	✓	✓		✓
Unexcused and/or Excused Absence Days	✓			✓	✓	✓
Discipline History						
Incident Date	✓		✓			✓
Discipline Action	✓	✓	✓	✓		✓
Discipline Action Description	✓		✓			✓
Incident ID	✓					✓
Academic Progress						
Scores/Levels Achieved from Reading Assessments	✓	✓	✓	✓	✓	✓
Scores/Levels Achieved from Math Assessments		✓			✓	✓
Substitute Data Elements Received						
In-School Suspensions					✓	
Out-of-School Suspensions					✓	
Days of In-School Suspension				✓		
Days of Out-of-School Suspension				✓		



APPENDIX C: MEASURES OF PROGRAM FIDELITY BY MONTH AND SITE

Average Monthly Hours with Youth – Goal is 16 Hours							
	Austin	Boston	Central Oregon	Charlotte	Los Angeles	San Francisco	Seattle
Average	15.9	14.9	14.1	11.9	12.2	12.4	12.2
Jul-17						10.7	6.8
Aug-17						14.2	8.5
Sep-17						17.1	9.2
Oct-17						13.3	12.2
Nov-17						13.2	12.7
Dec-17						10.0	11.4
Jan-18		11.4				13.2	13.7
Feb-18		11.1				11.6	14.1
Mar-18						11.7	13.6
Apr-18						14.2	12.0
May-18	4.7	11.0	12.1			14.9	12.2
Jun-18	8.7	11.7	10.8	2.5		11.1	10.1
Jul-18	10.9	13.2	15.4	4.3		6.6	11.5
Aug-18	17.7	10.7	15.1	7.8		6.6	10.3
Sep-18	16.3	10.0	12.9	12.1	9.9	10.2	10.9
Oct-18	20.5	14.1	18.2	13.4	11.1	13.7	12.9
Nov-18	14.7	9.5	15.4	12.2	11.6	12.7	10.3
Dec-18	12.9	10.2	10.8	10.4	12.5	10.7	9.8
Jan-19	18.8	16.5	12.4	14.2	13.0	14.2	12.8
Feb-19	17.0	16.0	12.1	11.8	11.7	14.2	10.0
Mar-19	12.9	17.3	16.4	9.2	10.3	15.1	14.5
Apr-19	19.3	19.2	16.4	11.3	11.8	14.1	14.2
May-19	18.5	19.6	18.5	12.8	16.2	14.2	13.2
Jun-19	18.4	9.5	13.5	7.7	12.8	10.0	8.9
Jul-19	18.6	19.5	19.0	14.9	14.0	10.2	12.8
Aug-19	17.3	12.2	17.4	16.6	15.1	12.8	14.0
Sep-19	16.2	16.6	11.8	12.2	11.1	12.4	12.4
Oct-19	22.3	22.0	14.4	16.8	14.1	14.9	15.8
Nov-19	17.8	18.0	12.8	10.5	11.5	10.5	14.2
Dec-19	15.9	14.8	12.1	15.8	9.0	12.0	11.8
Jan-20	17.0	19.3	14.0	16.1	13.2	12.9	14.0
Feb-20	17.1	19.0	14.5	16.4	12.7	13.6	16.6
Mar-20	12.9	19.2	8.0	12.7	10.9	11.4	16.2

Source: ETO Data, July 2017 – March 2020



Average Monthly Contacts with Caregivers – Goal is 2 Contacts							
	Austin	Boston	Central Oregon	Charlotte	Los Angeles	San Francisco	Seattle
Average	7.3	6.2	5.1	3.9	6.3	5.2	4.6
Jul-17						7.2	2.8
Aug-17						8.6	3.5
Sep-17						7.4	3.3
Oct-17						5.2	3.9
Nov-17						6.2	3.6
Dec-17						4.1	3.9
Jan-18						5.7	4.6
Feb-18						4.7	5.3
Mar-18						5.1	4.7
Apr-18						4.1	4.3
May-18	1.9	4.9	2.1			5.3	5.0
Jun-18	3.8	5.2	3.4	1.5		5.4	3.8
Jul-18	6.4	5.5	3.1	1.9		3.7	3.6
Aug-18	7.7	5.1	4.3	3.8		3.8	3.5
Sep-18	6.0	4.6	4.2	5.5	5.8	4.0	4.1
Oct-18	6.4	5.1	6.4	5.1	4.8	5.0	4.1
Nov-18	7.1	4.1	4.5	4.1	4.6	4.7	3.7
Dec-18	5.8	3.5	3.4	5.3	5.1	4.2	3.7
Jan-19	8.6	4.5	3.8	4.6	7.5	5.7	5.0
Feb-19	7.7	5.1	3.8	3.5	5.7	5.2	4.6
Mar-19	6.3	6.4	6.0	3.4	5.1	5.7	5.5
Apr-19	6.8	7.5	6.2	4.8	7.1	4.4	5.7
May-19	7.1	7.5	6.6	4.4	8.5	5.7	4.8
Jun-19	9.2	3.5	6.0	2.1	6.5	4.7	2.8
Jul-19	8.2	7.1	7.8	3.9	7.7	5.0	4.7
Aug-19	7.6	5.6	7.1	3.5	8.7	5.9	5.1
Sep-19	8.2	7.9	5.3	3.3	6.0	5.2	5.1
Oct-19	10.2	9.4	6.0	3.4	7.5	5.1	6.6
Nov-19	8.0	9.2	6.3	2.3	5.4	4.3	5.1
Dec-19	7.2	7.4	6.0	4.4	4.6	3.8	5.6
Jan-20	7.8	8.3	5.8	4.3	6.1	4.9	5.8
Feb-20	7.2	7.4	5.4	3.8	6.1	4.6	6.4
Mar-20	12.1	8.2	4.7	6.9	7.0	5.8	8.1

Source: ETO Data, July 2017 – March 2020



Average Monthly Contacts with Schools – Goal is 2 Contacts							
	Austin	Boston	Central Oregon	Charlotte	Los Angeles	San Francisco	Seattle
Average	5.7	4.6	6.6	4.4	2.9	6.7	3.6
Jul-17						N/A	0.4
Aug-17						N/A	0.2
Sep-17						13.1	2.2
Oct-17						12.2	4.3
Nov-17						9.9	5.5
Dec-17						7.3	3.9
Jan-18						8.5	6.6
Feb-18						7.3	5.7
Mar-18						7.6	6.8
Apr-18						8.8	3.8
May-18	1.8	6.3	4.0			10.8	3.9
Jun-18	0.1	3.2	6.0	0.1		2.7	2.5
Jul-18	0.0	0.1	N/A	N/A		N/A	0.1
Aug-18	0.0	0.5	N/A	N/A		N/A	0.4
Sep-18	6.8	1.9	7.4	6.2	0.2	5.1	3.3
Oct-18	9.7	6.7	9.3	5.0	1.5	6.5	4.1
Nov-18	7.4	4.8	6.6	4.6	1.7	5.5	2.8
Dec-18	7.0	4.9	6.0	3.4	2.6	4.3	3.8
Jan-19	8.2	7.1	6.7	4.5	4.9	6.1	3.8
Feb-19	8.7	5.1	5.2	5.5	4.9	6.9	1.8
Mar-19	5.4	9.0	6.0	4.6	3.7	6.5	4.9
Apr-19	9.3	5.8	9.3	5.1	3.1	6.7	3.9
May-19	8.4	6.7	10.4	6.4	5.2	6.6	4.7
Jun-19	0.1	2.0	4.4	0.6	2.3	1.6	2.9
Jul-19	0.0	0.2	N/A	N/A	N/A	N/A	0.7
Aug-19	0.0	0.2	N/A	N/A	N/A	N/A	0.4
Sep-19	8.4	5.3	8.2	3.9	3.2	6.0	3.9
Oct-19	11.1	7.0	7.3	6.1	3.1	6.1	6.3
Nov-19	7.9	4.7	5.0	2.9	2.2	4.4	4.7
Dec-19	7.1	3.2	5.0	3.9	2.4	4.2	4.7
Jan-20	8.8	7.6	6.0	5.6	3.1	5.7	5.1
Feb-20	8.6	6.1	7.9	6.3	3.8	6.4	5.6
Mar-20	5.5	7.6	4.2	3.7	2.4	3.3	3.5

Note: N/A Denotes summer months

Source: ETO Data, July 2017 – March 2020



APPENDIX D: SELECTING YOUTH FROM CHILD WELFARE AGENCIES

1. The child welfare agency creates an initial pool of potential participants based on FOTC criteria. This is an initial, de-identified list of 5- and 6-year-old children who are in kindergarten, organized by race/ethnicity/gender who:
 - a. were removed from a home in the chapter's service area and are placed in a foster/relative placement that is in the chapter's service area;
 - b. have kith/kin connections within the service area;
 - c. came into care due to neglect plus at least one additional factor (e.g., parental drug and/or alcohol use, domestic violence, incarcerated parent);
 - d. score a 1 or 2 on the Child and Adolescent Needs and Strengths (CANS) assessment of behavioral and environmental risk/protective factors, or comparable tool depending on the child welfare jurisdiction;
 - e. are not being served in the highest level of residential treatment;
 - f. are not being co-managed by a developmental disability program and/or are not receiving Medicaid for developmental disability services; and
 - g. meet site-specific criteria (e.g., children of color).
2. FOTC site staff then meet with child welfare agency staff to discuss the child's placement plan, if known. Children that are likely to remain, or be placed, in the geographic area served by FOTC are selected to ensure that the Friend and child can have a continuing relationship.
3. The child welfare regional/district office:
 - a. signs program enrollment forms;
 - b. signs/sends a letter of introduction to supervisors/caseworkers and the child's school, endorsing the child's enrollment in the program;
 - c. gives FOTC staff permission to conduct outreach to biological and foster/relative caregivers for sign-on/program enrollment; and
 - d. signs consent forms for treatment and comparison group participants.
4. FOTC staff reach out to the child's foster/relative caregiver and biological family to introduce the program and obtain consent for treatment and comparison group participants.



APPENDIX E: BASELINE EQUIVALENCE FOR WA DCYF GROUPS

Variable	Treatment		Comparison		P-Value
	N	% or mean	N	% or mean	
Gender					
Male	16	44%	13	48%	0.770
Female	20	56%	14	52%	
Race/Ethnicity					
White, Non-Hispanic	11	31%	13	48%	0.155
Other	25	69%	14	52%	
Clinically Diagnosed with a Disability					
Yes	3	8%	3	11%	0.924
No	20	56%	15	56%	
Not Yet Determined	13	36%	9	33%	
Child Has Been Adopted in the Past					
No	33	92%	24	89%	0.710
Unable to Determine	3	8%	3	11%	
Manner of Removal from Home					
Voluntary	1	3%	2	7%	0.393
Court Ordered	35	97%	25	93%	
Circumstances of Removal from Home: Physical Abuse					
Conditions Does Not Apply	32	89%	22	81%	0.406
Condition Applies	4	11%	5	19%	
Circumstances of Removal from Home: Sexual Abuse					
Condition Does Not Apply	36	100%	27	100%	-
Circumstances of Removal from Home: Neglect					
Conditions Does Not Apply	11	31%	11	41%	0.401
Condition Applies	25	69%	16	59%	
Circumstances of Removal from Home: Parent's Alcohol Abuse					
Conditions Does Not Apply	32	89%	24	89%	1.000
Condition Applies	4	11%	3	11%	
Circumstances of Removal from Home: Parent's Drug Abuse					
Conditions Does Not Apply	20	56%	15	56%	1.000
Condition Applies	16	44%	12	44%	
Circumstances of Removal from Home: Child's Alcohol Abuse					
Conditions Does Not Apply	36	100%	26	96%	0.244
Condition Applies	0	0%	1	4%	
Circumstances of Removal from Home: Child's Drug Abuse					
Conditions Does Not Apply	36	100%	27	100%	-
Circumstances of Removal from Home: Child's Disability					
Conditions Does Not Apply	35	97%	27	100%	0.383
Condition Applies	1	3%	0	0%	
Circumstances of Removal from Home: Child's Behavior Problem					
Conditions Does Not Apply	35	97%	27	100%	0.383
Condition Applies	1	3%	0	0%	
Circumstances of Removal from Home: Death of Parent(s)					
Conditions Does Not Apply	35	97%	27	100%	0.383
Condition Applies	1	3%	0	0%	
Circumstances of Removal from Home: Incarceration of Parent(s)					



Variable	Treatment		Comparison		P-Value
	N	% or mean	N	% or mean	
Conditions Does Not Apply	33	92%	25	93%	0.893
Condition Applies	3	8%	2	7%	
Circumstances of Removal from Home: Caregiver's Illness/Disabling Condition Affects Their Ability to Care for Child					
Conditions Does Not Apply	30	83%	25	93%	0.275
Condition Applies	6	17%	2	7%	
Circumstances of Removal from Home: Abandonment					
Conditions Does Not Apply	34	94%	27	100%	0.213
Condition Applies	2	6%	0	0%	
Circumstances of Removal from Home: Inadequate Housing					
Conditions Does Not Apply	30	83%	25	93%	0.275
Condition Applies	6	17%	2	7%	
Current Foster Care Placement Setting					
Foster Family Home (Relative)	18	50%	10	37%	0.591
Foster Family Home (Non-Relative)	17	47%	16	59%	
Trial Home Visit	1	3%	1	4%	
Number of Removals from Home	36	1.36	27	1.15	0.126
Number of Aggregate Placements	36	3.50	27	3.33	0.836



APPENDIX F: SELECTING YOUTH FROM PARTNER SCHOOLS

1. Orient school staff
2. Inform caregivers/guardians that child selection will occur by sending a passive permission letter to caregivers in the mail using school's envelopes.
3. Conduct school five-six weeks of observations only on those children whose caregivers have consented
4. Teachers Complete Questionnaires on Children with Consent
5. Turn in questionnaires for data entry and creation of the *Initial List*
6. Hold the Selection Meeting with Teachers, other School Staff (Principal, and/or School Counselor), observing Friends, and the Program Director.
7. Together, the Program Director and Friends complete the eight questionnaires needed to create the final *Selected List* of children
8. Turn in questionnaires for data entry and creation of the *Selected List*
9. Give selected list to school to make initial phone calls to get family consent to contact and tell more about the Program.
10. Meet with the caregivers of selected children to describe program and enroll.



APPENDIX G: BASELINE EQUIVALENCE FOR POOLED SCHOOL GROUPS - ALL

	Comparison		Treatment		P-Value
	N	%	N	%	
Gender					
Male	107	54%	100	53%	0.912
Female	91	46%	87	47%	
Race/Ethnicity					
White, Non-Hispanic	19	10%	19	10%	0.853
Other	179	90%	168	90%	
School Title 1 Status					
Yes	198	100%	187	100%	-
Grade (School Year)					
Pre-School (18-19)	6	3%	3	2%	<0.0001
Kindergarten (16-17)	39	20%	34	18%	
Kindergarten (17-18)	63	32%	63	34%	
Kindergarten (18-19)	25	13%	59	32%	
1st grade (16-17)	13	7%	7	4%	
1st grade (17-18)	11	6%	11	6%	
1st grade (18-19)	40	20%	9	5%	
2nd grade (18-19)	1	1%	1	1%	
English Language Learner (ELL)					
Yes	36	18%	35	19%	0.892
No	162	82%	152	81%	
Special Education					
Yes	29	15%	28	15%	0.928
No	169	85%	159	85%	
Spring Reading Performance					
At/Above Benchmark	57	29%	51	27%	0.937
Below Benchmark	80	40%	76	41%	
Not Provided	61	31%	60	32%	
Absent Rate					
1st Quartile	16	8%	16	9%	0.998
2nd Quartile	26	13%	25	13%	
3rd Quartile	31	16%	29	16%	
4th Quartile	125	63%	117	63%	
Disciplinary Incidents					
Yes	6	3%	11	6%	<0.0001
No	192	97%	97	52%	
Not Provided	0	0%	79	42%	
Total	198		187		



APPENDIX H: BASELINE EQUIVALENCE FOR POOLED SCHOOL GROUPS TO MAXIMIZE QUALITY OF COMPARISON GROUP

	Comparison		Treatment		P-Value
	N	%	N	%	
Gender					
Male	39	49%	32	46%	0.772
Female	41	51%	37	54%	
Race/Ethnicity					
White, Non-Hispanic	9	11%	9	13%	0.738
Other	71	89%	60	87%	
School Title 1 Status					
Yes	80	100%	69	100%	-
Grade (School Year)					
Kindergarten (16-17)	10	13%	5	7%	0.825
Kindergarten (17-18)	33	41%	33	48%	
Kindergarten (18-19)	7	9%	7	10%	
1st grade (16-17)	12	15%	6	9%	
1st grade (17-18)	9	11%	9	13%	
1st grade (18-19)	8	10%	8	12%	
2nd grade (18-19)	1	1%	1	1%	
English Language Learner (ELL)					
Yes	15	19%	14	20%	0.813
No	65	81%	55	80%	
Special Education					
Yes	14	18%	13	19%	0.832
No	66	83%	56	81%	
Spring Reading Performance					
At/Above Benchmark	23	29%	18	26%	0.933
Below Benchmark	34	43%	30	43%	
Not Provided	23	29%	21	30%	
Absent Rate					
1st Quartile	6	8%	6	9%	0.992
2nd Quartile	11	14%	10	14%	
3rd Quartile	13	16%	11	16%	
4th Quartile	50	63%	42	61%	
Disciplinary Incidents					
Yes	3	4%	3	4%	0.853
No	77	96%	66	96%	
Total	80		69		