THE STEPHEN GROUP

A SYSTEMATIC REVIEW OF FAMILY CENTERED TREATMENT

Title IV-E Transitional Payment Assessment for the Nebraska Department of Health and Human Services

Abstract

This document provides a summary of the process used by the Stephen Group to review Family Centered Treatment (FCT) for the purpose of claiming Title IV-E evidence-based prevention services Transitional Payments. Using a systematic approach to the review of multiple studies, FCT was found to have at least two contrasts with non-overlapping samples in studies carried out in usual care or practice settings that achieve a rating of moderate or high on design and execution and demonstrate favorable effects in a target outcome domain. At least one of the contrasts demonstrated a sustained favorable effect of at least 12 months beyond the end of treatment on at least one target outcome. As a result, the intervention has been determined to be well-supported.

The Stephen Group, LLC
814 Elm Street
Suite 309
Manchester, NH 03101
(603) 419-9592
www.stephengroupinc.com

A Systematic Review of Family Centered Treatment

Title IV-E Transitional Payment Assessment for the Nebraska Department of Health and Human Services

Table of Contents

A Systematic Review of Family Centered Treatment

Title IV-E Transitional Payment Assessment for the Nebraska Department of Health and Human Services

Overview

Under contract to the State of Nebraska Department of Health and Human Services, The Stephen Group (TSG) has completed a review of the evidence base for Family Centered Treatment (FCT) in accordance with the standards articulated in the *Title IV-E Prevention Services Clearinghouse Handbook of Standards and Procedures* ("the Handbook"). In collaboration with our subcontracted partner, MEF Associates, a systematic review of three published studies was completed and we have arrived at a rating of **WELL-SUPPORTED** for FCT. According to the Handbook, a well-supported program:

> "Has at least two contrasts with non-overlapping samples in studies carried out in *usual care or practice settings* that achieve a rating of moderate or high on design and execution and demonstrate favorable effects in a target outcome domain. At least one of the contrasts must demonstrate a sustained favorable effect of at least 12 months *beyond the end of treatment* on at least one target outcome (p. 43)."

We find these standards to have been met. This memo summarizes our review of the three FCT studies examined, providing supporting evidence for the well-supported rating we have assigned. In the course of this review, there were several occasions where the Handbook's guidance was unclear or left room for interpretation. In these instances, we made decisions based on our understanding of the guidance and our best judgment; we have detailed our choices and justifications for them here.

Review Team and Conflict of Interest Statement

Review Team

This evaluation consisted of an experienced team of researchers and evaluation professionals from TSG and MEF Associates. TSG was pleased to partner with MEF Associates who provided expert technical guidance and assistance in the validation of study designs and statistical outcomes. Project staff included:

David DeStefano, MA – Senior TSG Consultant and Project Manager: Mr. DeStefano has more than 17 years of experience designing, implementing and conducting outcome evaluation for various federally-funded projects including a National Resource Center, Quality Improvement Center and numerous Administration on Children, Youth and Families funded demonstration grants and collaborative agreements. His experience includes quasi-experimental research design, focus group studies, survey research and data analysis. He earned a Master of Public Policy from New England College and a BA from Purdue University.

Kate Stepleton, PhD – MEF Associates Senior Research Associate: Ms. Stepleton has expertise in child and family research and policy, particularly in the areas of child welfare, maltreatment prevention, early childhood, and child well-being. She is skilled in qualitative and quantitative methods, design of experimental and quasi-experimental studies, survey research, and data analysis. She has managed research projects at Rutgers University School of Social Work, served in the Administration on Children, Youth and Families in the U.S. Department of Health and Human Services, and was an Associate with the Center for the Study of Social Policy in Washington, D.C. She has a Ph.D. in social work from Rutgers, an MSW from the University of Chicago's School of Social Service Administration, and a BA in Sociology from Barnard College.

Marissa Putnam, PhD – MEF Associates Research Associate: Ms. Putnam has conducted random assignment experimental research projects at Georgetown University, funded by the National Science Foundation and has also worked as a Research Assistant and Programmer at Mathematica Policy Research where she contributed to federal and state program implementation and evaluation, as well as measure development, in early childhood and health areas. Marisa earned her Master of Public Policy from the McCourt School of Public Policy at Georgetown University and her PhD in Developmental Psychology at Georgetown University.

Conflict of Interest Statement

TSG is committed to integrity and fairness in the conduct of all of its activities. As such, we certify neither TSG, our subcontracted partner, MEF Associates, or staff of either organization have a relationship with the developer of FCT or study authors through employment, consultancies, stock ownership, honoraria, or other relationship, either directly or through immediate family, which may be considered a conflict of interest. As such, the resulting opinion presented in this document is impartial and independent of external influence which may bias our determination.

Family Centered Treatment

Family Centered Treatment (FCT) is a behavioral intervention for youth who are in need of intensive services to prevent placement or to be reunified. The treatment model was developed by practitioners and has been refined by provider wisdom and experience over thirty years. FCT engages members of youths' family systems, targeting multiple dimensions of family functioning. Services are delivered at home or in the community over approximately six months. FCT is a listed treatment intervention on the National Child Trauma Stress Network website and is a SAMHSA trauma grant awardee. Additional information about FCT's treatment model can be found at www.familycenteredtreatment.org.

Program or Service Area(s)

Family Centered Treatment was reviewed in the area(s) of:

- In-Home Parenting Skills Based Program
- Mental Health

Handbook, Manual and Program Documentation

Program implementation materials including an online manual, implementation guide and other documentation were made available to reviewers in digital copies. The implementation manual, Wheels of Change © is accessible as a digital training manual through the e-learning platform *Mindflash*. Access to this platform was provided to the reviewers by the model developer for the purpose of verification. In addition to access to the online training materials, the model developer provided the following documents for review:

- Program Design and Implementation Guide
- Path of Implementation for Providers
- FCT Readiness Assessment Interview Plan
- Implementation Driver Assessment – closed copy
- Fidelity Adherence Compliance Tracker (FACT) – copy
- Fidelity Implementation-strategy Tool
- Readiness Assessment Report – Example Redact

- Readiness Assessment Matrix (RAM)
- Definitive Report on FCT – 1 of 6 required readings for FCT to achieve Certification

Program materials have been archived by The Stephen Group and are available for review, upon request.

Eligible Studies

Based on a comprehensive literature review of bibliographic databases and public websites maintained by state and local governments, three studies were identified and deemed eligible for review. Table 1 lists these three studies:

Table 1. Studies and Publications Reviewed

Indiana Waiver Substudy: • The Indiana University Evaluation Team & The Department of Child Services. (2018). Indiana Department of Child Services Child Welfare Title IV-E Waiver Demonstration Project Final Report. Indianapolis, IN: Indiana University School of Social Work and Indiana Department of Child Services.
Sullivan, et al. • Sullivan, M.B., Bennear, L.S., & Honess, K. (revised 2011). A quasi-experimental evaluation of Family Centered Treatment in the Maryland Department of Juvenile Services Community Based Non-residential Program: A report to Maryland Department of Juvenile Services and Institute for Family Centered Services. Great Falls, VA: FamiliFirst. • Sullivan, M. B., Bennear, L. S., Honess, K. F., Painter Jr, W. E., & Wood, T. J. (2012). Family Centered Treatment®--an alternative to residential placements for adjudicated youth: outcomes and cost-effectiveness. *Journal of Juvenile Justice*, 2(1), 25-40.
Bright, et al. • Bright, C. L., Betsinger, S., Farrell, J., Winters, A., Dutrow, D., Lee, B.R. & Afkinich, J. (2017). Youth Outcomes Following Family Centered Treatment In Maryland. Baltimore, MD: University of Maryland School of Social Work. • Bright, C.L., Farrell, J., Winters, A.M., Betsinger, S., & Lee, B. (2017). Family Centered Treatment, juvenile justice, and the Grand Challenge of Smart Decarceration. *Research on Social Work Practice, 28*(5), 638-645.

Copies of these studies have been maintained by The Stephen Group and are available for review, upon request.

Steps undertaken in the review of each study are described, below. Documentation detailing communication with study authors or the developer of FCT has been maintained by TSG and is available upon request.

Study Reviews

The Indiana Waiver Substudy employs a quasi-experimental design (QED) to evaluate the impact of FCT in a sample of child welfare-involved youth in Indiana.

- **Population:** The intervention group contained all youth who received FCT in Indiana from January 1, 2015 to December 31, 2015. The authors use propensity score matching to construct a comparison group of youth receiving child welfare services during the same period who did not receive FCT. As a result of the matching, the comparison group is similar on demographic and risk factors to the treatment group.
- **Data:** Administrative child welfare data

Study Design and Execution Rating

We have assigned the Indiana Waiver Study a design and execution rating of **MODERATE** for all contrasts.

- **Statistical Models:** The statistical modeling measures are, according to the information provided, appropriate for the analysis task (section 5.9.1). The authors' propensity scoring model includes appropriate covariates, and matching procedures eliminated any statistically significant differences in groups. The baseline equivalence standard was met for all contrasts, and no adjustment to the impact model was needed. Matching was done without replacement.
- **Measurement Standards:** All outcome and pre-test measures meet the Handbook's measurement standards (section 5.9.2). All have face validity. Because the measures are drawn from administrative child welfare data, all are assumed reliable and to have been consistently measured across intervention and comparison groups.
- **Design Confounds:** No design confounds were identified (section 5.9.3). Intervention and comparison groups were successfully matched on demographics and child welfare case characteristics. Post-match comparisons demonstrated nonsignificant differences between groups. It is possible that the groups differed on unobserved characteristics, but we are satisfied that the groups are comparable based on what the authors present. The intervention was delivered statewide, so we assume no n=1 person-provider confound exists.
- **Missing Data:** There does not appear to be any missing data.
- **Baseline Equivalence:** The child welfare outcomes assessed in the study do not have direct pre-tests. As such, we needed to identify a suitable pre-test alternative for each outcome. We selected safety ranking as a plausible pre-test, seeing it as a "common precursor" (Handbook, p. 30) to all outcomes examined. Safety risk itself is a multi-level categorical variable, so we selected a single level of the variable, high risk, to use as the pre-test. We believe this to be an appropriate selection because (a) it is likely to be associated with the study's outcomes, and (b) approximately half of the study population were assessed as high risk (whereas fewer than two percent of the population were assessed as low risk). The baseline equivalence standard (section 5.7) across intervention and treatment groups was met (Table 2) and no adjustment was needed in the impact model.

Table 2. Indiana Waiver Substudy: Baseline Equivalence

Contrasts: Outcome Measures	Pre-Test or Pre-Test Alternative	Intervention Group		Matched Comparison Group		Effect Size	Equivalence Standard Met
		n	Proportion	n	Proportion		
Remaining in-home throughout involvement with DCS	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
No repeat maltreatment during case	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
No repeat maltreatment within 6 months of case closure	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
No re-entry after case closure	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
Days of DCS involvement	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes

Days until reunification	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
Safety ranking: safe	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
Safety ranking: conditionally safe	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes
Safety ranking: unsafe	Safety ranking: Very High Risk"	187	.51	187	.52	-0.03	Yes

Impact Estimates

The study had two significant contrasts, both of which were favorable (Table 3).

Table 3. Indiana Waiver Substudy: Impact Estimates for Favorable Contrasts

Contrasts: Outcome Measure	Intervention Group			Matched Comparison Group			p value	Effect Size
	n	m	sd	n	m	sd		
Remaining in-home throughout involvement with DCS	187	.56	N/A	187	.39	N/A	.001	.41
Days until reunification	69	341	238.42	83	417	229.81	.02	-.32

Sullivan, et al.

The Sullivan, et al. study employs a quasi-experimental design QED to evaluate the impact of FCT in a sample of child welfare-involved youth in Maryland.

- **Population:** The intervention group contained youth who received FCT between July 1, 2003 and December 31, 2007 in Maryland. A propensity score-matched comparison group was constructed from youth who were discharged from group homes, therapeutic group homes, and committed residential placements during the same time.
- **Data:** All data are drawn from administrative records from the state Department of Juvenile Services.

Study Design and Execution Rating

We have assigned the Sullivan, et al. study a design and execution rating of **MODERATE** for some but not all contrasts. While the authors present findings both one and two years post-treatment, the Handbook requires only one contrast with a sustained favorable effect for at least twelve months for an intervention to receive a rating of well-supported. Having identified favorable 12-month effects, we did not review year-two findings.

- **Statistical Models:** The statistical modeling measures are, according to the information provided, appropriate for the analysis task. The authors' propensity scoring model includes appropriate covariates, and matching procedures eliminated any statistically significant differences in groups. It is possible that the groups differed on unobserved characteristics, but we are satisfied that the groups are comparable based on what the authors present. However, as demonstrated below, the baseline equivalence standard was not met for several of the study's contrasts, and the authors do not appear to control for these post-matching group differences in their impact models. Therefore, in presenting impact estimates, we have only shown those statistically significant contrasts for which the baseline equivalence standard was met.

- **Measurement Standards:** All outcome and pre-test measures meet the Handbook's measurement standards. All have face validity. Because the measures are drawn from administrative juvenile justice data, all are assumed reliable and to have been consistently measured across intervention and comparison groups.
- **Design Confounds:** We have not identified any design confounds. The authors describe how selection into FCT takes place: judges make decisions informed by a structured assessment tool and the recommendations of case managers and probation officers. While this potentially introduces selection bias, the authors make two arguments about how they address this: first, they include an approximation of the measures from the structured assessment tool in the propensity score model. Assessments were not available for all youth in the sample, so the authors identified proxies for the measures drawn from pre-treatment juvenile justice data for youth. Sufficient detail is given to demonstrate that these proxy measures are suitable alternatives to the assessment's indicators. Second, the authors specify the matching model such that region is fixed. Intervention group youth may only be matched with comparison group youth in their region in an effort to hold constant the effect of geographic variation in how the child welfare system operates. We find these measures to be adequate for controlling for potential selection bias. Youth who are assigned to FCT are considered "at imminent risk for out of home placement (Sullivan, Bonnear, & Honess, p. 4)," suggesting that youth who did not receive FCT would otherwise have been placed in group or residential placements. As such, we find that the comparison group is conceptually suitable to the study. There is no n=1 person-provider confound.
- **Missing Data.** The authors appear to have complete data on all baseline and outcome variables. We did note that there is a slight discrepancy in the size of the matched comparison group used to calculate pre-treatment characteristics (n=1,785) and the size of the matched comparison group used to estimate treatment effects at one year (n=1,788). The authors note that they omitted some cases from the descriptive analysis because "they skewed the means of the matched groups on important characteristics" (p. 13). They go on to explain that this is "an artifact of using 4 matches for each treatment observation, with replacement, and an aggregation of matching characteristics via the propensity score" (p. 13). The skewed means in the matched comparison group were observed for measures relating to youth placements in secure confinement and special placements; these measures were not skewed in the unmatched comparison sample. When presenting descriptive statistics for the matched comparison group, the authors note that the observations responsible for skewing the noted means were dropped and assure the reader that differences between the full matched comparison sample (n=1,788) and the slightly smaller group (n=1,785) on other measures are "miniscule." We did not use measures related to secure confinement or special placements to establish baseline equivalence for any of the contrasts, so we are satisfied that this difference in reported comparison sample sizes does not threaten the validity of the study.
- **Baseline Equivalence:** We were able to find direct pre-tests for many of the outcomes examined in the study; for others, we identified pre-test alternatives that were conceptually similar or could be plausibly considered precursors to the outcomes in question. The baseline equivalence standard was met for some but not all contrasts (Table 4). Where baseline equivalence was not established, the effect sizes fell into the range requiring the researchers to adjust for the pre-tests or pre-test alternatives in the impact model; however, as noted below, the impact model did not appear to include any adjustment.

Table 24. Sullivan et. al: Baseline Equivalence

Contrasts: Outcome Measures (one year post-treatment)	Pre-Test or Pre-Test Alternative	Intervention Group			Matched Comparison Group			Effect Size	Equivalence Standard Met
		n	m	sd	n	m	sd		
Proportion of youth with residential placements	Proportion of youth with placements: community based residential	446	0.17	0.38	1785	0.17	0.37	0.00	Yes
Frequency residential placements	Placement frequency: community based residential	446	0.23	0.56	1785	0.21	0.51	0.04	Yes
Duration residential placements	Placement duration: community based residential	446	37.16	131.30	1785	25.16	83.93	0.13	Adjustment needed
Conditional duration residential placements	Placement duration: community based residential	446	37.16	131.30	1785	25.16	83.93	0.13	Adjustment needed
Proportion of youth with pending placements	Proportion of youth with placements: community based residential	446	0.17	0.38	1785	0.17	0.37	0.00	Yes
Frequency pending placements	Placement frequency: community based residential	446	0.23	0.56	1785	0.21	0.51	0.04	Yes
Duration pending placements	Placement duration: community based residential	446	37.16	131.30	1785	25.16	83.93	0.13	Adjustment needed
Conditional duration pending placements	Placement duration: community based residential	446	37.16	131.30	1785	25.16	83.93	0.13	Adjustment needed
Proportion of youth with community detention	Proportion of youth with placements: community detention	446	0.62	0.49	1785	0.65	0.48	-0.06	Adjustment needed
Frequency of community detentions	Placement frequency: community detention	446	1.11	1.19	1785	1.23	1.23	-0.10	Adjustment needed

Duration of community detentions	Placement duration: community detention	446	45.09	53.92	1785	48.46	59.58	-0.06	Adjustment needed
Conditional duration community detentions	Placement duration: community detention	446	45.09	53.92	1785	48.46	59.58	-0.06	Adjustment needed
Proportion of youth with secure detentions	Proportion of youth with placement: secure detention	446	0.63	0.48	1785	0.65	0.48	-0.04	Yes
Frequency of secure detentions	Placement frequency: secure detention	446	1.23	1.36	1785	1.23	1.29	0.00	Yes
Duration of secure detentions	Placement duration: secure detention	446	22.82	32.03	1785	20.27	28.21	0.09	Adjustment needed
Conditional duration secure detentions	Placement duration: secure detention	446	22.82	32.03	1785	20.27	28.21	0.09	Adjustment needed
Proportion of youth offending	Proportion of youth with offenses: category 1	446	0.28	0.45	1785	0.26	0.44	0.05	Yes
Frequency of offenses	Frequency of offenses: All categories	446	8.19	6.30	1785	7.96	6.34	0.04	Yes
Proportion of offending in category 1 and 2	Proportion of youth with offenses: category 1	446	0.28	0.45	1785	0.26	0.44	0.05	Yes
Frequency of category 1 and 2 offenses	Frequency of offenses: category 1	446	0.43	0.89	1785	0.37	0.76	0.08	Adjustment needed
Proportion of youth with adjudications	Proportion of youth with adjudications: all categories	446	1.70	1.76	1785	0.12	0.45	0.04	Yes
Frequency of adjudications	Frequency of adjudicated offenses: all categories	446	2.70	2.42	1785	2.67	2.30	0.01	Yes
Proportion adjudications category 1 and 2	Proportion of youth with adjudications: category 1	446	0.11	0.31	1785	0.09	0.28	0.07	Adjustment needed
Frequency of category 1 and 2 adjudications	Frequency of adjudications: category 1	446	0.14	0.47	1785	0.12	0.45	0.04	Yes

Impact Estimates

The impact model for all contrasts did not include any adjustment for variables that did not meet the baseline equivalence standard. For contrasts that met the baseline equivalence standard, we calculated *p* values according to the procedures described in Appendix A. However, the authors

conducted propensity score matching with replacement, meaning that once a youth in the comparison group was matched, they were returned to the sample and could be matched with additional treatment youth. As such, youth in the comparison group may be counted multiple times. The resultant downward biasing of standard errors yields potentially inflated *p* values in traditional *t*-tests unless a statistical correction is applied. The authors have corrected for this duplication in comparison observations in their reported findings; however, short of replicating their analyses with their raw data, we cannot do the same. The *p* values reported in Tables 5 and 6, which were calculated based on traditional *t*-tests, may therefore be artificially low. However, our review and the authors' findings agree that the contrasts listed in these tables are significant.

Among those contrasts for which pre-tests and pre-test alternatives met the baseline equivalence standard, four were significant.[1] Two significant contrasts were favorable (Table 5) and two were unfavorable (Table 6) when considered in isolation. Importantly, however, the unfavorable contrasts, when considered in the context of the other significant findings, tell a story about FCT's overall positive impact. This distinction is explained in greater detail below. All contrasts pertain to outcomes measured at 12 months post-treatment.

Table 5. Sullivan, et al.: Impact Estimates for Favorable Contrasts

Contrasts: Outcome Measure	Intervention Group			Matched Comparison Group			p value	Effect Size
	n	m	sd	n	m	sd		
Proportion of youth with residential placements at 12 months post-treatment	446	.38	0.49	1788	0.50	0.50	<.001	-.30
Frequency of residential placements at 12 months post-treatment	446	0.50	0.74	1788	0.63	0.70	<.001	-.18

Table 6. Sullivan, et al.: Impact Estimates for Unfavorable Contrasts

Contrasts: Outcome Measure	Intervention Group			Matched Comparison Group			p value	Effect Size
	n	m	sd	n	m	sd		
Proportion of youth with adjudications at 12 months post-treatment	446	.32	0.47	1788	.23	.44	<.001	.28
Frequency of adjudications at 12 months post-treatment	446	0.70	1.52	1788	0.45	1.04	<.001	.18

Results indicate that, compared to youth in the comparison group, more youth who received FCT had adjudications in the year following treatment. Youth in the treatment group also had a higher frequency of adjudications in the post-treatment year. However, considering these findings alongside the study's other results, the authors argue that higher adjudication rates among FCT, at minimum,

[1] One additional case requires special explanation. In our review, the contrast for the proportion of youth with secure detentions, for which the baseline equivalence standard was met, was significant at p<.001. However, in the report, the contrast does not meet even a minimum threshold for statistical significance. While our p value may be inflated due to duplication in the comparison sample, we suspect there may be a typo in the report and the article, as the average effect of treatment on the treated (SATT) reported does not seem appropriate given the treatment and comparison means presented. We have asked the authors to verify the values in the publications, but as of November 25, 2019, we have not received a response. As the Clearinghouse only requires one sustained favorable outcome for an intervention to be rated as well-supported, and as our review and the authors' analysis agree on the significance of two sustained favorable outcomes, we did not feel it necessary to pursue the issue further.

do not suggest unfavorable effects of FCT. Rather, they might be evidence of the program's capacity to change family system values. The authors write:

> [Higher adjudication rates] must be reflective of court decisions as applied to youth receiving FCT. This outcome may be interpreted as a manifestation of the emphasis on accountability in Family Centered Treatment; the model attempts to instill accountability by accepting responsibility for one's actions as a family system value. This may be exhibited in the family's interactions with the courts as an increase in the likelihood of an offense being adjudicated. Overall, however, the fact that residential placements and days in detention are significantly lower suggests that the average youth receiving FCT committed fewer offenses of a nature that would warrant a consideration of removal from the community (Sullivan, Bonnear, & Honess, pp. 12-13).

We agree with the authors that the adjudication findings must be considered in the context of the study as a whole and that they do not undermine the evidence for the program's effectiveness.

Bright, et al.

Similar to Sullivan, et al., the Bright, et al. study employs a quasi-experimental design to evaluate the impact of FCT in a sample of child welfare-involved youth in Maryland. The Bright et al. study covers a later, non-overlapping time period.

- **Population:** The study population consisted of youth who had been adjudicated delinquent in Maryland. The intervention group consisted of 1,246 youth who received FCT, initiating treatment between July 1, 2008 and June 30, 2013. The comparison group was drawn from the population of youth who were served in group homes or treatment group homes during the same period.
- **Data:** Data are drawn from the administrative records of the service provider and the state Department of Juvenile Services.

Study Design and Execution Rating

We have assigned the Bright, et al. study a design and execution rating of **LOW** for all contrasts.

- **Statistical Models:** The authors' propensity scoring model includes appropriate covariates, and matching resulted in nonsignificant differences between intervention and comparison groups. However, as demonstrated below, even after matching, none of the pre-test alternatives met the Handbook's standard for baseline equivalence; all fell into the adjustment range. In these instances, the Handbook requires that the impact model control for the group differences in the pre-test alternatives. However, the study's impact models do not include any such controls.
- **Measurement Standards:** All outcome and pre-test measures meet the Handbook's measurement standards. All have face validity. Because the measures are drawn from administrative juvenile justice data, all are assumed reliable and to have been consistently measured across intervention and comparison groups.
- **Design Confounds:** We have not identified any design confounds.
- **Missing Data:** There does not appear to be any missing data.
- **Baseline Equivalence:** The juvenile justice outcomes identified in this study do not have direct pre-tests. As such, we have identified pre-test alternatives, that, according to our judgment, are conceptually similar or could be plausibly considered precursors to the study's outcomes. None of the pre-test alternatives we identified met the baseline equivalence

standard (Table 7). Where baseline equivalence was not established, the effect sizes fell into the range requiring the researchers to adjust for the pre-tests or pre-test alternatives in the impact model; however, as noted below, the impact model did not appear to include any adjustment.

Table 7. Bright, et al: Baseline Equivalence

Contrasts: Outcome Measures	Pre-Test or Pre-Test Alternative	Intervention Group			Matched Comparison Group			Effect Size	Equivalence Standard Met
		n	m	sd	n	m	sd		
Re-adjudication	Number of prior delinquency complaints	1246	5.29	3.80	693	5.73	4.00	-.11	Adjustment needed
Commitment	Any prior committed placement	1246	.124	N/A	693	.144	N/A	-.10	Adjustment needed
Conviction	Any prior adjudication for a violent offense	1246	.167	N/A	693	.190	N/A	-.10	Adjustment needed
Incarceration	Any prior committed placement	1246	.124	N/A	693	.144	N/A	-.10	Adjustment needed

Summary

Based on our thorough review of the Indiana Waiver Substudy, Sullivan et al., and Bright et al., we find that FCT meets the Title IV-E Prevention Services Clearinghouse's standards for a rating of **WELL-SUPPORTED**. As all of the contrasts in the Bright et al. study were rated as low, we draw exclusively on the Indiana Waiver Substudy and Sullivan et al. in making this determination. These studies examine two non-overlapping samples in usual care or practice settings. Each had contrasts that were rated as moderate and were statistically significant. In addition to the favorable outcomes reported in the Indiana Substudy, the favorable outcomes presented in Sullivan et al. were sustained for at least 12 months after treatment.

According to these two studies, it appears FCT decreases out-of-home placement for youth. Compared to youth who did not receive FCT, those who did were less likely to be in residential placements during their involvement with child welfare (Indiana Waiver Substudy) and in the year after FCT ended (Sullivan et al.). Those who were in residential placements had fewer residential placements in the year after FCT ended (Sullivan et al.). Youth in out-of-home placement who received FCT also had shorter time to reunification than those who did not receive FCT (Indiana Waiver Substudy)

Appendix A: Calculation Methods

The Title IV-E Prevention Services Clearinghouse standards are substantially based on the standards created for the What Works Clearinghouse (WWC). We therefore used the formulas found in the WWC procedures handbook for calculating effect sizes and *p* values.

Calculating Effect Sizes: The Handbook specifies preferred statistics for effect sizes: Hedges' *g* for continuous outcomes and the Cox index for dichotomous outcomes. These are also the preferred effect size statistics for the WWC; formulas are presented on pages 13 and 14 of the WWC procedures handbook.

Calculating *p* values: The Handbook instructs reviewers to calculate *p* values for contrasts for which the baseline equivalence standard has been met. For continuous outcomes, we conducted *t*-tests using the means, standard deviations, and sample sizes reported. We also conducted *t*-tests for binary outcomes presented in the Sullivan, et al. study, as standard deviations were provided. To calculate the *p* value for dichotomous outcomes when standard deviations were not provided, as in the Indiana Waiver Substudy, we used the formula found on page 16 of the WWC Procedures Handbook.

Appendix B: Completed PI-19-06 Attachment

Attachment B: Checklist for Program or Service Designation for HHS Consideration

Instructions:

Section I: The state must complete Section I (Table 1) once to summarize all of the programs and services that the state reviewed and submitted and the designations for HHS consideration.

Section II: The state must complete Section II (Tables 2 and 3) once to describe the independent systematic review methodology used to determine a program or service (listed in Table 1) designation for HHS consideration. Section II outlines the criteria for an independent systematic review. To demonstrate that the state conducted an independent systematic review consistent with sections 471(e)(4)(C)(iii)(I), (iv)(I)(aa) and (v)(I)(aa) of the Act, the state must answer each question in the affirmative. If the independent systematic review used the Prevention Services Clearinghouse Handbook of Standards and Procedures, the relevant sections must be indicated in the "Handbook Section" column. If other systematic standards and procedures were used, states must submit documentation of the standards and procedures used to review programs and services. States should determine the standards and procedures to be used prior to beginning the independent systematic review process. If the state cannot answer each question in Table 2 and Table 3 in the affirmative, ACF will not make transition payments for the program or service reviewed by the state using those standards and procedures.

Section III: The state must complete Section III (Tables 4 and 5) for each program or service listed in Table 1, and provide all required documentation. Section III outlines the requirements for the review of the program or service. States should complete Table 4 prior to conducting an independent systematic review to determine if a program or service is eligible for review. For a program or service to be eligible for review, the answer to both questions in Table 4 must be affirmative and the state must provide the required documentation. If a program or service is eligible for review, the state must conduct the review and identify each study reviewed in Table 5, regardless of whether a study was determined to be eligible to be included in the review.

Section IV: The state must complete Section IV (Tables 6-10) for each program or service (listed in Table 1) reviewed and submitted and provide all required documentation. Section IV lists studies the state determined to be "well-designed" and "well-executed" and outlines characteristics of those studies. Do <u>not</u> include eligible studies that were <u>not</u> determined to be "well-designed" and "well-executed" in Tables 6 -10. States should complete Table 6 with a list of all eligible studies determined to be "well-designed" and "well-executed." States should complete Table 7 to describe the design and execution of each eligible "well-designed" and "well-executed" study. States should complete Table 8 to describe the practice setting and study sample. States must answer in the affirmative that the program or service included in each study was not substantially modified or adapted from the version under review. States must detail favorable effects on target outcomes present in eligible studies determined to be "well-designed" and "well-executed." States must detail unfavorable effects on target and non-target outcomes present in eligible studies determined to be "well-designed" and "well-executed."

Section V: The state must complete Section V (Table 11) for each program or service reviewed and submitted. Section V lists the program or service designation for HHS consideration and verification questions relevant to that designation. The state must answer the questions applicable to the relevant designation in the affirmative.

Section I: Summary of Programs and Services Reviewed and their Designations for HHS Consideration

Section I. Summary of Programs and Services Reviewed

Table 1. Summary of Programs and Services Reviewed

To be considered for transitional payments, list programs and services reviewed and provide designations for HHS consideration.

Program or Service Name *(if there are multiple versions, specify the specific version reviewed)*	**Proposed Designations for HHS consideration** *(Promising, Supported, or Well-Supported)*
Family Centered Treatment	Well-Supported

Section II: Standards and Procedures for an Independent Systematic Review

Section II. Standards and Procedures for a Systematic Review

(*Complete Table 2 and Table 3 to provide the requested information on the independent systematic review. The same standards and procedures should be used to review all programs and services.)*

Table 2. Systematic Review

Sections 471(e)(4)(C)(iii)(I), (iv)(I)(aa) and (v)(I)(aa) of the Act require that systematic standards and procedures must be used for all phases of the review process. In the table below, verify that systematic (i.e., explicit and reproducible) standards and procedures were used and submit documentation of reviewer qualifications. If the systematic review used the Prevention Services Clearinghouse Handbook of Standards and Procedures, indicate the relevant sections in the "Handbook Section" column. If other systematic standards and procedures were used, submit documentation of the standards and procedures.

Table 2. Systematic Review	☑ to Verify	Handbook Section
Were the same systematic standards and procedures used to review all programs and services?	✓	--
Were qualified reviewers trained on systematic standards and procedures used to review all programs and services?	✓	--
Were standards and procedures in accordance with section 471(e) of the Social Security Act?	✓	--
Were standards and procedures in accordance with the Initial Practice Criteria published in Attachment C of ACYF-CB-PI-18-09?	✓	--
Program or Service Eligibility: Were systematic standards and procedures used to determine if programs or services were eligible for review? At a minimum, this includes standards and procedures to:	✓	
• Determine if a program or service is a mental health, substance abuse, in-home parent-skill based, or kinship navigator program; and	✓	2.1
• Determine if there was a book/manual or writing available that specifies the components of the practice protocol and describes how to administer the practice.	✓	2.1.2
Literature Review: Were systematic standards and procedures used to conduct a comprehensive literature review for studies of programs and services under review? At a minimum, this includes standards and procedures to:	✓	
• Search bibliographic databases; and Search other sources of publicly available	✓	3.2
• Studies (e.g., websites of federal, state, and local governments, foundations, or other organizations).	✓	3.2
Study Eligibility: Were systematic standards and procedures used to determine if studies found through the comprehensive literature review were eligible for review? At a minimum, this includes standards and procedures to:	✓	
• Determine if each study examined the program or service under review (as described in the book/manual or writing) or if it examined an adaptation;	✓	4.1
• Determine if each study was published or prepared in or after 1990;	✓	4.1.1
• Determine if each study was publicly available in English;	✓	4.1.2, 4.1.3
• Determine if each study had an eligible design (i.e., randomized control trial or quasi-experimental design);	✓	4.1.4
• Determine if each study had an intervention *and* appropriate comparison condition;	✓	4.1.4
• Determine if each study examined impacts of program or service on at least one 'target' outcome that falls broadly under the domains of child safety, child permanency, child well-being, or adult (parent or kin-caregiver) well-being. Target outcomes for kinship navigator programs can instead or also include access to, referral to, and satisfaction with services; and	✓	4.1.5

Table 2. Systematic Review	☑ to Verify	Handbook Section
• Identify studies that meet the above criteria and are eligible for review.	✓	4.1
Study Design and Execution: Were systematic standards and procedures used to determine if eligible studies were well-designed and well-executed? At a minimum, this includes standards and procedures to:	✓	
• Assess overall and differential sample attrition;	N/A	
• Assess the equivalence of intervention and comparison groups at baseline and whether the study statistically controlled for baseline differences;	✓	5.7
• Assess whether the study has design confounds;	✓	5.9.3
• Assess, if applicable, whether the study accounted for clustering (e.g., assessed risk of joiner bias[1]);	N/A	
• Assess whether the study accounted for missing data; and	✓	5.9.4
• Determine if studies meet the above criteria and can be designated as well-designed and well-executed.	✓	5.1 – 5.9
Defining Studies: Sometimes study results are reported in more than one document, or a single document reports results from multiple studies. Were systematic standards and procedures used to determine if eligible, well-designed and well-executed studies of a program and service have non-overlapping samples?	✓	
Study Effects: Were systematic standards and procedures used to examine favorable and unfavorable effects in eligible, well-designed and well-executed studies? At a minimum, this includes standards and procedures to:	✓	
• Determine if eligible, well-designed and well-executed studies found a favorable effect (using conventional standards of statistical significance) on each target outcome; and	✓	5.10
• Determine if eligible, well-designed and well-executed studies found an unfavorable effect (using conventional standards of statistical significance) on each target or non-target outcome.	✓	5.10
Beyond the End of Treatment: Were systematic standards and procedures used to determine the length of sustained favorable effects beyond the end of treatment in eligible, well-defined and well-executed studies? At a minimum, this includes standards and procedures to:	✓	
• Identify (and if needed, define) the end of treatment; and	✓	
• Calculate the length of a favorable effect beyond the end of treatment.	✓	
Usual Care or Practice Setting: Were systematic standards and procedures used to determine if a study was conducted in a usual care or practice setting?	✓	6.2.2
Risk of Harm: Were systematic standards and procedures used to determine if there is evidence of risk of harm?	✓	6.2.1
Designation: Were systematic standards and procedures used to designate programs and services for HHS consideration (as promising, supported, well-supported, or does not currently meet the criteria)? At a minimum, this includes standards and procedures to:	✓	
• Determine if a program or service has one eligible, well-designed and well-executed study that demonstrates a favorable effect on a target outcome and should be considered for a designation of promising;	✓	6
• Determine if a program or service has at least one eligible, well-designed and well-executed study carried out in a usual care or practice setting that demonstrates a favorable effect on a target outcome at least 6 months beyond the end of treatment and should be considered for a designation of supported; and	✓	6
• Determine if a program or service has at least two eligible, well-designed and well-executed studies with non-overlapping samples carried out in usual care or practice	✓	6

[1]If a cluster randomized study permits individuals to join clusters after randomization, the estimate of the effect of the intervention on individual outcomes may be biased if individuals who join the intervention clusters are systematically different from those who join the comparison clusters.

Table 2. Systematic Review	☑ to Verify	Handbook Section
settings that demonstrate favorable effects on a target outcome; at least one of the studies must demonstrate a sustained favorable effect of at least 12 months beyond the end of treatment on a target outcome; and should be considered for a designation of well-supported.		
Reconciliation of Discrepancies: Were systematic standards and procedures used to reconcile discrepancies across reviewers? (applicable if more than one reviewer per study)	✓	
Author or Developer Queries: Were systematic standards and procedures used to query study authors or program or service developers? (applicable if author or developer queries made)	✓	

Table 3. Independent Review

The systematic review must be independent (i.e., objective and unbiased). In the table below, verify that an independent review was conducted using systematic standards and procedures by providing the names of each state agency and external partner that reviewed the program or service. States must answer all applicable questions in the affirmative. Submit MOUs, Conflict of Interest Policies, and other relevant documentation.

List all state agencies and external partners that reviewed programs and services.
The Stephen Group, Inc. in collaboration with MEF Associates

Table 3. Independent Review	☑ to Verify
Was the review independent (conducted by reviewers without conflicts of interest including those that authored studies, evaluated, or developed the program or service under review)?	✓
Was a Conflict of Interest Statement signed by reviewers attesting to their independence? If so, attach the statement.	✓
Was a Memorandum of Understanding (MOU) signed by external partners (if applicable)? If so, attach MOU(s).	✓

Sections III-V: Describe and Document Findings from Each Program and Service Reviewed and Submitted

Section III. Review of Programs and Services
(*Complete Tables 4-5 for each program or service reviewed.)*

Table 4. Determination of Program or Service Eligibility

Fill in the table below for each program or service reviewed.

Table 4. Determination of Program or Service Eligibility:	☑ **to Verify**
Does the program or service have a book, manual, or other available documentation specifying the components of the practice protocol and describing how to administer the practice? Provide information about how the book/manual/other documentation can be accessed OR provide other information supporting availability of book/manual/other documentation. Program implementation materials were made available to reviewers in digital copies. The implementation manual, *Wheels of Change* © was made available as a digital training manual through the e-learning platform *Mindflash*. Access to this platform was provided to the reviewers by the model developer for the purpose of verification. In addition to access to the online training materials, the model developer provided the following documents for review: • Program Design and Implementation Guide • Path of Implementation for Providers • FCT Readiness Assessment Interview Plan • Implementation Driver Assessment – closed copy • Fidelity Adherence Compliance Tracker (FACT) – copy • Fidelity Implementation-strategy Tool • Readiness Assessment Report – Example Redact • Readiness Assessment Matrix (RAM) • Definitive Report on FCT – 1 of 6 required readings for FCT to achieve Certification	Yes.
Is the program or service a mental health, substance abuse, in-home parent-skill based, or kinship navigator program or service? Identify the program or service area(s).	In-Home Parent Skill-Based Mental Health

Table 5. Determination of Study Eligibility

Fill in the table below for each study of the program or service reviewed. Provide a response in every column; N/A or unknown are not acceptable responses. The response in columns iii, v, vi, vii, and ix must be "yes" or "no." The response in column ix is "yes" only when the responses in columns iii, v, vi, and vii are "yes."

i. Study Title/Authors	ii. Publicly Available Location	iii. Is the study in English? (Yes/No)	iv. Design (RCT, QED, or other). If other, specify design.	v. Did the intervention condition receive the program or service under review in accordance with the book/manual/documentation? (Yes/No)	vi. Did the comparison condition receive no or minimal intervention or treatment as usual? (Yes/No)	vii. Did the study examine at least one target outcome? (Yes/No)	viii. Year Published	ix. Eligible for Review? (Yes/No)
Indiana Waiver Substudy: The Indiana University Evaluation Team & The Department of Child Services. (2018). Indiana Department of Child Services Child Welfare Title IV-E Waiver Demonstration Project Final Report. Indianapolis, IN: Indiana University School of Social Work and Indiana Department of Child Services.	https://www.in.gov/dcs/files/20180102FinalReportfromDCSandIU.pdf	Yes	QED	Yes	Yes	Yes	2018	Yes
Sullivan, et al.: Sullivan, M.B., Bennear, L.S., & Honess, K. (revised 2011). A quasi-experimental evaluation of Family Centered Treatment in the Maryland Department of Juvenile Services Community Based Non-residential Program: A report to Maryland	http://www.bscc.ca.gov/wp-content/uploads/JOJJVol2_Iss1.pdf ; full report available upon request from authors	Yes	QED	Yes	Yes	Yes	Report revised 2011; article published in 2012	Yes

i. Study Title/Authors	ii. Publicly Available Location	iii. Is the study in English? (Yes/No)	iv. Design (RCT, QED, or other). If other, specify design.	v. Did the intervention condition receive the program or service under review in accordance with the book/manual/documentation? (Yes/No)	vi. Did the comparison condition receive no or minimal intervention or treatment as usual? (Yes/No)	vii. Did the study examine at least one target outcome? (Yes/No)	viii. Year Published	ix. Eligible for Review? (Yes/No)
Department of Juvenile Services and Institute for Family Centered Services. Great Falls, VA: FamiliFirst. Abridged results reported in Sullivan, M. B., Bennear, L. S., Honess, K. F., Painter Jr, W. E., & Wood, T. J. (2012). Family Centered Treatment®--an alternative to residential placements for adjudicated youth: outcomes and cost-effectiveness. *Journal of juvenile justice*, *2*(1), 25-40.								

Section IV. Review of "Well-designed" and "Well-executed" Studies
(*Complete Tables 6-10 for each program or service reviewed.)*

Table 6. Studies that are "Well-Designed" and "Well-Executed"[2]

Provide an electronic copy of each of the studies determined to be eligible for review and determined to be "well-designed" and "well-executed."

List all eligible studies that are "well-designed" and "well-executed' (Study Title/Author)
Indiana Waiver Substudy: • The Indiana University Evaluation Team & The Department of Child Services. (2018). Indiana Department of Child Services Child Welfare Title IV-E Waiver Demonstration Project Final Report. Indianapolis, IN: Indiana University School of Social Work and Indiana Department of Child Services.
Sullivan, et al. • Sullivan, M.B., Bennear, L.S., & Honess, K. (revised 2011). A quasi-experimental evaluation of Family Centered Treatment in the Maryland Department of Juvenile Services Community Based Non-residential Program: A report to Maryland Department of Juvenile Services and Institute for Family Centered Services. Great Falls, VA: FamiliFirst. • Sullivan, M. B., Bennear, L. S., Honess, K. F., Painter Jr, W. E., & Wood, T. J. (2012). Family Centered Treatment®--an alternative to residential placements for adjudicated youth: outcomes and cost-effectiveness. Journal of juvenile justice, 2(1), 25-40.

[2] For reference, the Prevention Services Clearinghouse Handbook Chapter 5 defines "well-designed" and "well-executed" studies as those that meet design and execution standards for high or moderate support of causal evidence. Prevention Services Clearinghouse ratings apply to contrasts reported in a study. A single study may have multiple design and execution ratings corresponding to each of its reported contrasts.

Table 7. Study Design and Execution

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below. Provide a response in every column; N/A or unknown are not acceptable responses for columns i, ii, iii, v, vi, and vii. The response in column ii must be "yes."

i. Study Title/Authors	ii. Verify the Absence of all Confounds? (Yes/No)	iii. List Measures that Achieved Baseline Equivalence	iv. List Measures that did NOT Achieve Baseline Equivalence but were Statistically Controlled for in Analyses	v. Overall Attrition[3] (for RCTs only)	vi. Differential Attrition[4] (for RCTs only)	vii. Does Study Meet Attrition Standards?	viii. Notes, as needed
Indiana Waiver Substudy	Yes	Risk classification: very high risk	Treatment and comparison groups were satisfactorily equalized on baseline characteristics using propensity score matching. The impact model did not include any adjustment for pre-treatment characteristics.	The study is a QED.	The study is a QED.	The study is a QED.	A risk classification of "very high risk" was identified by the reviewer as a suitable pretest alternative for all outcome variables.
Sullivan, et al.	Yes	• Proportion of youth with placements: community based residential • Placement frequency: community based residential • Placement duration in days: community based residential • Proportion of youth with placements: secure detention	Treatment and comparison groups were satisfactorily equalized on baseline characteristics using propensity score matching. The impact model did not include any	The study is a QED.	The study is a QED.	The study is a QED.	

[3] For reference, the Prevention Services Clearinghouse Handbook section 5.6 defines o*verall attrition* as the number of individuals without post-test outcome data as a percentage of the total number of members in the sample at the time that they learned the condition to which they were randomly assigned.

[4] For reference, the Prevention Services Clearinghouse Handbook section 5.6 defines *differential attrition* as the absolute value of the percentage point difference between the attrition rates for the intervention group and the comparison group.

i. Study Title/Authors	ii. Verify the Absence of all Confounds? (Yes/No)	iii. List Measures that Achieved Baseline Equivalence	iv. List Measures that did NOT Achieve Baseline Equivalence but were Statistically Controlled for in Analyses	v. Overall Attrition[3] (for RCTs only)	vi. Differential Attrition[4] (for RCTs only)	vii. Does Study Meet Attrition Standards?	viii. Notes, as needed
		• Placement frequency: secure detention • Proportion of youth with offenses: category 1 • Frequency of offenses: all categories • Frequency of adjudicated offenses: category 1 • Proportion of adjudicated offenses: all categories • Frequency of adjudicated offenses: all categories	adjustment for pre-treatment characteristics.				

Table 8. Study Description

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below to describe the practice setting and study sample as well as affirm that the program or service evaluated was not substantially modified or adapted from the version under review. Provide a response in every column; N/A or unknown are not acceptable responses. The response in column v must be "yes."

i. Study Title/Authors	ii. Was the study conducted in a usual care or practice setting? (Yes/No)	iii. What is the study sample size?	iv. Describe the sample demographics and characteristics of the intervention group	v. Describe the sample demographics and characteristics of the comparison group	vi. Verify that the program or service evaluated in the study was ***NOT*** substantially modified or adapted from the manual or version of the program or service selected for review (Yes/No)
Indiana Waiver Substudy	Yes	N = 374 Intervention: 187 Matched Comparison: 187	• 49.2% male, 50.8% female • 89.3% white, 6.42% black, 0.00% American Indian • 75.4% designated CHINS (child in need of services) • 99.1% with reunification as permanency goal • 32.1% classified as very high risk	• 50.2% male, 49.7% female • 86.6% white, 13.4% black, 4.28% American Indian • 69.5% designated CHINs • 95.8% with reunification as permanency goal • 33.2% classified as very high risk	Yes
Sullivan, et al.	Yes	N = 2,234 Intervention: 446 Matched Comparison: 1,788	• Age at first offense: 12.85 • Age at intake: 15.20 • Proportion of males: .75 • Proportion African American: .31 • Proportion Caucasian: .31 • Proportion Hispanic: .08	• Age at first offense: 12.86 • Age at intake: 15.19 • Proportion of males: .73 • Proportion African American: .59 • Proportion Caucasian: .33 • Proportion Hispanic: .077	Yes

Table 9. Favorable Effects

*For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below listing only target outcomes with **favorable effects**. Provide a response in every column; N/A or unknown are **not acceptable** responses.*

i. Study Title/Authors	ii. List the Target Outcome(s)	iii. List the Outcome Measures	iv. List the Reliability Coefficients for Each	v. Are Each of the Outcome Measures Valid?	vi. Are Each of the Outcome Measures Systematically Administered?	vii. List the P-Values for Each of the Outcome Measures	viii. List the Size of Effect for Each of the Outcome Measures	ix. Indicate the Length of Effect Beyond the End of Treatment (in months)
Indiana Waiver Substudy	Permanency	Remaining in home throughout involvement with child welfare	Measure is drawn from administrative data and presumed reliable per section 5.9.2 of the Handbook.	Yes	Yes	< .001	.41	Minimum 0 months; the time between end of treatment and case closure would differ for each child. Treatment did not continue after case closure.
Indiana Waiver Substudy	Permanency	Days to reunification	Measure is drawn from administrative data and presumed reliable per section 5.9.2 of the Handbook.	Yes	Yes	< .001	-.32	Minimum 0 months; the time between end of treatment and reunification would differ for each child.
Sullivan, et al.	Permanency	Proportion of youth with residential placements	Measure is drawn from administrative data and presumed reliable per section 5.9.2 of the Handbook.	Yes	Yes	<.001	-.30	12 months post-treatment

i. Study Title/Authors	ii. List the Target Outcome(s)	iii. List the Outcome Measures	iv. List the Reliability Coefficients for Each	v. Are Each of the Outcome Measures Valid?	vi. Are Each of the Outcome Measures Systematically Administered?	vii. List the P-Values for Each of the Outcome Measures	viii. List the Size of Effect for Each of the Outcome Measures	ix. Indicate the Length of Effect Beyond the End of Treatment (in months)
Sullivan, et al.	Permanency	Frequency of residential placements	Measure is drawn from administrative data and presumed reliable per section 5.9.2 of the Handbook.	Yes	Yes	<.001	-.18	12 months post-treatment

[a] Of those youth who were pending placement, days spent pending placement.

Table 10. Unfavorable Effects

For each study eligible for review and determined to be "well-designed" and "well-executed," fill out the table below listing only target outcomes with ***unfavorable effects****. Provide a response in every column; N/A or unknown are not acceptable responses.*

i. Study Title/Authors	ii. List the Target or Non-Target Outcome(s)	iii. List the Outcome Measures	iv. List the Reliability Coefficients for Each	v. Are Each of the Outcome Measures Valid?	vi. Are Each of the Outcome Measures Systematically Administered?	vii. List the P-Values for Each of the Outcome Measures	viii. List the Size of Effect for Each of the Outcome Measures	ix. Indicate the Length of Effect Beyond the End of Treatment (in months)
Sullivan, et al.	Child Well-Being	Proportion of youth with adjudications	Measure is drawn from administrative data and presumed reliable per section 5.9.2 of the Handbook.	Yes	Yes	<.001	.20	12 months post-treatment
Sullivan, et al.	Child Well-Being	Frequency of adjudications	Measure is drawn from administrative data and presumed reliable per section 5.9.2 of the Handbook.	Yes	Yes	<.001	>22	12 months post-treatment

NOTE: The study authors argue that the increase in adjudications is evidence of FCT's emphasis on accountability, given that the number of offenses is the same across groups.

"Post-treatment offenses committed by the youth in this treatment sample are more likely to be adjudicated, however, and the [SATT] effect size is curiously large. The number of offenses committed over the follow-up period that were adjudicated were measured and the frequency of offenses is the same across groups. This must be reflective of court decisions as applied to the youth receiving FCT. This outcome may be interpreted as a manifestation of the emphasis on accountability in Family Centered Treatment; the model attempts to instill accountability by accepting responsibility for one's actions as a family system value. This may be exhibited in the family's interactions with the courts as an increase in the likelihood of an offense being adjudicated. Overall, however, the fact that residential placements and days in detention are substantially lower suggests that the average youth receiving FCT committed fewer offenses of a nature that would warrant a consideration of removal from the community."

Section V. Program or Service Designation for HHS Consideration

Table 11. Program or Service Designation for HHS Consideration

Fill out the table below for the program or service reviewed. Only select one designation. Answer questions relevant to the selected designation; relevant questions must be answered in the affirmative.

Table 11. Program or Service Designation for HHS Consideration	☑ **to Verify**
There is ***NOT*** sufficient evidence of risk of harm such that the overall weight of evidence does not support the benefits of the program or service.	Yes
	☑ **the Designation and Provide a Response to the Questions Relevant to that Designation**
Well-Supported	
• Does the program or service have at least two eligible, well-designed and well-executed studies with non-overlapping samples[5] that were carried out in a usual care or practice setting?	Yes
• Does one of the studies demonstrate a sustained favorable effect of at least 12 months beyond the end of treatment on at least one target outcome	Yes
Supported	
• Does the program or service have at least one eligible, well-designed and well-executed study that was carried out in a usual care or practice setting and demonstrate a sustained favorable effect of at least 6 months beyond the end of treatment on at least one target outcome?	Yes
Promising	
• Does the program or service have at least one eligible, well-designed and well-executed study and demonstrate a favorable effect on at least one 'target outcome'?	Yes

[5]Samples across multiple sources of a study are considered overlapping if the samples are the same or have a large degree of overlap. Findings from an eligible study determined to be "well-executed" and "well-designed" may be reported across multiple sources including peer-reviewed journal articles and publicly available government and foundation reports. In such instances, the multiple sources would have overlapping samples. The findings across multiple sources with these overlapping samples should be considered **one** study when designating a program or service as "well-supported," "supported," and "promising."