Washington State Institute for Public Policy
110 Fifth Avenue SE, Suite 214 • PO Box 40999 • Olympia, WA 98501 • 360.664.9800 • www.wsipp.wa.gov

June 2023

Washington State's Functional Family Therapy Program: *Outcome Evaluation*

In 2022, the Board of Directors for the Washington State Institute for Public Policy (WSIPP) approved a contract with the Washington State Department of Children, Youth, and Families (DCYF). This contract directed WSIPP to conduct an evaluation of the relationship between youth participation in Functional Family Therapy (FFT) and recidivism outcomes for juvenile court youth.

Functional Family Therapy is a family-based treatment program for at-risk youth intended for youth aged 11-18 and typically includes 12 weekly sessions, each lasting approximately one hour. FFT therapists customize the program to target the specific needs of each youth and their families. The program could include teaching problem-solving techniques, communication strategies, and conflict-resolution skills, among others.

Section I provides additional context for the study and background information about the FFT program model. Section II outlines the data sources. Section III reviews the evaluation methodology. Section IV discusses the findings. Finally, Section V discusses the limitations of our findings and offers final takeaways.

Summary

In Washington State, Functional Family Therapy (FFT) is one of the many evidence-based programs made available to court-involved youth on probation. In 2022, WSIPP was contracted to evaluate the effect of the program on recidivism.

Using administrative data, this study examined the likelihood of recidivism for youth participating in FFT relative to eligible youth who did not participate in FFT. In addition, we evaluated for whom, and under what conditions, the program was most effective.

Our findings indicate that participation in FFT is associated with an increased likelihood of recidivism. On average, youth who started FFT were 10.1 percentage points more likely to recidivate than youth in the comparison group. Of those who recidivated, there were no significant differences found in the rates of felony or violent felony recidivism. The association between participation in FFT and recidivism did not vary based on youth characteristics, geography, living situation, or competency of therapist.

Youth who completed FFT experienced a lower likelihood of recidivism in relation to youth who did not complete FFT but remained more likely to recidivate than youth in the comparison group.

Suggested citation: Spangler, M., & Gibson, C. (2023). *Washington State's Functional Family Therapy Program: Outcome evaluation* (Document Number 23-06-3901). Olympia: Washington State Institute for Public Policy.

I. Background

In 1997, the Washington State Legislature passed the Community Juvenile Accountability Act (CJAA), funding community-based programs to reduce the likelihood of recidivism among convicted juveniles.[1] One of these programs was Functional Family Therapy (FFT).[2]

In 2004, WSIPP released a report showing evidence of a significant reduction in felony recidivism when FFT is delivered competently.[3] Since then, in 2017, the Washington State Center for Court Research (WSCCR) reevaluated FFT and did not find evidence of a reduction in recidivism.[4] In 2022, WSIPP released a report showing some evidence of reductions in overall and felony recidivism for specific groups of male and female youth.[5]

Recent research suggests that the populations and needs of court-involved youth have evolved since the initial passage of the CJAA.[6] Consistent with the CJAA Advisory Committee's short- and long-term research goals, CJAA contracted with WSIPP to conduct an updated evaluation of FFT for court-involved youth.[7]

FFT Program Model

FFT is a structured home-based family intervention for moderate- to high-risk youth.[8] Over the course of 10 to 12 sessions, therapists work with families to move through key phases of treatment. Therapists aim to engage all family members with the program, understand how they relate to each other, teach them different ways to interact and make sustainable improvements by planning for risky situations. For more information on the program, see Exhibit 1.

[1] RCW 13.40.500 through 13.40.540.
[2] The remaining CJAA funded programs are Aggression Replacement Training (ART), Coordination of Services (COS), Education and Employment Training (EET), Family Integrated Transitions (FIT), and Multi-Systemic Therapy (MST).
[3] Barnoski, R. (2004). *Outcome evaluation of Washington State's research-based programs for juvenile offenders* (Doc. No. 04-01-1201). Olympia: Washington State Institute for Public Policy.
[4] Peterson, A. (2017) *Functional Family Therapy in a probation setting: Outcomes for youths starting treatment January 2010-September 2012*. Olympia, WA: Center for Court Research, Administrative Office of the Courts.
[5] Knoth-Peterson, L., Gibson, C., & Adams, N. (2022). *What works for whom? Juvenile court assessment tool and program eligibility* (Doc. No. 22-06-1902). Olympia: Washington State Institute for Public Policy.
[6] Knoth, L., Drake, E., Wanner, P., & Westley, E. (2020). *Washington State's juvenile justice system: Evolution of policies, populations, and practical research* (Doc. No. 20-01-1901). Olympia: Washington State Institute for Public Policy.
[7] Washington State Department of Children, Youth, and Families. (2022). *Report to the Washington State Legislature, block grant proviso report.*
[8] Barnoski, R. (2009). *Providing evidence-based programs with fidelity in Washington State juvenile courts: Cost analysis* (Doc. No. 09.12.1201). Olympia: Washington State Institute for Public Policy.

Research Questions

The current study examines the relationship between FFT and recidivism, with supplementary analyses examining youth subpopulations and particular program conditions. Our evaluation focuses on three research questions.

1) *Does FFT significantly reduce the likelihood of recidivism?*

We analyze the relationship between FFT participation and recidivism generally. We also test the relationship between FFT on types of recidivism, looking at the likelihood that youth who recidivate do so with a felony or a violent felony versus less serious offenses.

2) *For whom is FFT most effective?*

We complete the same analyses of general recidivism and type of recidivism (felony or violent felony) for subpopulations of youth. First, we look at subpopulations based on the youth's characteristics (race, gender, age, and risk level). Then, we look at subpopulations based on court geography and family composition when they enter FFT.[9]

3) *Under what conditions is FFT most effective?*

We test whether the relationship between FFT participation and recidivism varies by program characteristics (e.g., whether treatment was completed, what phase the youth reached, and the competency rating of the therapist).

Exhibit 1
Functional Family Therapy:
Overview and Eligibility

Program Overview

FFT is a family-based treatment program for moderate- to high-risk youth. FFT is intended for youth aged 11-18.

Over the course of 10 to 12 sessions, the therapist works through the five phases in the FFT program model: Engagement, motivation, relational assessment, behavior change, and generalization. Each phase has a set of goals to be accomplished before moving on to the next phase.

In our analysis, three distinct categories of phases are identified to signal progression through the FFT model. The three phase categories are,

1) Engagement/motivation—*Engaging all family members in the program.*
2) Behavior change—*Teaching a set of strategies based on the family's needs.*
3) Generalization—*Helping the family plan for situations that could cause relapse into bad behaviors.*

Eligibility

Eligibility for FFT is based on scores from a risk assessment administered to youth under supervision by a juvenile court. Eligible youth must—

- Be classified as moderate- or high-risk and
- Exceed a particular risk score on domain 7B on the full assessment, related to the youth's current living situation.

[9] Court geography is categorized as east–urban, east–rural, west–urban, and west–rural courts. A court is designated as urban or rural based on the county they are in, according to the rural county definition based on population density. This information is published by the Office of Financial Management and is based on RCW 82.14.370.

II. Data

This evaluation uses data from three sources: 1) the Juvenile Assessment Research Database (ARD), 2) the WSIPP Criminal History Database (CHD), and 3) therapist and session information provided by FFT, LLC. Exhibit 2 provides an overview of each data source.

We identified court-involved youth who were eligible for FFT from the ARD. For this evaluation, we focus on youth who were in the court system but not detained. We then linked each youth's associated court cases in the CHD. We also linked the ARD and CHD data to the FFT, LLC data on therapist adherence and session information where possible.

The datasets used lack a common identifier to link the court case to a specific risk assessment and, further, to a specific therapist and string of session information. We used multiple strategies to connect these data reliably, but exact precision is not possible. Appendix I provides details on the process used to create a final dataset.

Exhibit 2

FFT Outcome Evaluation Source Data Details

Dataset name	Data source	Information included
Assessment Research Database (ARD)	WAJCA; AOC	Assessment information for the Back on Track (BOT) and Positive Achievement Change Tool (PACT); referrals to and participation in evidence-based programs.
Criminal History Database (CHD)	AOC; DOC; DCYF; WSIPP	Combines court data from AOC, incarceration, and community supervision data from DOC, and residential confinement data from DCYF.
FFT, LLC therapist, youth, and session data	FFT, LLC	Contains employment and quality assurance (QA) data for therapists working with juvenile court cases. This data also houses youth session and demographic data.

Notes:
WAJCA = Washington Association of Juvenile Court Administrators.
AOC = Administrative Office of the Courts.
DOC = Department of Corrections.
DCYF = Department of Children, Youth, and Families.

III. Evaluation Methodology

The primary analysis in this report compares the recidivism outcomes of youth who were eligible and participated in FFT (the treatment group) with youth who were eligible but did not participate in FFT (the comparison group). In an ideal research setting, placement into the treatment is completely randomized. In the case of random assignment, we are assured that there are no significant differences in the treatment and control groups beyond the participation in the treatment. In this setting, we can attribute any differences in outcomes to the treatment itself.

Participation in FFT in today's juvenile justice setting is not randomized. Youth eligible for FFT choose whether to participate in consultation with their Juvenile Probation Counselor (JPC). Youth are often choosing from a menu of rehabilitative options.[10] This aspect of choice means that differences in outcomes may be explained by unobserved differences in the type of youth who decide to participate in FFT. In other words, differences in outcomes may not be wholly due to participation in FFT.

To combat this, we employ advanced statistical methods that reduce the effects caused by systematic differences between those in treatment and those in the comparison group.[11]

Treatment and Comparison Groups

The treatment group consists of youth who were eligible for and participated in FFT between November 2003 and July 2017. As a reminder, youth included in this evaluation were involved in the juvenile justice system but not detained in a facility. We include youth who started FFT regardless of whether they completed the program.

The comparison group is a "treatment-as-usual" group, which consists of youth who were eligible for FFT during the same time period but did not participate. The comparison group youth either participated in a non-FFT evidence-based program (EBP), participated in a local option, or did not elect to participate in any rehabilitative program.[12]

[10] 99.7% of the treatment and comparison groups were eligible for more than one EBP in our sample. This does not include other local options available in lieu of or in addition to an EBP.

[11] Even after using advanced statistical methods, we cannot completely account for any systemic differences between the treatment and comparison groups. Specifically, if there are differences between the two groups that are unobservable given the data available to us, we are not able to account for those differences.

[12] WSIPP's previous evaluation, Barnoski (2004), was able to take advantage of a waitlist design to induce randomization. Programs were also newer and less widely available during the period of that study. With that, the study findings were able to approximate the effect of FFT relative to a no-treatment control. Today, juvenile court youth are likely to receive treatment of some kind, whether that is an EBP or some other local treatment option. In Appendix II we test the robustness of our findings by limiting our comparison group to those who received no other EBP in the follow-up period and found the results remained consistent with what is reported in the main body of the report. This robustness check, however, is still unable to account for participation in non-EBP options or speak to the reasons why these youth might not participate in EBPs.

Methods

In our study, we use entropy balancing to construct a comparison group of youth equivalent to the youth who participate in FFT based on observable characteristics.[13] We then used regression analysis to estimate the likelihood of recidivism and compute the difference between the treatment and comparison groups.

Exhibit 3 summarizes the characteristics of youth who participated in FFT. Using entropy balancing, the comparison group is reweighted to match the treatment group.

Our findings are presented in Section IV. First, we present predicted probabilities of recidivism for the entire group of youth who participated in FFT relative to those who were eligible but did not participate. Then, we discuss variation in effects by gender, race, age, risk level, and court geography. Finally, we discuss further variation by program completion, phase reached, and therapist competence.

Our results represent average correlations between participation in FFT and recidivism after controlling for observable differences between our treatment and comparison groups. Given that our findings are not necessarily causal, if we find a negative effect, this simply suggests that other programs may be more effective than FFT at reducing recidivism. A negative finding would not necessarily mean that FFT actively harms individuals who participate.

[13] Entropy balancing is a weighting method that reweights the comparison observations such that the treatment (participated in FFT) and comparison (did not participate in FFT) groups have the same mean for the identified characteristics.

Exhibit 3

Court-Involved Youth FFT Sample Characteristics (N = 6,104)

Variable		Proportion
Recidivism rate		53.8%
Demographics		
Age		
	Under 15	26.3%
	15 and older	73.7%
Sex		
	Female	29.7%
	Male	70.3%
Race/Ethnicity		
	White	68.4%
	Black	13.0%
	Hispanic/Latino	12.8%
	American Indian / Alaska Native*	2.6%
	Asian / Hawaiian / Pacific Islander*	2.2%
	Other/unknown*	1.0%
Risk Level		
	Moderate	41.1%
	High	58.9%
Geography		
	East-Rural	9.8%
	East-Urban	20.6%
	West-Rural	18.4%
	West-Urban	51.2%
Living situation		
	Living with parents	91.9%
	Not living with parents	8.1%
Current conviction		
Crime category		
	Drug	6.7%
	Person	31.5%
	Property	43.4%
	Sex	0.8%
	Other	17.7%
Crime grade		
	Misdemeanor	74.2%
	Felony, violent and non-violent	19.2%
	Violent felony	6.6%

Notes:
*Groups listed were not included in subgroup analyses due to sample size considerations.
Proportions listed represent total observations, not proportions of unique individuals. Individuals can appear multiple times in our sample.

IV. Recidivism Analysis

We analyze the relationship between participation in FFT and three recidivism outcomes:

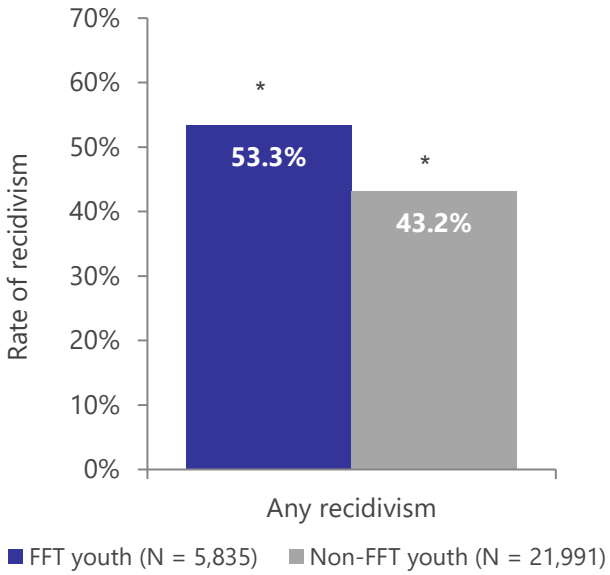
- Any recidivism,
- Felony recidivism (versus misdemeanor recidivism), and
- Violent felony recidivism (versus non-violent felony and misdemeanor recidivism).

For felony and violent felony recidivism, we classified youth based on the most serious offense in the youth's first recidivism event. For these two outcomes, we report the percentage of youth who committed that type of offense of the youth who recidivated.[14]

The results are presented in Exhibit 4 and Exhibit 5.[15] Overall, youth who participated in FFT were more likely to recidivate than similar youth who did not participate in FFT.[16] However, the recidivists who participated in FFT were about as likely as those in the comparison group to recidivate with a felony or violent felony charge. In other words, among recidivists, those who participated in FFT did not exhibit a different pattern in the severity of the crime.

Exhibit 4

18-Month Predicted Probabilities, by Treatment Status



Note:
* Significant at the 0.05 level.

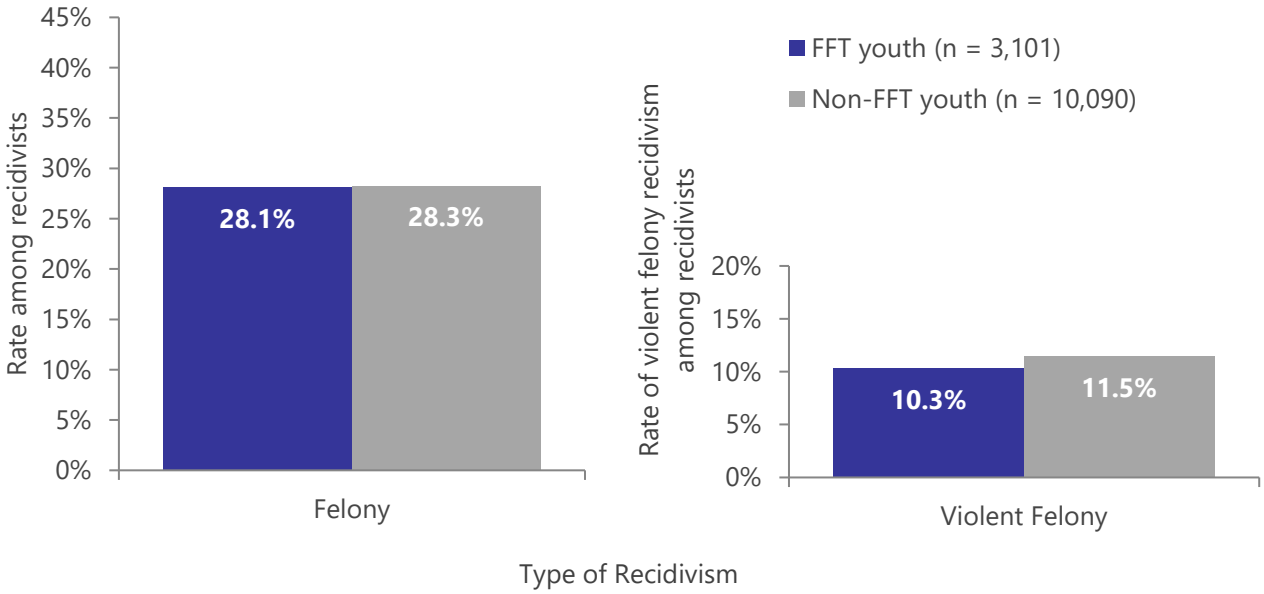
[14] Previous evaluations have conducted analyses by type of recidivism, see Knoth, L., Wanner, P., & He, L. (2019). *Washington State recidivism trends: FY 1995-2014.* (Doc. No. 19-03-1901). Olympia: Washington State Institute for Public Policy). These evaluations compared one type of recidivism (i.e., misdemeanor) against all other collective possible outcomes (i.e., felony recidivism and no recidivism). In this evaluation, we are comparing types of recidivism within the pool of individuals who recidivated.

[15] We conducted linear regression analyses on the balanced samples. More details on the covariates used in the entropy balancing process as well as the regression model can be found in Appendix II. Exhibit 4 shows the predicted likelihoods of recidivism.

[16] Statisticians often rely on a metric, the p-value, to determine whether an effect is significant. The p-value is a measure of the likelihood that the difference could occur by chance—values range from 0 (highly significant) to 1 (no significant difference). By convention, p-values less than 0.05 (a 5% likelihood that the difference could occur by chance) are considered statistically significant.

Exhibit 5

18-Month Predicted Felony Rates Among Recidivists, by Treatment Status



Note:
Neither of the results were significant at a standard statistical level.

Youth Subgroup Analyses

Although we did not find evidence that FFT lowers the likelihood of recidivism for youth overall, it may be effective for groups of individuals. To examine this, we replicated the main analyses with various subgroups of court-involved youth.[17]

Exhibit 6 displays the average relationship between participation in FFT and the rate of recidivism for all subgroups. The "Estimate" column indicates the estimated percentage point (pp) difference between the FFT group and non-participants in their likelihood of recidivism. Values greater than zero indicate that recidivism was higher in the FFT group.

We found that across all subgroups, youth who participated in FFT had a greater likelihood of recidivating than youth who did not. The magnitude of the relationship varied only slightly among the subgroups.[18]

Exhibit 6

Recidivism Outcomes, by Subgroup

Average relationship between FFT participation and recidivism (relative to eligible youth who did not participate in FFT)				
Category	Group	Estimate (pp)	*P-value*	N
Full group		9.8	*<0.001*	27,809
Gender	Male	9.5	*<0.001*	20,005
	Female	11.0	*<0.001*	7,803
Race	White	9.3	*<0.001*	17,196
	Black	11.1	*<0.001*	3,601
	Hispanic	11.5	*<0.001*	3,659
Age group	Under 15	14.5	*<0.001*	4,955
	15 and older	8.2	*<0.001*	22,853
Risk level	Moderate	10.4	*<0.001*	10,784
	High	9.4	*<0.001*	17,024
Court geography	East rural	8.5	*0.003*	3,487
	East urban	12.1	*<0.001*	4,045
	West rural	8.9	*<0.001*	4,731
	West urban	10.1	*<0.001*	15,545
Youth's living situation	Lives with parents	9.4	*<0.001*	23,866
	Does not live with parents	14.6	*<0.001*	3,896

Notes:
Coefficient estimates, p-values, and sample sizes represent averages across 100 sampling iterations.
Numbers in the "Estimate" column are percentage point (pp) differences relative to the comparison youth.
For example, male youth in FFT were 9.5 pp more likely to recidivate than male youth in the comparison group.

[17] For each subgroup, we rebalance the sample prior to running the regression. An overview of the characteristics used in entropy balancing and covariates controlled for in our regression model can be found in Appendix II.

[18] The largest variation was seen in the age category. The likelihood of recidivism for youth under 15 was 14.5 percentage points higher than those who did not participate. The likelihood of recidivism for youth 15 and older after participation in FFT was 8.2 percentage points higher than those who did not participate.

Exhibit 7 displays the estimated effect of participation in FFT on the rate of felony and violent felony recidivism for all subgroups.

Among recidivists, we found no significant relationship between participation in FFT and recidivating with a felony offense. For most subgroups, we did not find a significant relationship between participation in FFT and recidivating with a violent felony offense. However, in two subgroups, we found FFT participation was associated with a decreased rate of violent felony recidivism.

Exhibit 7

Felony and Violent Felony Recidivism Outcomes by Subgroup

Average relationship of FFT participation on types of recidivism, as a proportion of all recidivism (relative to eligible youth who did not participate in FFT)						
Category	Group	N	Felony		Violent Felony	
			Estimate (pp)	*P-Value*	Estimate (pp)	*P-Value*
Full group		13,191	-0.2	*0.762*	-1.2	*0.144*
Gender	Male	10,302	0.1	*0.772*	-0.6	*0.557*
	Female	2,889	0.7	*0.708*	-2.2	*0.107*
Race	White	7,644	-0.3	*0.760*	-1.5	*0.111*
	Black	2,004	-1.1	*0.688*	-0.4	*0.741*
	Hispanic	1,919	0.8	*0.730*	-0.5	*0.738*
Age group	Under 15	2,573	2.6	*0.322*	1.1	*0.534*
	15 and older	10,617	-1.6	*0.231*	**-2.3**	*0.013*
Risk level	Moderate	4,287	0.7	*0.676*	**-2.7**	*0.037*
	High	8,905	-1.0	*0.524*	-0.3	*0.706*
Court geography	East rural	1,779	0.4	*0.736*	0.6	*0.749*
	East urban	1,923	-3.9	*0.270*	-3.9	*0.142*
	West rural	2,111	-4.5	*0.164*	-3.4	*0.127*
	West urban	7,378	1.2	*0.452*	-0.5	*0.659*
Youth's living situation	Lives with parents	11,275	-0.3	*0.749*	-1.5	*0.088*
	Does not live with parents	1,890	0.6	*0.780*	1.8	*0.481*

Notes:
Coefficient estimates, p-values, and sample sizes represent averages across 100 sampling iterations.
Estimates which reach the 0.05 level of significance are in bold.
Numbers in the "Estimate" column are percentage point (pp) differences relative to the comparison youth. For example, youth in FFT were 0.2 pp less likely to have recidivated with a felony than youth who recidivated in the comparison group. Youth in FFT were 1.2 pp less likely to have recidivated with a violent felony than youth who recidivated in the comparison group.

First, youth aged 15 and older were 2.3 percentage points *less* likely to recidivate with a violent felony if they participated in FFT.[19] Second, youth classified as moderate risk were 2.7 percentage points *less* likely to recidivate with a violent felony if they participated in FFT.[20]

Previously, we examined differences in outcomes based on starting FFT, regardless of what happened after individuals began receiving treatment. Here, we assess differences in outcomes based on the youth experiences in FFT.

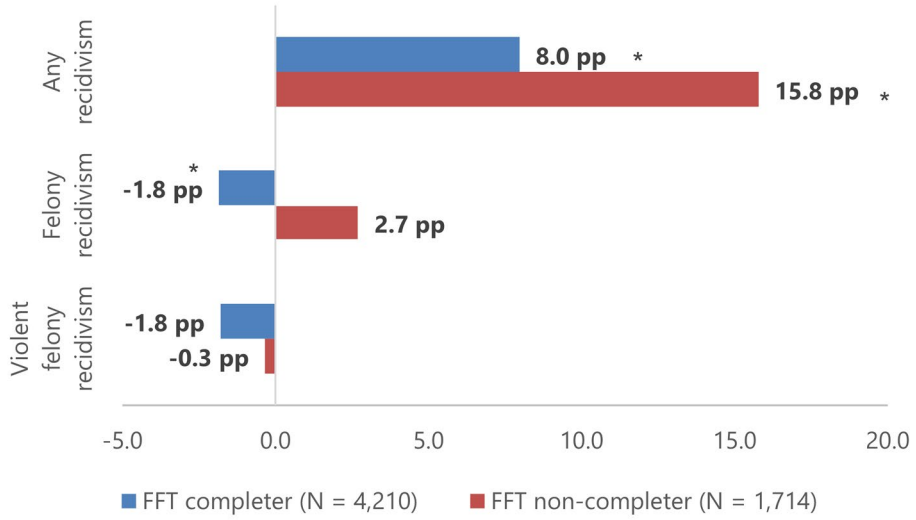
First, we examine whether program participation has a differential correlation with recidivism based on whether the youth completed the program or not.

Program Completion

Exhibit 8 presents the probability of recidivism by program completion status for all three recidivism outcomes. The vertical line at zero represents the recidivism rate of the comparison group. Bars to the right of the vertical line indicate higher recidivism rates than those in the comparison group. Bars to the left of the vertical line indicate lower recidivism rates than those in the comparison group.

Exhibit 8

Estimated Difference in 18-Month Recidivism Rates, by Program Completion Status



Notes:
* Significant at the 0.05 level.
Amounts are differences in likelihood of recidivism relative to the comparison group (FFT completers were 7.98 percentage points (pp) more likely to recidivate than those in the comparison group, FFT non-completers were 15.78 pp more likely to recidivate than those in the comparison group.)

[19] Of youth 15 and older who recidivated, 12.2% of those in the comparison group and 9.9% of those who participated in FFT recidivated with a violent felony offense, after statistical weighting and regression analysis. This difference was statistically significant at the 5% level.

[20] Of youth classified as moderate risk who recidivated, 10.7% of the comparison group and 8% of the youth who participated in FFT recidivated with a violent offense, after statistical weighting and regression analysis. This difference in predicted probability was statistically significant at the 5% level.

Looking at recidivism generally, the comparison group exhibited a 44% recidivism rate, on average.[21] Of those who recidivated, approximately 29% of the comparison group recidivated with a felony.[22] Approximately 12% of those who recidivated in the comparison group did so with a violent felony.[23]

Youth who completed FFT were less likely to recidivate relative to youth who started but did not complete the program. However, youth who completed FFT were still more likely to recidivate relative to youth in the comparison group.

Rates of felony recidivism among the three groups (youth who completed FFT, started but did not complete FFT, and the comparison group) were not statistically significantly different. However, youth who completed FFT were less likely to recidivate with a violent felony relative to youth in the comparison group.

This next set of analyses was done on a subset of youth who started FFT between 2011-2017 and were able to be additionally linked to the FFT, LLC supplied data. Adding this data allows us to look more in-depth at the youth's time in FFT (e.g., how many sessions they had and who their therapist was). More details on the way this sample was constructed can be found in Appendix I.

First, we examine whether program participation has a differential relationship with recidivism based on the phase a youth reached prior to exiting the program. Second, we examine whether therapist competency has a differential relationship with recidivism.

Phase Completion

Exhibit 9 presents the estimated difference in recidivism rates by the last phase the youth reached in the FFT program model. The phases are grouped into three categories—engagement/motivation, behavior change, and generalization.

Once again, the vertical line at zero represents the recidivism rate of the comparison group. At each phase, youth participating in FFT recidivate at higher rates than the comparison group. Those who participate through the generalization phase experience better outcomes than those who exit the program prior to this phase, but those youth still recidivate at a higher rate than the comparison group.

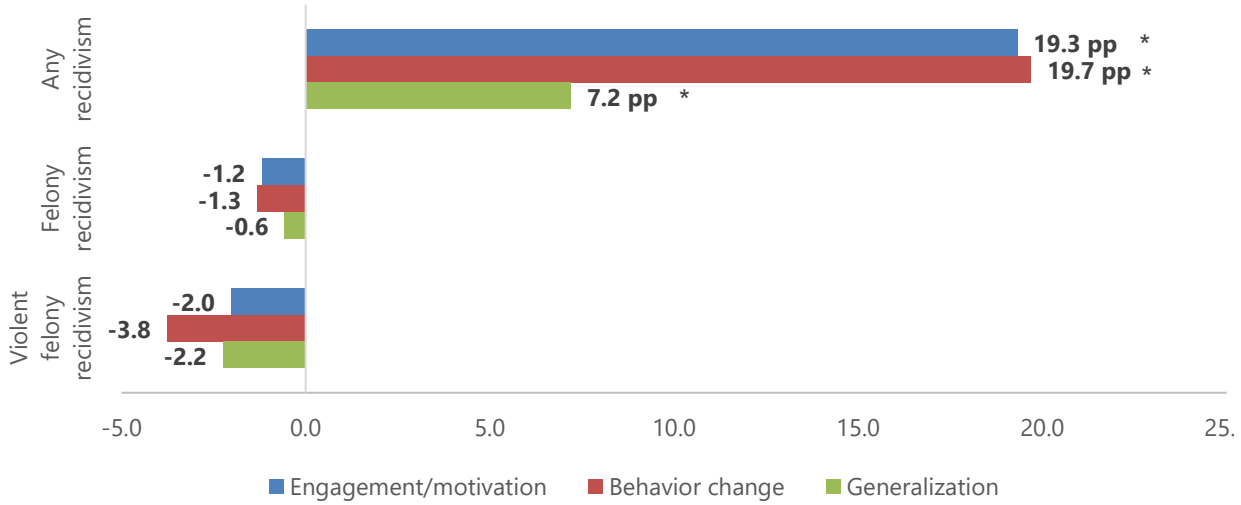
[21] These numbers reflect predicted probabilities after weighting and regression analysis. Given that, the predicted probability of the comparison group's recidivism rate differs slightly between each group (completers or non-completers). When comparing FFT completers to the comparison group, the predicted probability of recidivism was 42.4% for the comparison group and 50.3% for the FFT completers. When comparing FFT non-completers to the comparison group, the predicted probability of recidivism was 45.6% for the comparison group and 61.4% for the FFT non-completers.

[22] When comparing FFT completers to the comparison group, the predicted probability of felony recidivism was 27.3% for the comparison group and 25.5% for the FFT completers. When comparing FFT non-completers to the comparison group, the predicted probability of felony recidivism was 30.4% for the comparison group and 33.1% for the FFT non-completers.

[23] When comparing FFT completers to the comparison group, the predicted probability of violent felony recidivism was 11.2% for the comparison group and 9.4% for the FFT completers. When comparing FFT non-completers to the comparison group, the predicted probability of violent felony recidivism was 12.4% for the comparison group and 12% for the FFT non-completers.

Exhibit 9

Estimated Difference in 18-Month Recidivism Rates, by Phase Completion



Notes:
* Significant at the 0.05 level.
Amounts are differences in likelihood of recidivism relative to the comparison group (FFT participants that only reached the engagement/motivation stage were 19.3 percentage points (pp) more likely to recidivate than those in the comparison group, those that reached the behavior change stage were 19.7 pp more likely to recidivate than those in the comparison group, and those that reached the generalization phase were 7.2 pp more likely to recidivate than those in the comparison group.)

Therapist Competency

Previous research on EBPs, including FFT, has shown that the level of fidelity to the program model being delivered can influence the likelihood of recidivism among those who start treatment.[24] In other words, if the way a program is delivered deviates from the original structure of the model, it may lead to less favorable outcomes.

To enable this type of analysis, FFT, LLC provided WSIPP with session data for court-involved youth. From these data, we could identify the therapist that served a particular youth and their family and any fidelity scores for that therapist.[25] Therapist fidelity scores range from 0-6 and measure how closely the therapist adheres to the model as intended.[26] All therapists are required to score a three or higher by the end of their first year to maintain certification.[27] The compiled ARD/CHD data were linked with the data provided by FFT, LLC to conduct additional analyses for youth who participated in FFT from 2011-2017.

[24] Barnoski (2004).
[25] Therapists receive regular fidelity and dissemination adherence scores from their supervisor. The scores are based on the therapist's work in the prior week. The fidelity score is intended to measure how closely the therapist is adhering to the model as intended.
[26] The median score in our dataset was 4.22.
[27] WA State Functional Family Therapy Project – Therapist Standards.

We combined the fidelity scores into three groups: Highly competent (scores of 4.5 and above), competent (scores equal to or greater than 3 and less than 4.5), and less competent (scores of less than 3).[28]

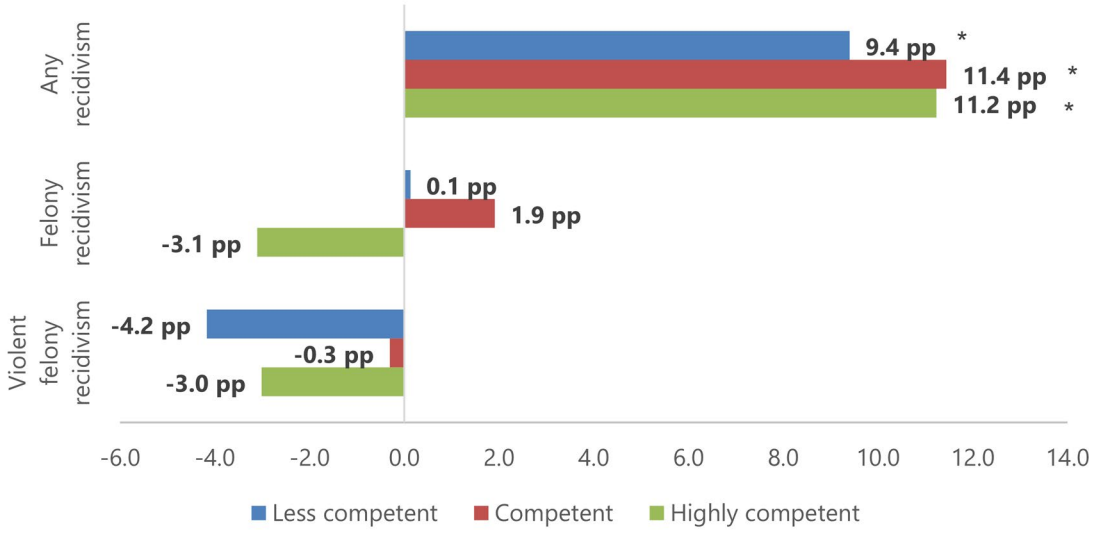
Exhibit 10 presents the estimated difference in recidivism rates by therapist competency level.

Youth experienced higher rates of recidivism relative to youth in the comparison group, regardless of therapist competency. The magnitude of the effect varied only slightly with respect to the different levels of therapist competence.

It is possible that therapists who are known to be highly competent are matched with the youth who are most in need and have a higher likelihood of recidivism. Therefore, this selection process may mask the therapeutic effects of highly competent therapists.

Exhibit 10

Estimated Difference in 18-Month Recidivism Rates, by Therapist Competency



Note:
* Significant at the 0.05 level.
Amounts are differences in likelihood of recidivism relative to the comparison group (FFT participants with less competent therapists were 9.4 percentage points (pp) more likely to recidivate than those in the comparison group, FFT participants with competent therapists were 11.4 pp more likely to recidivate than those in the comparison group, and FFT participants with highly competent therapists were 11.2 pp more likely to recidivate than those in the comparison group.)

[28] We selected 3 as the cutoff for competency because that is the standard that a therapist must meet within their first year working as an FFT certified therapist to maintain their certification.

V. Conclusion

This section discusses the limitations of this study and the impacts those limitations have on the ability to draw meaningful conclusions from our findings. Then we summarize our findings in connection with the original research questions outlined in Section I of the report.

Limitations

The ability of this study to speak to the effectiveness of FFT is limited by the way individuals select to participate in FFT and by data availability.

Selection Bias

Because participation in FFT is not randomized, our analysis may be affected by selection bias. It may be the case that youth who start FFT are different from youth who do not start FFT on unidentifiable characteristics. This would mean that, even after implementing advanced statistical methods to account for bias, some bias would still be present due to unobservable factors. If these differences are related to our outcome measures, then the resulting differences in the likelihood of recidivism cannot be simply attributed to starting FFT. At face value, our results represent average correlations between participation and recidivism conditional on observable characteristics. In practical terms, we are unable to fully distinguish between the effects of the program and differences in the types of individuals who chose to participate in FFT.

Generalizability

Our findings are only generalizable to the kind of youth we observe in our sample. This means that our results may not hold for FFT programs delivered in different states, youth in different age groups, or youth who look otherwise different from those in our final sample population.

Further, our findings are unique to the way that FFT was administered during the time period of our data. Meaning if administrative rules or methods governing the delivery of FFT were changed, our findings would not necessarily hold.

Data Availability

Our main analyses were limited to the youth who were able to be reliably matched to a court case in the Criminal History Database.[29] Additional analyses were limited to the youth who were able to be further reliably matched to a case in the FFT, LLC database.[30] There is no unique identifier in the ARD database that can be used to identify the corresponding court case in the Administrative Office of the Courts (AOC's) court records. Similarly, no unique identifier in the FFT, LLC data can be used to identify the corresponding assessment and intervention record in the ARD database.

[29] 87.06% of cases were able to be reliably matched from the ARD to the CHD.

[30] 50.87% of the 4,871 identified court-involved youth trips in FFT, LLC's data were able to be reliably matched to the ARD and CHD, resulting in a final dataset of 2,478 observations for the therapist competency and phase completion analyses.

In addition to CJAA-funded EBPs, court-involved youth may participate in treatment programs funded and administered by the local juvenile court or other forms of treatment, such as in-patient drug treatment. We did not have access to data capturing participation in non-CJAA-funded programs. It is possible that FFT-eligible youth participated in alternative forms of treatment that affected the overall recidivism outcomes. Future research would benefit from having greater access to data regarding participation in these alternative forms of treatment.

Comparability to Prior Evaluations

Several factors inhibit the ability to directly compare this evaluation with previous evaluations of FFT in Washington State.[31] First, WSIPP's 2004 study of FFT used a wait-list comparison group, meaning youth in the comparison group received no treatment. In this study, we use a treatment-as-usual approach to select our comparison group. In today's juvenile court environment, in contrast with the environment back in the late 1990s, most youth receive some rehabilitative treatment either as an EBP or another local option. If these alternative forms of treatment are effective, then our comparison group may have a lower likelihood of recidivism compared to previous evaluations. This phenomenon could mask the ability to see overall therapeutic effects from FFT in our study.

Unlike previous evaluations, we measure the type of recidivism as the most serious offense in the first recidivism event during the follow-up period.[32]

Finally, this study includes a wider breadth and depth of data in its analyses. This evaluation used a sample of youth eligible for FFT from November 2003 to July 2017. This resulted in significantly more observations than in previous evaluations. Also, this evaluation was able to complete supplementary analyses using FFT, LLC data on therapist competency and session information.

In 2022, WSIPP released a report showing FFT exhibited therapeutic effects for male and female participants who gave specific responses on the initial risk assessment that determines eligibility.[33] This is not something we could explore further in this study, but it suggests that FFT may indeed be beneficial for specific categories of youth, even though it is associated with higher levels of recidivism for youth *on average*. Future studies could more deeply explore a differential relationship between participation in FFT and recidivism by specific response patterns on the risk assessment.

Summary of Findings

Overall, this study found that court-involved youth participating in FFT in Washington State had a higher likelihood of recidivism than eligible, non-participating youth.

In Section I, three research questions were outlined which have been addressed in this study.

1) *Does FFT significantly reduce the likelihood of recidivism?*

[31] Barnoski (2004) and Peterson (2017).
[32] For more information, see Appendix I.
[33] Knoth-Peterson et al. (2022).

No. In fact, youth participating in FFT were *more* likely to recidivate than similar eligible, non-participating youth by 10.1 percentage points on average. Court-involved youth participating in FFT who recidivated were not more or less likely to have recidivated with a felony or violent felony offense relative to eligible, non-participating youth who recidivated.

2) For whom is FFT most effective?

We could not identify any subgroup of youth in this study for whom FFT was beneficial relative to the comparison group in terms of overall recidivism. The relationship between participation in FFT and the likelihood of recidivating did not vary by race, gender, age, or risk level. Across all subgroups, youth who participated in FFT were more likely to recidivate than youth who were eligible but did not participate in FFT.

Again, in keeping with the overall findings, there were no significant differences in the likelihood of recidivating with a felony for youth who recidivated between those who participated in FFT and those who were eligible but did not participate. This result held for all subgroups (race, gender, age, and risk level).

Similarly, for those who recidivated, there was no overall significant difference in the likelihood of recidivating with a violent felony between those who participated in FFT and those who were eligible but did not participate.

However, two groups of recidivists who participated in FFT were less likely than comparison recidivists to commit a violent felony. Youth ages 15 and older and youth who were classified as moderate risk were both less likely to recidivate with a violent felony if they participated in FFT relative to those who didn't participate. In other words, looking specifically at youth who recidivated, older youth and youth deemed at lower risk for recidivating had a lower likelihood of recidivating with a violent felony relative to the comparison group.

3) Under what conditions is FFT most effective?

We did not identify any conditions under which FFT participants had better recidivism outcomes than those in the comparison group. Our analyses did not identify any geographical areas or living situations in which FFT was associated with lower rates of recidivism relative to non-FFT programming.

Although prior evaluations found therapist competence moderates the effect of FFT, our analyses found that even youth working with highly competent therapists experienced higher rates of recidivism than the comparison group.

Our evaluation found youth who complete FFT, or progress further in the model, were less likely to recidivate than those who did not complete or progress as far in the program. However, those youth who completed FFT still recidivated at statistically significantly higher rates than those who did not participate in FFT.

Discussion

Our analyses consistently found that youth who participated in FFT were more likely to be convicted of subsequent offenses than youth who did not participate in FFT, often participating instead in alternative programs. This does not mean that FFT leads to worse outcomes than no treatment at all, but rather that, on average, FFT is not reducing recidivism compared to all the other treatment options available to youth in the court system.

In WSIPP's 2004 evaluation of FFT, we found that FFT is associated with reductions in felony and violent felony recidivism only when delivered competently and otherwise had insignificant findings.[34] In this study, randomization was induced from a waitlist design, and the researchers were able to take advantage of a no-treatment comparison group.

Then, in 2022, WSIPP published a report that, in part, evaluated program effectiveness on recidivism by responses to specific questions on the Positive Achievement Change Tool (PACT) assessment.[35] While not the primary focus of this report, the researchers found that participation in FFT did not result in a therapeutic effect when looking at recidivism generally for male and female youth who participated in FFT.[36] Similar to our study, this report looked at a treatment-as-usual comparison group made up of individuals participating in other rehabilitative programs and used similar statistical methods. In this 2022 study, the researchers found evidence of therapeutic effects for male and female youth after participating in FFT when limiting their analyses to specific answer types on the full assessment.[37]

In the time between those two studies, two key factors changed the juvenile justice landscape; the programs available to youth involved in the juvenile justice system and the youth that participate in it.

First, the programs offered to youth involved in the juvenile justice system have expanded. At the time of the first report, funding was moved from intensive probation to EBPs, specifically FFT and Aggression Replacement Training (ART).[38] In 2007, funding for EBPs was expanded, and additional programs were implemented in juvenile courts.[39]

---

[34] Barnoski (2004).

[35] Knoth-Peterson et al. (2022).

[36] For male youth, FFT participation led to a decrease in recidivism in 0.003% of the model iterations and led to an increase in recidivism in 50.7% of the iterations. For female youth, FFT participation led to a decrease in recidivism in 0% of the model iterations and led to an increase in recidivism in 11.05% of the iterations. The remainder of the iterations were either null, failed to converge, or were inconclusive.

[37] For male youth, the study found that FFT reduced likelihood of recidivism for youth who were 17 or older at their first offense. They also found FFT reduced the likelihood of felony recidivism for male youth who believed they had no control over their own anti-social behavior. For female youth, the study found that participation in FFT lead to reduced likelihoods of felony recidivism for youth who felt close to at least one family member, who had no current attention deficit disorder (ADD)/ attention deficit hyperactivity disorder (ADHD) diagnosis, or who scored in the second lowest quartile on the attitudes/behavior domain.

[38] Coordination of Services (COS) and Multi-Systemic Therapy (MST) were also funded but were less widely adopted.

[39] Family Integrated Transitions (FIT) and Victim Offender Mediation as well as the use of drug courts.

Second, the population of court-involved youth has changed.[40] In a previous WSIPP report, researchers found that between FY 1995 and FY 2014, recidivism rates for youth involved in juvenile courts declined. They also note that declining rates of felony recidivism largely drove the declining rates of recidivism.

As mentioned previously, it is possible that selection bias is still present due to unobservable factors—characteristics not apparent in our data that Juvenile Probation Counselor (JPCs) have in mind when recommending EBPs to youth. It is possible that JPCs are successfully identifying youth who would most benefit from FFT, and those youth are also at the highest risk of recidivism. If this is happening, that would explain the higher predicted rates of recidivism. Aside from the potential of selection bias, there are other alternative explanations for why we might expect participants in FFT to recidivate at a higher rate.

One potential explanation is simply that other programs are more effective than FFT at reducing recidivism for their participants. Given that juvenile court youth today are likely to receive some intervention, a therapeutic effect of FFT could be masked if the average alternative intervention that a comparison youth receives has a larger therapeutic effect.[41]

Another potential explanation is that the youth eligible for FFT are unable to internalize the treatment effectively.[42] As mentioned in the 2022 report, youth eligible for FFT exhibit the highest needs in the current family domain. It is possible that these high needs also come with characteristics that prohibit the treatment of the underlying risks in the family domain, serving as a barrier to effective treatment.

[40] Knoth et al. (2019).

[41] The research team conducted a robustness check, in Appendix II, looking just at youth who did not participate in other EBPs. This robustness check found consistent results with what is reported in the main body of the report. However, we cannot guarantee that the youth who received no other EBP intervention did not receive some alternative local treatment. WSIPP does not have data on participation in non-EBP alternative, local options.

[42] Knoth-Peterson et al. (2022).

Appendices

Washington State's Functional Family Therapy Program: *Outcome Evaluation*

I. Data

This appendix discusses the data sources, processing, and sampling procedures used for the main analysis, as well as additional data processing procedures used for subsequent analyses.

Data Sources, Processing, and Sampling

Information regarding evidence-based program (EBP) referrals and completion is housed in the Juvenile Assessment Research Database (ARD). Criminal history and recidivism data came from WSIPP's Criminal History Database (CHD). The sample was linked to the CHD to create a single data set of all identifiable youth eligible for FFT within our date range. For part of our analysis, we also link to data from FFT, LLC. Additional processing and sampling steps were then taken to address any additional irregularities and to ensure minimal selection bias prior to analysis. Many of the data sources and processing procedures used in this study were identical to procedures used in previous reports.[43]

Juvenile Assessment Research Database (ARD)

The Juvenile Assessment Research Database (ARD), housed and maintained by the Administrative Office of the Courts (AOC), contains information from the actuarial risk assessment tool used in the 33 juvenile courts in Washington State. The risk assessment used by the courts was developed to determine the risk of recidivism, identify those who would most benefit from rehabilitation efforts, and aid in developing case management plans to rehabilitate justice-involved youth. The ARD contains usable assessment information starting in November 2003, and our evaluation used data available as of July 2021.

Individuals may be administered many different types of assessments, including an abridged prescreen, an initial, reassessments, and final assessments. FFT eligibility depends on a youth's score on a full assessment, and eligibility can change between initial assessments and reassessments. However, since reassessments were not consistently administered, using eligibility tied to results from a reassessment may bias our sample. Additionally, final assessments were administered prior to release from community supervision and would not be used to determine program eligibility or placement. Consequently, *we limited our sample of observations to only initial assessments and reassessments for a new offense.*

[43] Ibid.

We selected all FFT eligibility or participation records corresponding to initial assessments or reassessments for new offenses in our sample. During our sample time period, courts were able to "turn off" program eligibility for programs that were not offered in their court. To account for this missingness, we also manually calculated eligibility using assessment data and eligibility criteria.

We omitted trips if the reason a youth did not start a program indicated that they were systematically different from those in the program. Specifically, we omitted those who were "awaiting or involved in in-patient drug treatment," "committed to JRA," "deceased," "incarcerated," had "whereabouts unknown," "on warrant status," or who "moved or is moving out of state." These are individuals that would be a poor comparison for the treatment group. For example, those who are deceased are unable to recidivate, so including them in the comparison group would lead to bias if the comparison group consists of individuals who cannot recidivate. We then assigned individuals to the treatment group if they started FFT and assigned individuals to the comparison group if they did not start FFT.

Multiple Cases During the Study Period

Individuals may have multiple, distinct criminal cases for which they were potentially eligible for FFT. There were many instances where the same person had a case for which they participated in FFT and a separate case for which they were eligible but did not participate in FFT. As a result, an individual could appear multiple times in both the treatment and comparison groups.

To address the dependence of cases within our sample arising from this issue, we first removed records from the comparison group if they ever participated in FFT (5,340 records). This eliminates the dependence between treatment and comparison groups and ensures that our comparison group is limited to those who never participated in FFT during the sample time period. Second, to address dependence within the treatment or comparison groups, we randomly selected one record for each individual with multiple records. This process of random selection was repeated 100 times to test the sensitivity of results due to record selection.

In addition to assessment data, the ARD contains separate databases on EBP eligibility, program start and completion dates, and reasons for non-starts and non-completions. It also contains demographic information such as person characteristics (first name, last name, date of birth, sex, and race), court information, and youth identifier variables for linking with criminal history data.

Criminal History Database (CHD)

WSIPP's Criminal History Database (CHD) combines data from several Washington State agencies: court data from the Administrative Office of the Courts, residential confinement data from Juvenile Rehabilitation at the Department of Social and Health Services, and incarceration in state prisons and community supervision data from the Department of Corrections. The CHD allows researchers to create criminal history and recidivism measures for all justice-involved youth in the state. WSIPP updates the CHD quarterly, and our analyses used information from the CHD up to March 2023.

Consistent with standards previously published by WSIPP, we measured recidivism as any criminal offense committed following the youth's at-risk date that resulted in a court legal action (i.e., conviction, diversion, or deferred disposition) in a Washington State court using an 18-month follow-up period and a 12-month adjudication period. This means that, for youth to be in our dataset, we made sure that we had at least 18 months of data while they are considered "at-risk" in the community and at-risk for recidivating, and then allowed an additional 12-month lag to allow for the court legal system to adjudicate any arrests. This time

lag allows for our data to have enough months of youth on supervision to capture recidivism events and for those events to be accurately reflected in the data we collect.
For this evaluation, we measured types of recidivism based on the most serious offense in the first recidivism event during the follow-up period. Previous evaluations measured recidivism as the most serious offense type of all recidivism events during the follow-up period. By selecting the first recidivism event during the follow-up period, we are restricting our measurement to the recidivism event that is most directly related to treatment status.

For youth who participated in FFT, the at-risk date was the FFT start date. We constructed an equivalent at-risk date to compare youth who did not participate in FFT. If we were to measure recidivism for the comparison group beginning at the adjudication date, we would potentially include recidivism events for the comparison group that would not be considered for similar youth in the treatment group. That is, if a youth in the treatment group committed an offense prior to starting FFT, it would not be identified as a recidivism event. To ensure that we are not introducing bias in the comparison group, we calculated the average time between the adjudication date and the start of FFT for youth in the treatment group. We used this average length of time to FFT participation as the lag time between adjudication and the at-risk date for the comparison group. In our analyses, this meant the at-risk date for the comparison group was 51 days after the assessment date.

ARD/CHD Compiled Dataset

ARD and CHD records can be reliably associated at the youth level using the Judicial Information System (JIS) and Juvenile and Correction System (JCS) numbers. However, both data sources can include multiple records for a given youth—e.g., the ARD can include multiple assessments for a given case or separate cases, and the CHD can include multiple cases over time. Unfortunately, there is no court case identifier to link records between these two data sources (there is a court case number in the CHD but not the ARD). We use the same method we used in previous reports to merge these datasets by linking ARD assessment dates and CHD adjudication dates.[44]

Other Data Exclusions

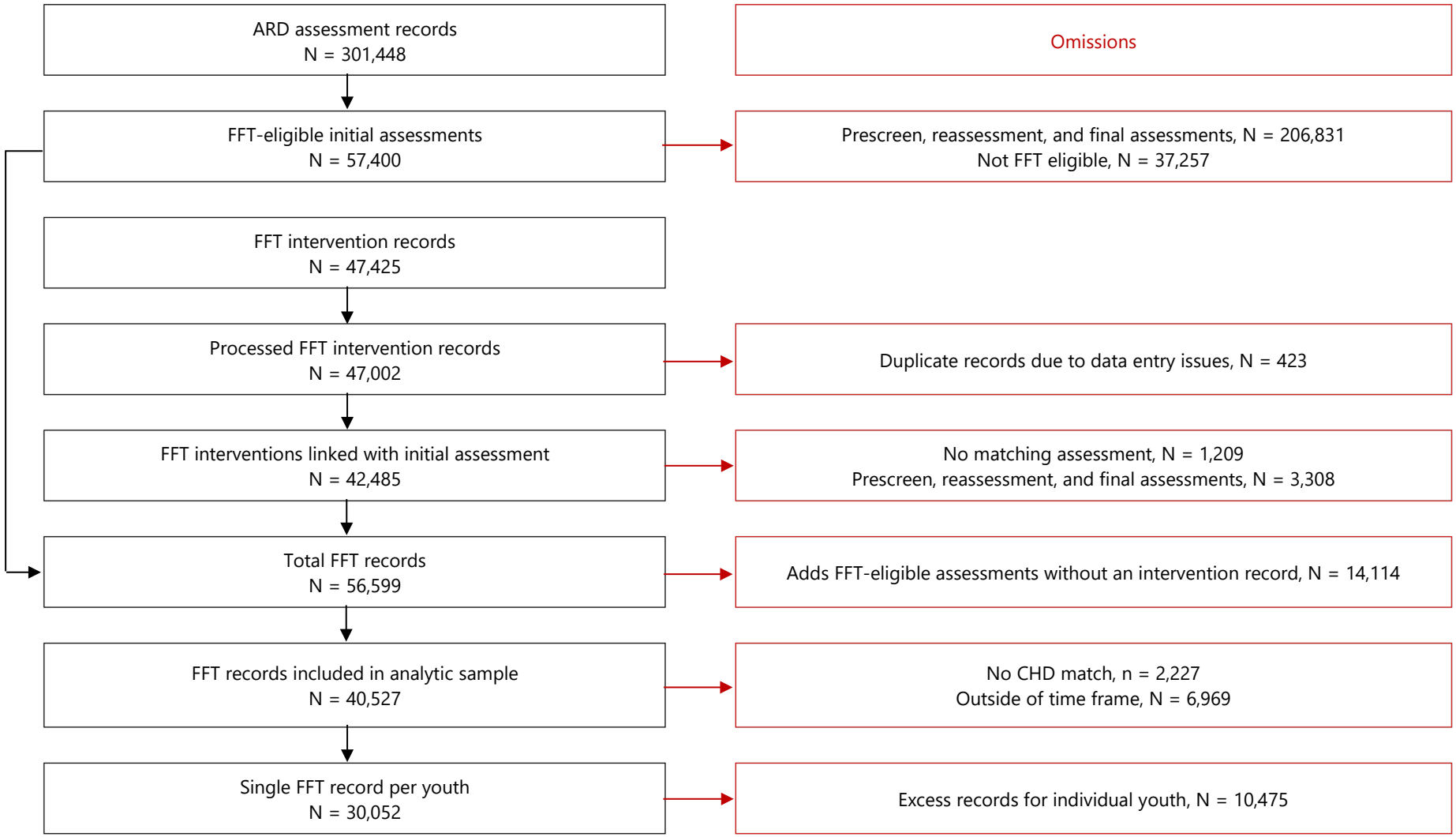
While coding the data for analyses, we made additional selection decisions that removed some youth from the treatment and comparison groups.[45] First, we excluded youth who were listed as starting FFT but for whom we did not have an FFT start date available in the data. Second, we removed youth under 10 or greater than 21 years old on the date they were assessed.[46] Third, we removed youth for whom necessary independent variables were missing. Specifically, we removed youth who were missing information about the index offense characteristics and youth who were missing risk assessment information.

Exhibit A1 walks through these data processing steps. Additional details can be found in Appendix II in *What works for whom? Juvenile court assessment tool and program eligibility*.[47]

[44] Knoth-Peterson et al. (2022).
[45] Further information regarding these data exclusion decisions are available upon request.
[46] During the timeframe of the sample included in our evaluation, Washington State juvenile courts were able to maintain jurisdiction of youth up to age 21. Thus, it is possible that some youth were assessed and assigned to an EBP up to age 21.
[47] Knoth-Peterson et al. (2022).

Exhibit A1

ARD Date Processing Procedure



FFT, LLC Data

FFT, LLC provided WSIPP with client session, therapist, and outcome data from its case management system for court-involved youth who participated in FFT from 2011 onward. Therapists working with FFT, LLC receive information for court-involved youth from a youth's Juvenile Probation Counselor (JPC) via paper or electronic form(s). These include demographic information, case information, and family contact information. This information is input by the therapist in a case management system internal to FFT, LLC, called CSS. This system is where all case information is stored, including youth and family details, session dates and locations, assessment information, and the therapist's session notes. Supervisors will check in with therapists weekly and, using the CSS, provide scores for fidelity and dissemination adherence.[48]

Therapist Competency Analysis. Therapists receive weekly fidelity and dissemination adherence scores from their supervisors while they are staffing a case. From discussions with individuals at FFT, LLC, the research team decided to use the fidelity scores as the basis from which to create the competency scores for therapists.

For each youth's unique trip through FFT, there is a first and last recorded session date associated with a unique therapist. To measure program fidelity for a given youth trip through FFT, we used the youth's first and last session dates plus a 6-month bandwidth, then took an unweighted average over all fidelity scores for the corresponding therapist during that time period.[49] The average fidelity score for that therapist and time period was saved as the youth's therapist's "score" for their trip through FFT.

Once the scores were constructed at the youth level, we created three levels of competency to perform our analyses; highly competent (4.5 and greater), competent (from 3 up to 4.5), and less competent (less than 3).

Phase Completion Analysis. Youth session information is recorded in the CSS by the therapist. Youth session information includes the location, the date it occurred, and what phase of the program they were in during that session. Due to the overlapping nature of the program model phases, the five phases are consolidated into three more distinct phase categories tracked in the CSS; engagement/motivation, behavior change, and generalization. These are the categories we used to conduct the analysis of phase completion.

Linking the FFT, LLC Data to the ARD/CHD Compiled Dataset

We linked cases from the FFT, LLC data to the ARD/CHD compiled data. Because a unique identifier does not exist to link these databases reliably, this linkage resulted in a smaller dataset of individuals who were linked to all three original datasets.[50] This smaller dataset was used for a portion of the analysis done in Section IV.[51]

---

[48] FFT, LLC supervisors use both a therapist's fidelity and dissemination adherence scores to measure progress and overall adherence to the FFT program model. For purposes of this study, we chose to focus on the fidelity measure as that most closely measures the level to which the therapist adhered to the FFT program model within the youth's session time.

[49] The bandwidth is desirable to account for the fact that some therapists were not rated during a youth's entire trip and prevents losing observations as a result.

[50] The FFT, LLC data was matched to the ARD data on JCS number when available, youth name, estimated year of birth from recorded age, court, completed status, and program start date and/or case open date.

[51] This smaller dataset (N = 2,478) was used to perform the analyses related to phase completion and therapist competency, presented in Section IV.

II. Recidivism Analysis

In an ideal outcome evaluation, with randomization into treatment and control groups, we could confidently conclude that participation in FFT causes different outcomes. Unfortunately, we cannot make this claim in this report because individuals are not randomly assigned to participate in FFT. There are many reasons why an individual might end up in the treatment program, and those reasons are not always random. Individuals must meet the existing eligibility requirements as well as self-select into the program. If individuals who are less likely to recidivate are also more likely to participate in the program, then at least part of an estimated reduction in recidivism for the treatment group cannot be attributed to the program itself.

Therefore, we use an empirical approach to remove as many dissimilarities as possible between the treatment and comparison groups. Our preferred approach does this in two separate steps: we preprocess the data using entropy balancing, and then we control for other variables in a regression analysis. We follow this process for each subgroup and outcome in our analysis. We then repeat this process 100 times following the randomization procedure described in Appendix I to ensure the robustness of our results.

Entropy Balancing

Entropy balancing is a data preprocessing method that weights observations in the comparison group. When these weights are applied, covariates in the treatment and comparison groups will have similar distributions (i.e., the covariates are "balanced").[52] This weighting is done so that every covariate of interest is balanced simultaneously.

For our analyses, we entropy balanced on the following variables:

- Gender
- Race
- Age & Age^2
- Crime category (person, property, sex, drug, and other)
- Crime grade (misdemeanor, felony, violent felony)
- Disposition
- Assessment version
- Risk level category
- Any prior EBP participation
- Any other EBP eligibility
- Assessment year[53]
- Domain scores[54]
- Current living arrangement domain items[55]

[52] For a full discussion of this method, see Hainmueller, J. (2012). Entropy balancing for casual effects: a multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis, 20,* 25-46.
[53] ARD models include BOT and PACT domain scores while FFT models include only PACT domain scores, thus assessment version is not balanced on for the FFT therapist quality (fidelity score) regression and dosage (phase) models.
[54] Domain scores include static/dynamic risk/protective scores. Domain scores are numeric variables unless the number of available scores is less than five, in which case they are categorical variables.
[55] Current living arrangement item responses were consolidated where appropriate.

If entropy balancing succeeds, we are guaranteed that the included covariates balance across the treatment and comparison groups, so differences across these variables will not bias our overall results. The entropy balancing step can fail when it is not possible to balance the requested covariates. This can happen when there is no way to weight observations to get similar distributions, as in the situation where everyone in the treatment group has a certain characteristic, and everyone in the comparison group does not. The only subgroup and outcome combination with a convergence rate of less than 100% was when we performed analyses specifically within the group of recidivists for those under 15 years of age. This pattern is likely explained by the decrease in sample size.

In this report, we test the effectiveness of FFT on several subgroups. We calculate entropy balancing weights for every subsample of individuals so that the covariates are balanced within each subsample.

For each subsample of individuals, we run two different entropy balancing models. The first is run on the entire sample. The second is run on the subset of those individuals who did recidivate. This second entropy balancing model is used when examining changes in the share of recidivism, that is, felony recidivism or violent felony recidivism. In all subgroups and outcomes, we redo the entropy balancing step for each iteration to account for differences across sampling randomizations.

Regression Analysis

Next, we run a regression analysis. In this stage, we can control for additional covariates that we were unable to balance in the entropy balancing stage.

We estimate the following regression equation:

$$P(Recid_i = 1) = \alpha + \beta_1 Treat_i + \boldsymbol{\gamma X_i} + \delta_t + \theta_c$$

$Recid_i$ is a binary variable that is one if individual i recidivated during the period of interest, $Treat_i$ is a binary variable that is one if an individual i participated in the program, $\boldsymbol{X_i}$ is a vector of individual control variables, δ_t are year-fixed effects and θ_c are court-fixed effects.[56] Observations are weighted on the entropy balance weights calculated in the previous step.

We estimate a linear probability model (LPM) because the dependent variable of recidivism is binary. We selected the LPM over a non-linear model such as a logistic model mainly because of the high number of control variables and the often-small sample sizes. In practice, these logit models did not consistently successfully converge. We use robust standard errors to account for heteroskedasticity.

In an LPM, coefficients can be interpreted as percentage point changes. For example, if the program reduced the probability of recidivism from 40% to 30%, this 10-percentage point decrease would show up as an estimated coefficient of -0.10.

Exhibit A2 displays a summary of the repeated analyses. Displayed are the number of runs per subgroup that exhibited therapeutic (FFT participation led to a decrease in the likelihood of the recidivism outcome), iatrogenic (FFT participation led to an increase in the likelihood of the recidivism outcome), and null (statistically insignificant) results.

[56] The control variables used were *court, prior participation in individual EBPs, eligibility for other individual EBPs,* and *follow-up participation in other individual EBPs.*

Exhibit A2

Summary of Repeated Subgroup Analyses

Subgroup characteristic	Any recidivism			Felony recidivism			Violent felony recidivism		
	Therapeutic	Null	Iatrogenic	Therapeutic	Null	Iatrogenic	Therapeutic	Null	Iatrogenic
Full sample									
Full sample	0	0	100	0	100	0	7	93	0
Gender									
Female	0	0	100	0	100	0	13	87	0
Male	0	0	100	0	100	0	0	100	0
Race/ethnicity									
White	0	0	100	0	100	0	20	80	0
Black	0	0	100	0	100	0	0	100	0
Hispanic/Latino	0	0	100	0	100	0	0	100	0
Age									
Under 15	0	0	100	0	64	1	0	65	0
15 or older	0	0	100	1	99	0	98	2	0
Assessment version									
WAJCA	0	0	100	0	100	0	0	100	0
PACT	0	0	100	0	100	0	1	99	0
Risk level category									
Moderate risk	0	0	100	0	100	0	79	21	0
High risk	0	0	100	0	100	0	0	100	0
East-rural	0	0	100	0	100	0	0	100	0
East-urban	0	0	100	4	96	0	11	89	0
West-rural	0	0	100	14	86	0	14	86	0
West-urban	0	0	100	0	100	0	0	100	0
Living arrangements									
Living with parents	0	0	100	0	100	0	36	64	0
Not living with parents	0	0	100	0	100	0	0	100	0

Testing the Sensitivity of Results

By Varying the At-Risk Date

We tested the sensitivity of our results to the choice of at-risk date. To do this, we replicated our main analysis on the full set of youth and changed the at-risk date for the treatment and comparison groups to their assessment date. The results remained consistent with those presented in the report.[57]

By Varying the Participation in Other EBPs

Some youth eligible for FFT complete additional EBPs during the follow-up period. If individuals in the treatment group (who started FFT) and the comparison group (who were eligible for but did not start FFT) are receiving other kinds of treatment during the follow-up period, it is possible that this participation in alternative programs is influencing the rate of recidivism. To look for this kind of effect, we performed a robustness check replicating our main analysis on the subsample of individuals who did not participate in other EBPs during the follow-up period.

Our results remained consistent with those presented in the main body of the report.[58] Youth who participated in FFT were more likely to recidivate than similar youth who did not participate in FFT. In addition, the group of recidivists that participated in FFT were not differentially likely to recidivate with a felony or violent felony.

By Varying the Sampling Process

In our main analyses, any youth who ever participated in FFT was removed from the comparison group. This eliminated the dependence between treatment and comparison groups to ensure that our comparison group was limited to those who never had FFT exposure during the sample period.

However, some youth are eligible for FFT multiple times, sometimes participating in FFT and sometimes not participating. These individuals may be systematically different from those who never participate in FFT, and by limiting our sampling such that these individuals who switch back and forth can only ever be in our treatment group, we may be introducing bias to our sample.

In this robustness check, we allow youth who are eligible multiple times to go back and forth between the treatment and comparison groups as their trips through the juvenile court system get randomly selected. With this, we are allowing our estimates to be more unstable in order to test for the possible presence of sampling bias.

Our results remained consistent with those presented in the main body of the report.[59] Youth who participated in FFT were more likely to recidivate than similar youth who did not participate in FFT. In addition, the group of recidivists that participated in FFT were not differentially likely to recidivate with a felony or violent felony.

[57] Additional details are available upon request.
[58] Ibid.
[59] Ibid. The direction of results remained consistent throughout the analyses, however the magnitude and significance of the estimated effect size varied.

Acknowledgments

The authors would like to thank staff at the Administrative Office of the Courts (AOC), the Department of Children, Youth, and Families (DCYF), and FFT, LLC for helping us obtain the data needed to conduct this evaluation. We are particularly grateful to Cory Redman, David Sattler, Karen Brown, Helen Midouhas, and Michael Robbins for answering our questions regarding juvenile court operations and the FFT, LLC program model. We specifically want to thank Karen Brown for her assistance in preparing the FFT, LLC data for use in our analyses. Finally, we are thankful to the CJAA Advisory Committee for their feedback throughout the study.

For further information, contact:
Morgan Spangler at 360.664.9807, morgan.spangler@wsipp.wa.gov

Document No. 23-06-3901