Improving Outcomes for Children Exposed to Violence

Safe Start Promising Approaches

Appendix I. Worcester, Massachusetts: Intervention, Study, and Results

Dana Schultz, Lisa H. Jaycox, Lynsay Ayer, Claude Messan Setodji, Ammarah Mahmud, Aaron Kofner, Dionne Barnes-Proby

For more information on this publication, visit www.rand.org/t/RR1728

Published by the RAND Corporation, Santa Monica, Calif.
© Copyright 2017 RAND Corporation

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Support RAND

Make a tax-deductible charitable contribution at

www.rand.org/giving/contribute

www.rand.org

Appendix I. Worcester, Massachusetts: Intervention, Study, and Results

Introduction

The Worcester Safe Start program was a partnership between the American Institutes for Research's National Center on Family Homelessness (NCFH), the Central Massachusetts Housing Alliance (CMHA), and the University of Maryland's Family Informed Trauma Treatment Center. The research project overall, project management, and portions of the intervention were run through NCFH, and CMHA and NCFH staff implemented the primary program component. The university center provided training and ongoing supervision for the group therapy component of the program.

The Worcester Safe Start program was located in Worcester County, Massachusetts, with the comparison shelters based in Framingham. According to the U.S. Census Bureau's 2014 American Community Survey (U.S. Census Bureau, 2016), this catchment area had a combined population of 876,704 residents, 23 percent of whom were younger than age 18 and 6 percent of whom were younger than age 5. Approximately 79 percent of the population were white, 4 percent were black, 10 percent were Hispanic, and 7 percent were Native American, Asian, or other. The 2014 median household income was $81,631, and about 8 percent of the population were living at or below the poverty threshold. Of families with children under the age of 18, 14 percent were living at or below the poverty threshold. In 2014, the violent crime rate in the city of Worcester was 498.2 per 100,000 residents, compared with the national average of 202.6 per 100,000 (City-Data.com).

The Worcester Safe Start program, Shelter-Based, Trauma-Informed Assessment, Referral, and Treatment (START with Kids), was designed as a comprehensive approach to providing trauma-informed, clinically driven care for children in homeless shelters. Project components included (1) training on trauma and trauma-informed care for children, (2) child-specific assessments and service plans, (3) SFCR group intervention (Kiser, Donohue, et al., 2010), and (4) clinical support and consultation. This intervention was designed enhance homeless service providers' capacity to assess, refer, and support the children in their programs; expand collaboration among the shelters within the system and community; adapt and implement a clinically based intervention; and solidify best practices for serving children within the homeless service system in Worcester and beyond.

The outcome evaluation reported here presents data relevant to the question of whether the Worcester Safe Start program, as implemented within this project, improved outcomes for CEV, compared with usual care in the shelter.

Worcester Safe Start

- **Intervention components:** Child assessment; child-focused service plan; SFCR group sessions (optional)
- **Intervention length:** two sessions over two weeks for the child assessment and child-focused service plan; ten weeks for the SFCR group therapy sessions
- **Intervention setting:** Homeless shelter
- **Target population:** CEV
- **Age range, in years:** 0–18
- **Primary referral sources:** Village at Cambridge Street and Friendly House

Intervention Summary

Referrals

As shown in Table I.1, the Worcester Safe Start program received the majority of its study referrals from shelters in the Worcester emergency shelter system, particularly the Village at Cambridge Street at CMHA. Not surprisingly, the larger shelters contributed the most referrals into the study, with the Village at Cambridge Street accounting for 43 percent of referrals and Friendly House accounting for 27 percent. Outreach activities included conducting interviews with shelter residents to elicit feedback on engagement strategies. Families were referred when they entered the shelter and were working on their rehousing plans (usually within four weeks of entering the shelter). All families were offered the opportunity to participate in the study. Within two weeks of a family agreeing to participate, a program staff person conducted the baseline study assessment with the family.

Table I.1. Worcester Safe Start Referral Sources

Referral Source	Shelter Capacity or Capacity Range	Number of Referrals	Percentage of Total (*n* = 345)
Village at Cambridge Street	60–71	148	43
Friendly House	36	93	27
Henry Lee Willis Community Center	20–39	43	12
Catholic Charities of Worcester County	12	34	10
South Middlesex Opportunity Council[a]	39	27	8

[a] The site added this agency with several emergency shelters partway through the study.

Intervention Components

The Worcester Safe Start program for CEV ages 0–18 consisted of child assessments, child-focused service plans, and SFCR group therapy sessions. The program with these components, START with Kids, was delivered by shelter case managers (for the child assessments and child-

focused service plans) and a clinical child specialist hired for the project (for the SFCR group therapy sessions and some child assessments and service plans).

Child Assessments

Approximately seven to eight weeks after a family entered the shelter, shelter case managers conducted a comprehensive child assessment to identify necessary child-specific services and supports. The child assessments were typically completed in two sessions conducted over a two-week period, taking about two to three hours to complete. Once completed, the child assessment was included with the family's rehousing plan. NCFH developed the child assessment by identifying and integrating existing child-level assessment questions and tools. The child assessments focused on the areas of physical health, the caregiver–child relationship, development, social–emotional health and behavior, caregiver stress, and family functioning. Within each area, the assessment included questions to assess past and current status, stressors that might be affecting the area, strengths, challenges, and plans for addressing.

Child-Focused Service Plans

The shelter case manager created an individualized child-focused service plan based on the child assessment to address the needs of each child in the family, as well as the overall family unit. The child-focused service plan was meant to connect children to needed community-based services (e.g., trauma-based clinical services, health services, EHS, education supports). The child-focused service plan included specific referrals needed according to the assessment and provided a place to track referrals and follow-up contacts. The child-focused service plans were typically completed in two sessions over a two-week period, taking about two to three hours to complete. After the plan was complete, the shelter case manager reviewed the child-focused service plan with the family during weekly meetings.

Strengthening Family Coping Resources Group Sessions

For the optional group therapy component, Worcester Safe Start implemented the ten-week, closed-group version of SFCR (Kiser, Backer, et al., 2015; Kiser, Donohue, et al., 2010). This multifamily group intervention is designed to help families reestablish routines and rituals, enhance safety and coping, and manage stress. SFCR is usually a 15-week closed group that includes sessions that explore traumatic experiences within the family and allow for processing these experiences. Because of the shelter setting, Worcester decided to implement the ten-week version of SFCR. The shorter version of SFCR focuses on family coping-skill development and does not include the sessions related to traumatic experiences. The ten-week version of SFCR consists of ten two-hour sessions over a ten-week period. Groups are flexible and are designed to accommodate families who leave early or join in-progress groups.

Design Overview

The design of this study was a quasi-experimental design, with the comparison group drawn from other family shelters within the Worcester emergency shelter system. At the beginning of the study, there were four homeless shelters in the system with initial capacity ranging from 12 to 71. A fifth agency was added partway through the study. The largest shelter started as the first intervention group shelter, while the others recruited families for the comparison group. The original plan was to convert the other shelters to intervention shelters over the course of the study. However, because of issues with recruitment in the comparison shelters, the largest shelter remained the only intervention shelter throughout the study.

Families in the intervention shelter received the child assessment and child-focused service plan and were offered the opportunity to participate in the SFCR group sessions. Families in the comparison group received the normal services and supports that the shelter offered. For both groups, we assessed child outcomes and contextual information at baseline, six, and 12 months. Chapter Two summarizes and Appendix A fully describes the measures used in this study. The measures were uniform across the national evaluation but prioritized within each site as to the relevance to the intervention under study. Given the nature of the Worcester Safe Start intervention, the outcomes were prioritized as shown in Table I.2.

Table I.2. Worcester Prioritized Outcomes

Domain	Source or Measure	Child's Age, in Years	Respondent
Primary outcome measures			
Social and emotional competence	ASQ:SE	0–3	Caregiver
	SSIS assertion	3–12	Caregiver
	SSIS self-control	3–12	Caregiver
Secondary outcome measures			
Behavior and conduct problems	BITSEA problem scale	1–3	Caregiver
	Child behavior problems (BPI Total Scores and BPI Externalizing subscale)	3–18	Caregiver
Social and emotional competence	SSIS cooperation	3–12	Caregiver
		13–18	Child
	SSIS assertion	13–18	Child
	SSIS self-control	13–18	Child
	BERS-2 Affective Strength subscale	6–12	Caregiver
		11–18	Child
School behavior and attitudes	BERS-2 School Functioning subscale	6–12	Caregiver
		11–18	Child

Domain	Source or Measure	Child's Age, in Years	Respondent
Family functioning	BERS-2 Family Involvement subscale	6–12	Caregiver
		11–18	Child
	APQ positive involvement, negative or ineffective discipline, and deficient monitoring	6–18	Caregiver
	APQ parental involvement, positive parenting, poor monitoring and supervision, inconsistent discipline, and corporal punishment	8–18	Child
	FES Conflict scale	0–18	Caregiver
		11–18	Child
Tertiary outcome measures			
Background and contextual	Attitudinal Barriers to Care	0–18	Caregiver
	ESI total stressors, resource problems, and personal problems	0–18	Caregiver
Child PTSD symptoms	TSCYC PTSD scale	3–10	Caregiver
	CPSS total PTSD symptoms, reexperiencing symptoms, avoidance symptoms, and arousal symptoms	8–18	Child
Behavior and conduct problems	Delinquency	11–18	Child
	Substance use	11–18	Child
	Gangs	11–18	Child
Child depression	RADS	13–18	Child
	BPI internalizing problems	3–18	Caregiver
School behavior and attitudes	Self-reported grades	13–18	Child
Violence exposure	JVQ child victimization experiences (total, child maltreatment, child assault, child sexual abuse, and child witnessing violence)	0–11	Caregiver
		10–18	Child
	Caregiver victimization (total, DV, and non-DV)	0–18	Caregiver
Caregiver mental health	PHQ-8 depression scale	0–18	Caregiver
	PC-PTSD	0–18	Caregiver

Study enrollment took place for three years between July 2012 and June 2015. From the beginning, the Worcester Safe Start program had difficulty tracking families as they left the shelters, which resulted in poor retention rates for the 12-month study assessment. As a result, the 12-month study assessments were discontinued in November 2014.

Analysis Plan and Power Calculations

We conducted descriptive analyses to summarize the sample baseline characteristics: age, gender, race or ethnicity, family income level, child's violence exposure, and the outcome variables. Because this was a quasi-experimental design, we wanted to determine whether there were any major differences between the two groups at baseline. To do this, we tested for differences in child and caregiver characteristics between intervention and comparison group

children using *t*-tests and chi-squared tests. Because the retention rate was so different between the two groups, we also examined whether those who were lost to follow-up differed in any systematic way from those who were retained, using *t*-tests and chi-squared tests.

To assess the effect of the Safe Start intervention, we primarily examined differences between children in the intervention and comparison groups at six and 12 months postbaseline. It is important to consider the power this study has for such an analysis. One way to describe power is by using the effect size difference between the two groups being compared. The effect size is a standardized measure of the strength of association between an intervention and an outcome and is defined as the average difference in an outcome between the intervention and comparison groups divided by the common SE. The effect size measure is commonly classified as small if it is about 0.2, medium if it is about 0.5, and large if it is about 0.8 (Jacob Cohen, 1988).

The START for Kids model had not been studied previously, with no prior estimates of its intervention effect. Children of families in the intervention shelters were to receive child-focused assessments and service plans, both of unknown effect. Families also could participate in the SFCR group therapy sessions of unknown effect. Given its multiple components and uncertainty about what kind of effect to expect for this site, we estimated a medium effect for this intervention. Table I.3 shows the enrollment by group, comparing the actual enrollment with the target enrollment needed for power, assuming an 80-percent retention rate and 95-percent comparability between the intervention and comparison groups. With total enrollment of 345, Worcester enrolled 132 percent of the sample size required to detect a medium intervention effect. Worcester's overall six-month retention rate of 62 percent for both groups meant that it retained a total of 214 families in the study, representing 102 percent of the retained sample required to have an 80-percent chance of detecting a medium intervention effect (assuming 95-percent comparability between the intervention and comparison groups). As noted above, Worcester stopped collecting 12-month assessment partway through the data collection period because of difficulty tracking families after they left the shelters. At 12 months, Worcester retained 52 families in the study, representing 25 percent of the retained sample needed to have an 80-percent change of detecting a medium intervention effect, excluding the later enrollees for whom a 12-month assessment was not planned. With those retention rates, the power analysis indicated that we can expect a 95-percent chance at six months and a 42-percent chance at 12 months to detect a medium effect. Given the sample size here, there was sufficient power to detect a small to medium intervention effect of size 0.39 at six months and a large intervention effect of size 0.79 at 12 months, according to Cohen's 1988 effect size classification.

Table I.3. Worcester Required Versus Actual Enrollment for a Medium Effect Size

Requirement	Intervention Group	Comparison Group	Total
Enrolled sample needed for power	131	131	262
Total enrollment	149	196	345
Percentage of needed enrollment	114	150	132
Retained sample needed for power	105	105	210
Retained sample, six months	93	121	214
Percentage of needed retention, six months	89	115	102
Retained sample, 12 months	26	26	52
Percentage of needed retention, 12 months	25	25	25

Several factors other than overall sample size dampened statistical power. The range of children's ages meant that the full data were not available for some measures because not all children were in the age range eligible to complete that measure. Further, the corrections for the multiple statistical tests being conducted also reduced power.

We examined differences between the intervention and comparison groups using an intent-to-treat approach, which includes in analyses all those assigned to the intervention group, regardless of the amount of services received. As discussed in Chapter Two, comparisons between a comparison group and only those who complete services (or receive a predetermined amount of services) are likely to provide biased results given that those who do not engage in services or drop out prior to completion might differ systematically from those who remain. Ideally, analyses would take into account the type and amount of services received to account for dosage variability. We explore this issue of dose of intervention as described below.

To examine differences between the intervention and comparison groups using the intent-to-treat approach described above, we present baseline and follow-up estimates of primary, secondary, and tertiary outcomes for both groups when the sample size is greater than or equal to five. We compare groups via chi-squared or *t*-tests at each time point, compare means within groups across time, and examine difference in differences to compare the two groups on changes over time between baseline and the six-month and 12-month assessments (when the sample size is at least ten per group). At the six-month follow-up, we conducted multiple linear regressions on the continuous outcomes and linear probability regressions on the dichotomous outcomes to test for the difference in differences via main effects and the interaction between intervention status and time after controlling for baseline characteristics (child age, child gender, child race and ethnicity, and child's exposure to violence). We selected these baseline characteristics to correct for any potential imbalance in the groups by relevant demographic characteristics. We present the adjusted models when the sample size is less than 20 per group. At the 12-month follow-up, we conducted multiple linear regressions on the continuous outcomes and linear probability regressions on the dichotomous outcomes to test for the difference in differences via

main effects and the interaction between intervention status and time after controlling for those same baseline characteristics, as well as six-month violence exposure.

To examine outcomes related to the as-treated sample, those families who took part in the intervention services offered, we examined the outcome means for families who took part in the intervention services offered, broken down into two groups with one group consisting of those who participated in any of the three intervention components (child assessment, child-focused service plan, or SFCR group therapy) and the other group those who had not participated in any of the Safe Start intervention services at six months. Because children with more need are likely to receive more services, we would expect this selection scheme to possibly present an unforeseeable bias, with families more in need receiving more services. To account for this selection bias related to service dosage, we used the propensity score matching method to pair families in each dosage group with families with similar needs in the comparison group. The matching paired families based on similar baseline scores on the outcome measure of interest. The analyses then examined the difference in differences between the intervention and comparison groups for each dosage group at follow-up. Note that, in this analysis, the full comparison group is used in the matching of each of the dosage levels. We examine only primary outcomes with this method, in recognition that it is exploratory and preliminary. Further, the power for the dosage analyses is reduced because of the smallness of the samples for the different dosage groups.

When conducting large numbers of simultaneous hypothesis tests, as we did in this study, it is important to account for the possibility that some results will achieve statistical significance simply by chance. The use of a traditional 95-percent confidence interval, for example, will result in one out of 20 comparisons achieving statistical significance as a result of random error. We therefore adjusted for false positives using the FDR method (Benjamini and Hochberg, 1995). We based our assessments of statistical significance on applying the FDR procedure separately to the primary, secondary, and tertiary outcome tests in this report (as reported in Tables I.9 and I.10) using an FDR of 0.05. In the discussion of results, we have also identified significant trends in the data, defined as those tests with *p*-values of less than 0.05 without adjusting for multiple significance tests. Although these results might suggest a practical difference that would be statistically significant with a larger sample, they must be interpreted with caution because we cannot rule out that the difference was due to chance because of the multiple significance tests being conducted.

Study Results

Enrollment and Retention

As noted above, Worcester enrolled a total of 345 families in the study, with 149 in the intervention group and 196 in the comparison group. In Table I.4, we present the number and

percentage of all enrollees who were eligible for participation at each data collection time point. As shown, 62 percent of families enrolled in both groups were retained for the six-month caregiver assessment. Retention for the caregiver assessment at 12 months was 46 percent for the intervention group and 38 percent for the comparison group, taking into account those families for whom a 12-month assessment was not planned.

Table I.4. Retention of Enrollees Eligible to Participate in Assessments at Each Time Point

	Caregiver Assessment		Child Assessment	
Enrollees	**Six Months**	**12 Months**	**Six Months**	**12 Months**
Intervention				
Received	93	26	29	12
Expected	149	57	48	20
Retention rate, as a percentage	62	46	60	60
Comparison				
Received	121	26	24	5
Expected	196	68	39	14
Retention rate, as a percentage	62	38	62	36

Worcester's relatively low retention at the follow-up assessments increases the potential for biased results. Low retention overall might be related to treatment factors that lead to selection bias causing the characteristics of the study sample to shift. For example, if families in more distress are more likely to leave the study and be lost to follow-up, the results can be misleading and would allow us to generalize only to families with less distress.

Baseline Descriptive Statistics

For the descriptive statistics, we provide the characteristics for the full sample enrolled at baseline. As shown in Table I.5, children who participated in the study were, on average, 5.5 years old (range 0–18), with similar numbers of males (52 percent) and females (48 percent). The racial and ethnic background of families was largely minority (52 percent Hispanic, 18 percent black, and 16 percent other or mixed race), with 15 percent identifying as white. The vast majority of families (86 percent) reported very low family incomes, less than $10,000 per year. Families reported an average exposure to just under one type of violence in the past six months. We observed one statistically significant difference between the intervention and comparison groups at baseline, with the average age in the intervention group 6.3 years old compared with 4.8 years old for the comparison group ($p < 0.01$). The difference between groups in child gender approached significance ($p = 0.05$), with intervention group children more likely to be male.

In the sample of families who were retained in the study at six months, there was the same statistically significant difference between the groups, with children in the intervention group about 1.5 years older than children in the comparison group ($p = 0.01$; data not shown). There were no significant differences between the intervention and comparison groups at 12 months (data not shown). Comparison of those who were lost to follow-up and those who were retained revealed one significant difference between the two groups at six months, with children in the comparison group reporting more violence exposure (an average of 1.2 events) in the past six months than those who did not (an average of 0.8 events) ($p = 0.01$; data not shown).

Table I.5. Worcester Safe Start Baseline Sample Characteristics

Characteristic	Combined		Intervention		Comparison		Test for Comparison *p*-Value
	N	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)	
Child							
Age	345	5.5 (4.4)	149	6.3 (4.6)	196	4.8 (4.1)	<0.01*
CR of violence exposure	301	0.9 (1.5)	123	0.9 (1.3)	178	1.0 (1.6)	0.80
SR of violence exposure	59	1.3 (1.7)	31	1.2 (1.4)	28	1.4 (1.9)	0.65
		Percentage		**Percentage**		**Percentage**	
Child							
Gender							0.05[a]
Male	181	52	87	58	94	48	
Female	164	48	62	42	102	52	
Race and ethnicity							0.41
Hispanic	169	52	71	50	98	53	
White	49	15	21	15	28	15	
Black	58	18	31	22	27	15	
Other	51	16	20	14	31	17	
Caregiver							
Family income level							0.50
Less than $10,000	295	86	131	88	164	84	
$10,001–20,000	43	12	15	10	28	14	
$20,001–30,000	7	2	3	2	4	2	
More than $30,000	—	—	—	—	—	—	
Relationship to child							—
Parent or guardian	344	100	148	99	196	100	
Other relationship	1	<1	1	1	—	—	

NOTE: * = significant difference ($p < 0.05$). — = Cell is too small to show. Because of missing data, some numbers might not sum as shown.
[a] A significant trend in the chi-squared statistic.

We also examined the Worcester sample at baseline on three outcomes (PTSD symptoms, child depressive symptoms, and caregiver depressive symptoms) to describe the level of severity on these indexes among families entering the project (Table I.6). At baseline, nearly one-third of caregivers (29 percent) reported symptoms of PTSD in the significant range for their children (28 percent for boys and 30 percent for girls), and 38 percent of children ages 8–18 reported high levels of PTSD symptoms. At baseline, 25 percent of children ages 13–18 self-reported

depressive symptoms in the moderate clinical range. Eighteen percent of caregivers reported their own depressive symptoms in the moderately severe or severe range.

Table I.6. Baseline Assessment Estimates for Worcester Families

	Combined		Boys		Girls	
Assessment	***N***	**Percentage**	***N***	**Percentage**	***N***	**Percentage**
CR of child PTSD symptoms (ages 3–10)						
Normal	114	63	66	66	48	59
Borderline	15	8	6	6	9	11
Significant	52	29	28	28	24	30
SR of child PTSD symptoms (ages 8–18)						
Low	54	62	25	57	29	67
High	33	38	19	43	14	33
SR of child depression (ages 13–18)						
Normal range	21	58	15	71	6	40
Mild clinical range	6	17	4	19	2	13
Moderate clinical range	9	25	2	10	7	47
Severe clinical range	0	0	0	0	0	0
Caregiver depression						
None or minimal	114	33	70	39	44	27
Mild	102	30	51	28	51	31
Moderate	69	20	28	15	41	25
Moderately severe	40	12	22	12	18	11
Severe	20	6	10	6	10	6

Finally, we examined differences between the intervention and comparison groups at baseline for Worcester's primary, secondary, and tertiary outcomes (see Table I.12). At baseline, there were no differences between groups for the primary, secondary, or tertiary child outcomes after adjusting for multiple comparisons (aside from one tertiary outcomes that did not withstand the correction for multiple comparisons), indicating that the quasi-experimental design resulted in balanced groups.

Table I.12 summarizes cross-sectional differences between intervention and comparison groups at the six- and 12-month follow-up for Worcester's primary, secondary, and tertiary outcomes. There are also no significant differences between groups for the primary, secondary, or tertiary outcomes (aside from three secondary outcomes at six months that did not withstand the correction for multiple comparisons).

Uptake, Dosage, and Process of Care

The program recorded family-level service data on the follow-up service survey submitted at six and 12 months after the baseline assessment. Table I.7 shows the type and amount of services received for all families who were initially enrolled in the intervention group, regardless of whether they continued to participate in the ongoing research assessment. The data displayed in Table I.7 include services received by summing all time points that the program reported, which was six months for most participants in Worcester Safe Start. Worcester did not complete 12-month service surveys for families after stopping the 12-month assessments early.

Table I.7. Services That Worcester Safe Start Intervention Families Received

	With Service		Number of Sessions		
Service	***N***	**Percentage**	**Range**	**Mean**	**Median**
Baseline sample (*n* = 149)					
Child assessment sessions	78	52	0–3	0.8	1
Child-focused service plan sessions	59	40	0–2	0.5	0
Family group therapy sessions	11	7	0–10	0.6	0
Six-month analysis sample (*n* = 93)					
Child assessment sessions	56	60	0–3	0.9	1
Child-focused service plan sessions	46	49	0–2	0.6	0
Family group therapy sessions	9	10	0–10	0.8	0
12-month analysis sample (*n* = 26)					
Child assessment sessions	12	46	0–2	0.8	0
Child-focused service plan sessions	10	38	0–2	0.7	0
Family group therapy sessions	5	19	0–10	1.6	0

NOTE: A participant could receive more than one kind of service, so some numbers might not total as shown.

As shown in the top portion of Table I.7, just over one-half of intervention group families (52 percent) received individual clinical child assessments, 40 percent received child-focused service plans, and only 7 percent received the SFCR group therapy sessions. Among families who received at least one session, the average number of sessions was 1.5 for child assessment sessions, 1.4 for child-focused service plan sessions, and 8.5 for family group therapy sessions. The middle portion of Table I.7 shows the services received by the subgroup of intervention group families who participated in the six-month follow-up research assessment. These are the 93 families included in the intervention group in the outcome analysis sample for the Worcester program at six months. As shown in Table I.7, 60 percent of those assessed at six months received child assessments, 49 percent received child-focused service planning, and 10 percent received family group therapy sessions. Among only families who received each type of service, they had an average of 1.5 child assessment sessions, 1.3 child-focused service plan sessions, and

8.3 family group therapy sessions. The bottom portion of Table I.7 shows the services received by the subgroup of intervention group families who participated in the 12-month follow-up research assessment. These are the 26 families included in the intervention group in the outcome analysis sample for the Worcester program at 12 months. As shown in Table I.7, 46 percent of those assessed at 12 months received child assessments, 38 percent received child-focused service planning, and 19 percent received family group therapy sessions. Among families who received at least a session, they participated in an average of 1.8 child assessment sessions, 8.4 child-focused service plan sessions, and 1.7 family group therapy sessions. We did not collect service information on usual care in the comparison group.

On the follow-up caregiver assessment, we asked caregivers in the intervention group about their satisfaction with Safe Start services (Table I.8). The eight-item measure focused on the level of satisfaction with services received using a four-point Likert scale, with a higher score indicating greater level of satisfaction with services. Overall, caregivers who received Safe Start services were moderately satisfied with different aspects of Safe Start services, with ratings ranging from a low of 2.7 (program met needs) to a high of 3.3 (would recommend program to a friend).

Table I.8. Satisfaction with Services That Worcester Safe Start Intervention Families Received: Six Months

Satisfaction	*N*	Mean	SD
Rate quality of service	60	3.2	0.9
Got the kind of service wanted	60	2.9	0.94
Program met needs	60	2.7	0.99
Would recommend to a friend	60	3.3	0.8
Satisfied with help received	60	3.0	0.91
Helped deal more effectively with problems	60	3.0	0.75
Satisfied with service	60	3.0	0.82
Would come back to program	60	3.2	0.84

Key Outcome Findings

We begin by analyzing changes in mean scores over time both within the intervention and comparison groups and between the groups. For these analyses, we used an intent-to-treat approach that included all families allocated to the intervention, regardless of the level of service they received. The first set of columns of numbers in Table I.9 describes differences within groups between the baseline and the six-month assessment, with paired *t*-tests comparing each person's score at each follow-up wave with his or her own score at the baseline assessment and adjusting for multiple testing. At six months, for the primary outcome measures of different

aspects of social–emotional competence, there were no statistically significant within-group differences at six months. After adjusting for multiple comparisons, there were also no significant within-group differences among the 24 secondary outcomes. There were within-group differences for several of the 27 tertiary outcomes. At six months, both groups had significantly fewer caregiver total stressors, caregiver resource problems, CRs of the child's total victimization experiences, child's witnessing violence experiences, and caregiver's total number of traumatic experiences. Caregivers in both groups also reported significantly fewer of their own depressive symptoms at six months. The intervention group also improved significantly on CR of caregivers' PTSD symptoms, while the comparison group had significantly fewer caregiver experiences of DV.

Table I.9. Changes in Means for Outcome Variables Between Baseline and Six-Month Assessment and Group-Level Comparison of Mean Changes

						Group-Level Comparison of Mean Changes					
		Within-Family Mean Change[a]				Unadjusted Model[b]			Adjusted Model[c]		
Outcome		***N***	**Estimate**	**SE**	***p*-Value**	**Estimate**	**SE**	***p*-Value**	**Estimate**	**SE**	***p*-Value**
Primary											
CR of child social–emotional competence (ages 0–2)	Intervention	22	–1.14	5.90	0.85	2.26	11.31	0.84	4.85	10.99	0.66
	Comparison	50	–3.40	3.00	0.26						
CR of child assertion (ages 3–12)	Intervention	54	0.00	0.47	1.00	–0.61	0.98	0.53	–0.74	0.96	0.44
	Comparison	57	0.61	0.39	0.12						
CR of child self-control (ages 3–12)	Intervention	54	0.09	0.63	0.88	–0.64	1.25	0.61	–0.67	1.19	0.58
	Comparison	57	0.74	0.49	0.13						
Secondary											
CR of child behavioral problems (ages 1–3)	Intervention	24	0.04	1.09	0.97	1.00	2.87	0.73	1.20	2.92	0.68
	Comparison	52	–0.96	0.77	0.22						
CR of child total behavior problems (ages 3–18)	Intervention	68	–1.68	0.72	0.02[d]	–0.50	1.69	0.77	–0.60	1.66	0.72
	Comparison	67	–1.18	0.65	0.08						
CR of child externalizing behavior problems (ages 3–18)	Intervention	68	–0.94	0.49	0.06	–0.34	1.16	0.77	–0.40	1.14	0.73
	Comparison	67	–0.60	0.47	0.21						
CR of child cooperation (ages 3–12)	Intervention	54	1.09	0.48	0.03[d]	0.44	1.08	0.68	0.31	1.03	0.77
	Comparison	57	0.65	0.30	0.03[d]						
SR of child cooperation (ages 13–18)	Intervention	12	1.17	0.61	0.08	—	—	—	—	—	—
	Comparison	9	—	—	—						
SR of child assertion (ages 13–18)	Intervention	14	1.43	0.91	0.14	3.33	2.00	0.10	—	—	—
	Comparison	10	–1.90	1.13	0.13						
SR of child self-control (ages 13–18)	Intervention	14	0.57	0.69	0.42	0.97	2.24	0.67	—	—	—
	Comparison	10	–0.40	1.14	0.73						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
CR of child affective strengths (ages 6–12)	Intervention	26	0.15	0.40	0.71	0.12	1.18	0.92	0.17	1.13	0.88
	Comparison	28	0.04	0.51	0.94						
SR of child affective strengths (ages 11–18)	Intervention	18	–0.61	0.85	0.48	–1.88	1.78	0.30	—	—	—
	Comparison	11	1.27	0.63	0.07						
CR of child school functioning (ages 6–12)	Intervention	22	0.77	0.91	0.41	–0.05	2.16	0.98	0.05	2.08	0.98
	Comparison	23	0.83	0.81	0.32						
SR of child school functioning (ages 11–18)	Intervention	14	0.50	0.65	0.46	0.40	2.80	0.89	—	—	—
	Comparison	10	0.10	1.46	0.95						
CR of family involvement (ages 6–12)	Intervention	26	0.88	0.55	0.12	0.31	1.64	0.85	0.45	1.43	0.76
	Comparison	28	0.57	0.67	0.40						
SR of family involvement (ages 11–18)	Intervention	18	–0.17	0.51	0.75	–0.62	2.09	0.77	—	—	—
	Comparison	11	0.45	0.92	0.63						
CR of positive involvement (ages 6–18)	Intervention	36	–2.92	1.07	0.01[d]	–2.24	3.03	0.46	–2.15	3.13	0.49
	Comparison	34	–0.68	1.27	0.60						
CR of negative or ineffective discipline (ages 6–18)	Intervention	40	–1.05	0.78	0.19	–0.89	1.62	0.58	–0.94	1.55	0.54
	Comparison	37	–0.16	0.81	0.84						
CR of deficient monitoring (ages 6–18)	Intervention	40	0.35	0.32	0.29	0.08	0.78	0.92	0.05	0.64	0.94
	Comparison	37	0.27	0.34	0.44						
SR of mother involvement (ages 8–18)	Intervention	26	–2.58	1.20	0.04[d]	–2.58	3.37	0.45	–2.52	3.27	0.44
	Comparison	22	0.00	1.34	1.00						
SR of father involvement (ages 8–18)	Intervention	29	0.21	1.38	0.88	2.54	4.36	0.56	2.65	4.10	0.52
	Comparison	24	–2.33	2.41	0.34						
SR of positive parenting (ages 8–18)	Intervention	29	0.66	0.79	0.42	1.91	1.80	0.29	1.96	1.72	0.26
	Comparison	24	–1.25	0.77	0.12						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
SR of poor monitoring and supervision (ages 8–18)	Intervention	29	–0.90	0.60	0.15	–1.31	1.79	0.47	–1.35	1.31	0.31
	Comparison	24	0.42	0.41	0.32						
SR of inconsistent discipline (ages 8–18)	Intervention	29	0.21	0.88	0.82	–1.21	1.76	0.49	–1.02	1.56	0.51
	Comparison	24	1.42	0.94	0.15						
SR of corporal punishment (ages 8–18)	Intervention	29	0.07	0.25	0.79	–0.10	0.40	0.81	–0.10	0.37	0.80
	Comparison	24	0.17	0.16	0.29						
CR of family conflict (ages 3–18)	Intervention	93	–0.45	0.19	0.02[d]	0.08	0.36	0.82	0.09	0.37	0.81
	Comparison	120	–0.53	0.19	0.01[d]						
SR of family conflict (ages 11–18)	Intervention	18	0.00	0.41	1.00	1.73	0.82	0.04[d]	—	—	—
	Comparison	11	–1.73	0.54	0.01[d]						
Tertiary											
CR of attitudinal barriers to care	Intervention	93	–0.59	0.24	0.01[d]	–0.17	0.41	0.68	–0.13	0.42	0.75
	Comparison	121	–0.42	0.21	0.05[d]						
CR of total stressors	Intervention	93	–4.53	1.06	<0.01[d]*	–1.50	1.95	0.44	–1.83	1.93	0.34
	Comparison	120	–3.03	0.81	<0.01[d]*						
CR of resource problems	Intervention	93	–2.77	0.50	<0.01[d]*	–0.32	0.92	0.73	–0.28	0.91	0.76
	Comparison	120	–2.46	0.43	<0.01[d]*						
CR of personal problems	Intervention	93	–1.75	0.72	0.02[d]	–1.19	1.24	0.34	–1.56	1.24	0.21
	Comparison	120	–0.57	0.50	0.26						
CR of child PTSD symptoms (ages 3–10)	Intervention	50	–2.24	1.02	0.03[d]	–1.47	3.28	0.65	–1.63	3.25	0.62
	Comparison	56	–0.77	1.37	0.58						
SR of child PTSD symptoms (ages 8–18)	Intervention	29	–1.69	2.11	0.43	–0.52	3.59	0.88	–0.51	3.16	0.87
	Comparison	24	–1.17	1.68	0.49						
SR of child delinquency (ages 11–18)	Intervention	18	–0.06	0.15	0.72	0.04	0.19	0.85	—	—	—
	Comparison	11	–0.09	0.09	0.34						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
SR of child drug use (ages 11–18)	Intervention	18	–0.06	0.06	0.33	0.04	0.11	0.74	—	—	—
	Comparison	11	–0.09	0.09	0.34						
SR of child gang involvement (ages 11–18)	Intervention	18	–0.06	0.06	0.33	–0.06	0.06	0.32	—	—	—
	Comparison	11	0.00	0.00	—						
SR of child depressive symptoms (ages 13–18)	Intervention	14	–2.43	1.07	0.04[d]	–2.53	3.85	0.51	—	—	—
	Comparison	10	0.10	2.04	0.96						
CR of child internalizing problems (ages 3–18)	Intervention	68	–0.74	0.31	0.02[d]	–0.15	0.67	0.82	–0.20	0.65	0.76
	Comparison	67	–0.58	0.29	0.05						
SR of child grades (ages 13–18)	Intervention	14	–1.50	0.49	0.01[d]	–2.10	0.83	0.01[d]	—	—	—
	Comparison	10	0.60	0.43	0.19						
CR of total child victimization experiences (ages 0–11)	Intervention	77	–0.58	0.14	<0.01[d]*	–0.18	0.21	0.40	–0.13	0.21	0.53
	Comparison	110	–0.41	0.13	<0.01[d]*						
CR of child maltreatment (ages 0–11)	Intervention	77	–0.06	0.04	0.10	0.00	0.05	0.98	0.00	0.06	0.97
	Comparison	110	–0.06	0.04	0.11						
CR of child assault (ages 0–11)	Intervention	77	–0.08	0.06	0.18	–0.05	0.08	0.51	–0.04	0.08	0.61
	Comparison	110	–0.03	0.05	0.57						
CR of child sexual abuse (ages 0–11)	Intervention	77	0.03	0.02	0.16	0.03	0.02	0.25	0.03	0.02	0.26
	Comparison	110	0.00	0.01	1.00						
CR of child witnessing violence (ages 0–11)	Intervention	77	–0.38	0.09	<0.01[d]*	–0.08	0.13	0.57	–0.05	0.14	0.74
	Comparison	110	–0.30	0.09	<0.01[d]*						
SR of total child victimization experiences (ages 10–18)	Intervention	18	–0.78	0.32	0.03[d]	–0.19	0.65	0.77	—	—	—
	Comparison	17	–0.59	0.27	0.05						
SR of child maltreatment (ages 10–18)	Intervention	18	–0.22	0.13	0.10	0.07	0.19	0.71	—	—	—
	Comparison	17	–0.29	0.14	0.06						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
SR of child assault (ages 10–18)	Intervention	18	–0.28	0.14	0.06	–0.16	0.17	0.35	—	—	—
	Comparison	17	–0.12	0.08	0.16						
SR of child sexual abuse (ages 10–18)	Intervention	18	0.00	0.00	—	0.00	0.00	—	—	—	—
	Comparison	17	0.00	0.00	—						
SR of child witnessing violence (ages 10–18)	Intervention	18	–0.17	0.27	0.55	0.07	0.45	0.88	—	—	—
	Comparison	17	–0.24	0.18	0.22						
CR of caregiver total number of traumatic experiences	Intervention	93	–0.34	0.11	<0.01[d]*	0.07	0.45	0.63	0.08	0.14	0.57
	Comparison	120	–0.42	0.09	<0.01[d]*						
CR of caregiver experience of any DV	Intervention	93	–0.12	0.05	0.01[d]	0.05	0.06	0.42	0.05	0.06	0.35
	Comparison	120	–0.17	0.04	<0.01[d]*						
CR of caregiver experience of any non-DV trauma	Intervention	93	–0.05	0.04	0.20	0.00	0.06	0.94	–0.02	0.06	0.80
	Comparison	120	–0.06	0.03	0.07						
CR of caregiver depression	Intervention	93	–1.82	0.54	<0.01[d]*	–0.23	1.09	0.83	–0.27	1.08	0.81
	Comparison	120	–1.58	0.41	<0.01[d]*						
CR of caregiver PTSD	Intervention	93	–0.63	0.16	<0.01[d]*	–0.39	0.29	0.17	–0.36	0.29	0.22
	Comparison	120	–0.24	0.14	0.09						

NOTE: * = the significant difference over time remains after adjustment for multiple comparisons. — = Cell is too small to show. Mean change estimates are not shown when the group size is fewer than ten, and comparisons are not shown when the group size is fewer than ten for either group. Adjusted model results are not shown when the group size is fewer than 20 for either group.

[a] Within-family mean changes between the baseline and six-month scores for each group separately.

[b] Group-level comparison of within-family mean changes from baseline to six months.

[c] Estimate of the difference between the two groups' within-family mean changes from baseline to six months, controlling for child age, gender, race and ethnicity, and violence exposure (baseline exposure plus exposure between baseline and six months).

[d] A significant paired *t*-test of differences over time ($p < 0.05$). In cells containing *0.05*, we have rounded the value to 0.05, but it is still less than 0.05.

Table I.9 also shows intervention effects over time using an intent-to-treat approach in which all families in the intervention are compared with all those in the comparison group, regardless of the actual amount of intervention received in the intervention group. Because any change in outcomes observed can potentially be the result of a time trend observed in all children in the study, we used a difference-in-differences method to assess the unadjusted impact of the program. For this quasi-experimental study, the adjusted models are important because they control for the apparent differences between the two groups, so the unadjusted models need to be interpreted with caution. In the difference-in-differences models presented in the second set of columns in Table I.9, there was no evidence of intervention effects for the primary outcomes related to CR of different aspects of child social–emotional competence. Among secondary and tertiary outcomes, none of the unadjusted models showed evidence of an intervention effect. The adjusted models in the third set of columns control for child age, child gender, child race and ethnicity, and caregiver report of violence exposure. There was no evidence of an intervention effect for any of the primary, secondary, or tertiary outcomes with a large enough sample size to run the model. The results in Table I.9 can also be discussed in terms of effect sizes for Worcester's primary outcomes. Within-group changes were very small for the intervention for child social–emotional competence (–0.04 [–0.65 – 0.57]), assertion (0.00 [–0.38 – 0.38]), and self-control (0.02 [–0.36 – 0.40]). The adjusted between-group effect size was also very small for child social–emotional competence (0.08 [–0.28 – 0.45]), assertion (0.10 [–0.36 – 0.15]), and self-control (–0.07 [–0.33 – 0.18]), indicating again that the two groups did not differ much in terms of their improvement.

Table I.10 presents changes in outcomes from baseline to 12 months for each group, as well as group comparisons in the 12-month outcomes. These show a similar pattern of results except with much smaller numbers of families included in the analysis and therefore much less power to detect differences. For primary, secondary, and tertiary outcomes, there were some statistical trends in the within-group changes, but none of them withstands the correction for multiple testing. There was no evidence of intervention effects at 12 months for any of the outcomes in both unadjusted and adjusted models.

Table I.10. Changes in Means for Outcome Variables Between Baseline and 12-Month Assessment and Group-Level Comparison of Mean Changes

			Within-Family Mean Change[a]			Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
Outcome		*N*	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
Primary											
CR of child social–emotional competence (ages 0–2)	Intervention	5	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						

Outcome		N	Within-Family Mean Change[a] Estimate	SE	p-Value	Group-Level Comparison of Mean Changes: Unadjusted Model[b] Estimate	SE	p-Value	Adjusted Model[c] Estimate	SE	p-Value
CR of child assertion (ages 3–12)	Intervention	14	–1.79	0.68	0.02[d]	–3.06	1.73	0.08	—	—	—
	Comparison	11	1.27	1.02	0.24						
CR of child self-control (ages 3–12)	Intervention	14	0.29	1.28	0.83	–0.99	3.09	0.75	—	—	—
	Comparison	11	1.27	0.92	0.19						
Secondary											
CR of child behavioral problems (ages 1–3)	Intervention	3	—	—	—	—	—	—	—	—	—
	Comparison	6	—	—	—						
CR of child total behavior problems (ages 3–18)	Intervention	19	–2.47	1.58	0.13	2.13	3.43	0.54	—	—	—
	Comparison	15	–4.60	1.28	<0.01[d]						
CR of child externalizing behavior problems (ages 3–18)	Intervention	19	–1.89	1.04	0.09	1.37	2.39	0.57	—	—	—
	Comparison	15	–3.27	0.80	<0.01[d]						
CR of child cooperation (ages 3–12)	Intervention	14	–0.29	1.01	0.78	–0.56	2.61	0.83	—	—	—
	Comparison	11	0.27	0.84	0.75						
SR of child cooperation (ages 13–18)	Intervention	2	—	—	—	—	—	—	—	—	—
	Comparison	3	—	—	—						
SR of child assertion (ages 13–18)	Intervention	3	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
SR of child self-control (ages 13–18)	Intervention	3	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
CR of child affective strengths (ages 6–12)	Intervention	9	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of child affective strengths (ages 11–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
CR of child school functioning (ages 6–12)	Intervention	8	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
SR of child school functioning (ages 11–18)	Intervention	4	—	—	—	—	—	—	—	—	—
	Comparison	3	—	—	—						

Outcome		N	Within-Family Mean Change[a] Estimate	SE	p-Value	Group-Level Comparison of Mean Changes: Unadjusted Model[b] Estimate	SE	p-Value	Adjusted Model[c] Estimate	SE	p-Value
CR of family involvement (ages 6–12)	Intervention	9	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of family involvement (ages 11–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
CR of positive involvement (ages 6–18)	Intervention	11	–2.91	2.13	0.20	—	—	—	—	—	—
	Comparison	8			—						
CR of negative or ineffective discipline (ages 6–18)	Intervention	14	–2.79	1.45	0.08	—	—	—	—	—	—
	Comparison	9	—	—	—						
CR of deficient monitoring (ages 6–18)	Intervention	14	0.07	0.96	0.94	—	—	—	—	—	—
	Comparison	9									
SR of mother involvement (ages 8–18)	Intervention	11	1.18	2.41	0.63	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of father involvement (ages 8–18)	Intervention	12	2.50	2.98	0.42	—	—	—	—	—	—
	Comparison	5									
SR of positive parenting (ages 8–18)	Intervention	12	1.00	0.89	0.28	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of poor monitoring and supervision (ages 8–18)	Intervention	12	–1.50	1.17	0.23	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of inconsistent discipline (ages 8–18)	Intervention	12	–1.00	1.23	0.43	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of corporal punishment (ages 8–18)	Intervention	12	–0.83	0.53	0.15	—	—	—	—	—	—
	Comparison	5	—	—	—						
CR of family conflict (ages 3–18)	Intervention	26	–1.85	0.43	<0.01[d]	—	—	—	—	—	—
	Comparison	26	–2.58	0.46	<0.01[d]						
SR of family conflict (ages 11–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
Tertiary outcomes											
CR of attitudinal barriers to care	Intervention	26	–1.85	0.53	<0.01[d]	–0.96	0.88	0.28	–0.96	0.88	0.28
	Comparison	26	–0.88	0.47	0.07						

Outcome		N	Within-Family Mean Change[a] Estimate	SE	p-Value	Group-Level Comparison of Mean Changes: Unadjusted Model[b] Estimate	SE	p-Value	Adjusted Model[c] Estimate	SE	p-Value
CR of total stressors	Intervention	26	-8.92	2.50	<0.01[d]	–3.92	3.68	0.29	–3.92	3.67	0.29
	Comparison	26	–5.00	1.29	<0.01[d]						
CR of resource problems	Intervention	26	–5.73	1.09	<0.01[d]	–1.88	1.85	0.31	–1.88	1.84	0.31
	Comparison	26	–3.85	0.85	<0.01[d]						
CR of personal problems	Intervention	26	–3.19	1.67	0.07	–2.04	2.24	0.37	–2.04	2.25	0.37
	Comparison	26	–1.15	0.72	0.12						
CR of child PTSD symptoms (ages 3–10)	Intervention	12	–3.83	2.40	0.14	1.37	7.04	0.85	—	—	—
	Comparison	10	–5.20	4.25	0.25						
SR of child PTSD symptoms (ages 8–18)	Intervention	12	–2.58	3.40	0.46	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of child delinquency (ages 11–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
SR of child drug use (ages 11–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
SR of child gang involvement (ages 11–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
SR of child depressive symptoms (ages 13–18)	Intervention	3	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
CR of child internalizing problems (ages 3–18)	Intervention	19	–0.58	0.71	0.42	0.75	1.34	0.57	—	—	—
	Comparison	15	–1.33	0.73	0.09						
SR of child grades (ages 13–18)	Intervention	3	—	—	—	—	—	—	—	—	—
	Comparison	4	—	—	—						
CR of total child victimization experiences (ages 0–11)	Intervention	19	–0.21	0.29	0.48	—	—	—	—	—	—
	Comparison	22	–0.59	0.31	0.07						
CR of child maltreatment (ages 0–11)	Intervention	19	–0.05	0.09	0.58	—	—	—	—	—	—
	Comparison	22	–0.14	0.15	0.38						
CR of child assault (ages 0–11)	Intervention	19	–0.05	0.09	0.58	—	—	—	—	—	—
	Comparison	22	–0.14	0.10	0.19						
CR of child sexual abuse (ages 0–11)	Intervention	19	0.00	0.00	.	—	—	—	—	—	—
	Comparison	22	–0.05	0.05	0.33						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
CR of child witnessing violence (ages 0–11)	Intervention	19	0.00	0.24	1.00	—	—	—	—	—	—
	Comparison	22	–0.27	0.18	0.14						
SR of total child victimization experiences (ages 10–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of child maltreatment (ages 10–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of child assault (ages 10–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of child sexual abuse (ages 10–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
SR of child witnessing violence (ages 10–18)	Intervention	7	—	—	—	—	—	—	—	—	—
	Comparison	5	—	—	—						
CR of caregiver total number of traumatic experiences	Intervention	26	–0.23	0.23	0.33	0.35	0.34	0.31	0.35	0.34	0.31
	Comparison	26	–0.58	0.24	0.03[d]						
CR of caregiver experience of any DV	Intervention	26	–0.04	0.10	0.71	0.19	0.15	0.20	0.19	0.15	0.20
	Comparison	26	–0.23	0.10	0.03[d]						
CR of caregiver experience of any non-DV trauma	Intervention	26	–0.12	0.06	0.08	0.08	0.11	0.50	0.08	0.11	0.50
	Comparison	26	–0.19	0.10	0.06						
CR of caregiver depression	Intervention	26	–2.38	1.59	0.15	–0.08	2.67	0.98	–0.08	2.63	0.98
	Comparison	26	–2.31	1.21	0.07						
CR of caregiver PTSD	Intervention	26	–0.65	0.27	0.02[d]	–0.04	0.60	0.95	–0.04	0.59	0.95
	Comparison	26	–0.62	0.34	0.08						

NOTE: * = the significant difference over time remains after adjustment for multiple comparisons. — = Cell is too small to show. Mean change estimates are not shown when the group size is fewer than ten, and comparisons are not shown when the group size is fewer than ten for either group. Adjusted model results are not shown when the group size is fewer than 20 for either group.

[a] Within-family mean changes between the baseline and 12-month scores for each group separately.

[b] Group-level comparison of within-family mean changes from baseline to 12 months.

[c] Estimate of the difference between the two groups' within-family mean changes from baseline to 12 months, controlling for child age, gender, race and ethnicity, and violence exposure (baseline exposure plus the average of exposure at 6 and 12 months).

[d] A significant paired *t*-test of differences over time ($p < 0.05$).

Safe Start Service Dosage and Changes in Primary Outcomes

To examine any intervention effects that Safe Start service dosage has on outcomes, we divided the intervention families into two dosage groups because families tend to receive different amounts of services. Families who received either the child assessment or child-focused service plan or who participated in the SFCR group therapy sessions were categorized as having received any Safe Start services. The other group consisted of families who did not receive or participate in any of the Safe Start services. About one-half of the sample fell into each group at six months.

Because children and families with more need are likely to receive more services, we would expect a selection bias, with higher-need families receiving more services. To account for this selection bias, we used the propensity score matching method to pair families in each dosage group with families with similar needs in the comparison group. The matching paired families based on baseline scores on the outcome measure of interest. The analyses examined the difference in mean score changes between the intervention and comparison groups for each dosage group. Note that, in this analysis, the full comparison group is used in the matching of each of the dosage levels.

At six months, the propensity score matching analyses for primary outcomes (with the exception of the no-service group for child social–emotional competence, for which the sample size was too small) showed that, across the outcome measures, there was a lot of variability in the direction of change for the any-service dosage group. Some scores decreased when we expected them to increase, while others changed in the expected direction but only by a small amount (see Table I.11). The statistical test comparing the two groups on changes in mean scores between baseline and six months revealed no statistically significant differences between the intervention and comparison groups.

Table I.11. Changes in Means, by Dosage Group, for Primary Outcome Variables Between Baseline and Six-Month Assessment

Primary Outcome	Dosage	Group	*N*	Baseline		Six Months		Difference in Differences
				Mean	SD	Mean	SD	
CR of child social–emotional competence (ages 3–12)	None	Intervention	8	31.88	18.89	33.75	27.22	—
		Comparison	8	30.63	15.68	29.38	18.02	
	Any	Intervention	12	38.75	22.58	37.08	29.35	7.08
		Comparison	12	40.00	25.05	31.25	28.29	
CR of child assertion (ages 3–12)	None	Intervention	15	14.47	2.88	15.07	4.03	0.20
		Comparison	15	14.53	3.29	14.93	2.94	
	Any	Intervention	28	13.75	2.46	13.18	3.93	–0.96
		Comparison	28	14.21	3.52	14.61	3.24	
CR of child self-control (ages 3–12)	None	Intervention	20	7.90	4.56	10.30	4.60	1.70
		Comparison	20	7.25	4.12	7.95	3.44	
	Any	Intervention	19	8.84	3.98	8.63	4.60	–0.11
		Comparison	19	9.00	4.19	8.89	4.23	

NOTE: Data are not shown for outcomes when cell size is less than five for the group. Comparisons were not tested when the group size is less than ten for either group. — = Cell is too small to show.

Conclusions

Worcester's quasi-experimental study in homeless shelters compared usual shelter services augmented by child-focused assessment and service planning plus optional SFCR group therapy sessions with usual shelter services alone. The study enrolled 345 families but retained only 62 percent of them at six months and about 40 percent of them at 12 months. The participants in the study were all residing in the homeless shelters and were largely minorities (52 percent Hispanic, 18 percent black, 16 percent other) and impoverished (86 percent with family incomes less than $10,000). Families in the sample had moderate but variable levels of symptoms at the beginning of the study: About 30 to 40 percent of children had PTSD symptoms in the clinical range, about one-quarter of children had depressive symptoms in the clinical range, and 18 percent of caregivers reported depression in the clinical range. Uptake of services was rather low: Just over one-half of intervention group families received individual clinical child assessments, 40 percent received child-focused service plans, and only 7 percent received the SFCR group therapy sessions.

Worcester enrolled and retained enough families to detect a medium intervention effect of size 0.39 at six months, so the study was fully powered to detect the medium effect size expected. However, intent-to-treat analyses showed that mean scores in the intervention and comparison groups were in the expected direction across most outcome measures, but, overall,

there was no evidence of intervention effects at six or 12 months for any of the outcomes because the comparison group experienced similar improvements over time, with small between-group effects. Examination of families who received some services within the intervention, as compared with those who did not, showed a lot of variability in the direction of change for the group that received any services, without a clear pattern. Most of the any-service group received the child-focused assessment, child-focused service plan, or both, with very few receiving the SFCR group therapy sessions. Although we expected those in the intervention group who received the assessment or service plan to improve more than the matched comparison group, we saw very little change in either group among those who received any services. Because only a very small portion of Worcester families received the SFCR group therapy sessions, this study could not test the effectiveness of that portion of the intervention.

Despite the quasi-experimental design, we cannot draw firm conclusions about the effectiveness of the Worcester Safe Start program based on these results, however. Although it is possible that the program was ineffective in improving the outcomes for children in the intervention group relative to comparison families, there are several other possible explanations for the results presented here. First, the uptake of services within the intervention group was low, and it is possible that the families did not receive an adequate dose that would have led to improved outcomes. Second, in this program, the comparison group received shelter services as usual, including a wide array of social services. These services might have served to reduce the amount of difference between the two groups on the tertiary outcomes that showed some improvements. The failure to detect significant differences between the groups on tertiary outcomes might also have been due to the particular outcomes measured. That is, the Worcester Safe Start program might have improved the lives of children and families in ways that we did not measure (or measured adequately) in this study. Finally, the baseline status of families in the study showed a relatively healthy sample on some of the measures, leaving little room for improvement.

In sum, the Worcester Safe Start intervention did not show improved outcomes for children and families, despite the power to detect a medium effect. Implementation problems and family factors leading to low uptake of services impeded the study, with a minority of families receiving the planned intervention. Thus, further research is needed to fully test the intervention both in terms of improving implementation and in terms of measuring outcomes.

Table I.12. Comparison of Means for Worcester Outcome Variables

		Baseline				Six Months				12 Months			
Outcome		***N***	**Mean**	**SD**	**Difference**	***N***	**Mean**	**SD**	**Difference**	***N***	**Mean**	**SD**	**Difference**
Primary													
CR of child social–emotional competence (ages 0–2)	Intervention	30	39.33	31.40	4.81	22	39.77	35.67	8.17	5	58.00	44.1	—
	Comparison	73	34.52	22.07		50	31.60	22.84		7	28.57	17.49	
CR of child assertion (ages 3–12)	Intervention	93	14.15	3.05	0.20	54	13.89	3.84	–0.32	16	13.19	2.74	–1.53
	Comparison	99	13.95	3.60		57	14.21	3.91		14	14.71	3.81	
CR of child self-control (ages 3–12)	Intervention	93	9.58	4.72	0.99	54	9.48	4.73	0.57	16	10.06	5.54	0.99
	Comparison	99	8.59	4.68		57	8.91	4.23		14	9.07	5.40	
Secondary													
CR of child behavioral problems (ages 1–3)	Intervention	33	14.48	9.13	1.62	24	13.75	8.29	1.46	7	11.14	8.11	—
	Comparison	76	12.87	6.77		52	12.29	7.02		10	11.60	6.26	
CR of child total behavior problems (ages 3–18)	Intervention	114	10.94	6.87	–0.08	68	10.07	7.18	0.09	21	10.90	7.11	3.46
	Comparison	115	11.02	6.98		67	9.99	6.84		18	7.44	6.22	
CR of child externalizing behavior problems (ages 3–18)	Intervention	114	7.70	4.70	–0.17	68	7.32	4.95	–0.02	21	7.81	4.83	2.48
	Comparison	115	7.87	4.74		67	7.34	4.76		18	5.33	4.30	
CR of child cooperation (ages 3–12)	Intervention	93	11.49	4.04	0.39	54	12.15	4.16	0.76	16	11.13	4.30	0.05
	Comparison	99	11.10	4.13		57	11.39	3.79		14	11.07	3.91	
SR of child cooperation (ages 13–18)	Intervention	18	14.06	3.78	–2.01	13	15.69	3.75	—	3	—	—	—
	Comparison	15	16.07	4.65		9	15.78	3.46		3	—	—	
SR of child assertion (ages 13–18)	Intervention	20	11.65	2.78	–2.04	14	12.43	2.79	1.63	5	—	—	—
	Comparison	16	13.69	4.06		10	10.80	3.65		4	—	—	
SR of child self-control (ages 13–18)	Intervention	20	9.95	4.33	–1.18	14	9.43	4.13	–0.87	5	—	—	—
	Comparison	16	11.13	3.59		10	10.30	3.65		4	—	—	
CR of child affective strengths (ages 6–12)	Intervention	45	15.58	3.6	–0.84	26	16.04	3.07	–0.28	9	16.33	3.39	—
	Comparison	46	16.41	3.11		28	16.32	2.99		6	16.00	1.55	

Outcome		Baseline				Six Months				12 Months			
		N	Mean	SD	Difference	*N*	Mean	SD	Difference	*N*	Mean	SD	Difference
SR of child affective strengths (ages 11–18)	Intervention	28	13.29	2.95	–1.56	18	12.39	3.85	–3.07[a]	7	13.86	4.74	—
	Comparison	19	14.84	3.47		11	15.45	3.27		5	15.60	3.36	
CR of child school functioning (ages 6–12)	Intervention	40	19.60	4.53	–0.01	23	19.26	5.48	–0.66	9	19.33	4.00	–2.27
	Comparison	38	19.61	5.15		26	19.92	5.69		5	21.60	5.03	
SR of child school functioning (ages 11–18)	Intervention	24	18.83	4.88	–0.40	16	19.50	5.01	0.00	5	—	—	—
	Comparison	17	19.24	4.88		10	19.50	4.25		4	—	—	
CR of family involvement (ages 6–12)	Intervention	45	22.40	4.29	–0.62	26	22.54	4.16	–1.35	9	22.00	3.81	0.33
	Comparison	46	23.02	4.42		28	23.89	3.62		6	21.67	3.14	
SR of family involvement (ages 11–18)	Intervention	28	21.79	4.39	–1.06	18	21.56	4.76	–1.99	7	23.71	4.68	—
	Comparison	19	22.84	4.00		11	23.55	3.50		5	23.60	3.05	
CR of positive involvement (ages 6–18)	Intervention	61	63.75	9.64	–3.57[a]	18	21.56	4.76	–4.96[a]	12	63.42	7.03	—
	Comparison	56	67.32	7.17		34	65.91	8.09		9	70.78	7.43	
CR of negative or ineffective discipline (ages 6–18)	Intervention	66	22.36	4.71	0.91	40	21.85	5.09	0.69	14	20.93	5.99	2.53
	Comparison	62	21.45	4.73		37	21.16	5.75		10	18.4	6.22	
CR of deficient monitoring (ages 6–18)	Intervention	66	9.58	3.10	0.35	40	9.73	3.16	0.43	14	9.50	2.28	0.40
	Comparison	62	9.23	2.12		37	9.30	1.81		10	9.10	1.37	
SR of mother involvement (ages 8–18)	Intervention	45	33.87	8.11	–3.19	27	31.44	8.40	–4.82[a]	11	38.09	8.88	—
	Comparison	35	37.06	7.96		23	36.26	6.92		5	33.40	3.58	
SR of father involvement (ages 8–18)	Intervention	48	25.73	11.35	–1.94	29	23.86	10.83	–0.80	12	31.25	11.73	—
	Comparison	39	27.67	11.5		24	24.67	11.19		5	24.00	10.65	
SR of positive parenting (ages 8–18)	Intervention	48	23.31	4.12	–1.10	29	23.59	4.11	0.13	12	25.58	4.23	—
	Comparison	39	24.41	4.98		24	23.46	4.94		5	23.80	2.39	
SR of poor monitoring and supervision (ages 8–18)	Intervention	48	14.85	4.46	1.24	29	13.62	3.63	–0.46	12	13.42	4.40	—
	Comparison	39	13.62	5.02		24	14.08	4.95		5	16.60	6.54	
SR of inconsistent discipline (ages 8–18)	Intervention	48	14.52	4.81	0.11	29	14.14	4.63	–1.32	12	13.67	3.52	—
	Comparison	39	14.41	4.84		24	15.46	3.72		5	14.4	6.19	

Outcome		Baseline				Six Months				12 Months			
		N	Mean	SD	Difference	*N*	Mean	SD	Difference	*N*	Mean	SD	Difference
SR of corporal punishment (ages 8–18)	Intervention	48	3.60	1.3	0.40	29	3.55	1.59	0.26	12	3.33	0.89	—
	Comparison	39	3.21	0.57		24	3.29	0.69		5	3.00	0.00	
CR of family conflict (ages 3–18)	Intervention	149	2.23	1.92	0.07	93	1.83	1.64	0.27	26	1.73	1.78	0.69
	Comparison	196	2.16	2.17		120	1.56	1.56		26	1.04	1.08	
SR of family conflict (ages 11–18)	Intervention	28	2.75	1.92	–0.04	18	2.44	1.82	1.08	7	2.00	1.91	—
	Comparison	19	2.79	2.02		11	1.36	1.12		5	1.40	2.07	
Tertiary outcomes													
CR of attitudinal barriers to care	Intervention	149	2.02	2.26	0.08	93	1.34	2.09	–0.12	26	0.92	2.15	–0.42
	Comparison	197	1.94	2.09		121	1.46	1.97		26	1.35	1.94	
CR of total stressors	Intervention	149	44.21	9.75	2.15[a]	93	38.97	11.68	–0.14	26	37.65	11.47	0.81
	Comparison	196	42.06	9.28		120	39.11	9.79		26	36.85	8.83	
CR of resource problems	Intervention	149	18.53	4.33	0.44	93	15.33	5.01	–0.48	26	14.31	5.81	–0.46
	Comparison	196	18.09	4.74		120	15.82	4.83		26	14.77	5.09	
CR of personal problems	Intervention	149	25.68	6.67	1.72[a]	93	23.63	7.46	0.34	26	23.35	6.44	1.27
	Comparison	196	23.97	5.72		120	23.29	6.06		26	22.08	4.92	
CR of child PTSD symptoms (ages 3–10)	Intervention	85	41.44	11.88	–0.14	50	39.98	12.41	0.68	14	42.43	13.12	6.27
	Comparison	96	41.57	12.57		56	39.3	11.37		13	36.15	5.76	
SR of child PTSD symptoms (ages 8–18)	Intervention	48	9.77	10.40	–0.20	29	7.17	7.63	–1.49	12	6.92	7.27	—
	Comparison	39	9.97	8.74		24	8.67	9.88		5	0.20	0.45	
SR of child delinquency (ages 11–18)	Intervention	28	0.32	0.48	0.11	18	0.33	0.49	0.33[a]	7	0.14	0.38	—
	Comparison	19	0.21	0.42		11	0	0		5	0	0	
SR of child drug use (ages 11–18)	Intervention	28	0.11	0.31	–0.05	18	0	0	—	7	0	0	—
	Comparison	19	0.16	0.37		11	0	0		5	0	0	
SR of child gang involvement (ages 11–18)	Intervention	28	0.04	0.19	0.04	18	0	0	—	7	0	0	—
	Comparison	19	0.00	0.00		11	0	0		5	0	0	

Outcome		Baseline				Six Months				12 Months			
		N	Mean	SD	Difference	*N*	Mean	SD	Difference	*N*	Mean	SD	Difference
SR of child depressive symptoms (ages 13–18)	Intervention	20	75.75	6.11	1.06	14	74.14	6.63	–0.76	5	—	—	—
	Comparison	16	74.69	6.86		10	74.90	5.8		4	—	—	
CR of child internalizing problems (ages 3–18)	Intervention	114	3.24	2.69	0.09	68	2.75	2.82	0.11	21	3.10	2.74	0.98
	Comparison	115	3.15	2.7		67	2.64	2.76		18	2.11	2.72	
SR of child grades (ages 13–18)	Intervention	20	4.10	1.92	1.35[a]	14	2.50	1.29	–0.80	5	—	—	—
	Comparison	16	2.75	1.13		10	3.30	1.34		4	—	—	
CR of total child victimization experiences (ages 0–11)	Intervention	123	0.91	1.27	–0.04	77	0.27	0.62	–0.05	19	0.63	0.96	0.36
	Comparison	178	0.96	1.58		110	0.32	0.68		22	0.27	0.55	
CR of child maltreatment (ages 0–11)	Intervention	123	0.11	0.37	–0.07	77	0.01	0.11	–0.02	19	0.05	0.23	–0.04
	Comparison	178	0.19	0.53		110	0.04	0.19		22	0.09	0.29	
CR of child assault (ages 0–11)	Intervention	123	0.14	0.39	–0.04	77	0.09	0.29	0.00	19	0.05	0.23	0.05
	Comparison	178	0.17	0.51		110	0.09	0.29		22	0	0	
CR of child sexual abuse (ages 0–11)	Intervention	123	0.02	0.18	0.01	77	0.03	0.16	0.02	19	0	0	—
	Comparison	178	0.01	0.11		110	0.01	0.1		22	0	0	
CR of child witnessing violence (ages 0–11)	Intervention	123	0.50	0.87	–0.01	77	0.06	0.25	–0.05	19	0.37	0.76	0.28
	Comparison	178	0.51	0.95		110	0.12	0.35		22	0.09	0.43	
SR of total child victimization experiences (ages 10–18)	Intervention	31	1.16	1.44	–0.20	18	0.67	1.03	0.22	11	0.55	0.82	—
	Comparison	28	1.36	1.89		18	0.44	0.86		5	0.20	0.45	
SR of child maltreatment (ages 10–18)	Intervention	31	0.16	0.45	–0.23	18	0	0	0.00	11	0	0	—
	Comparison	28	0.39	0.74		18	0	0		5	0	0	
SR of child assault (ages 10–18)	Intervention	31	0.19	0.48	0.05	18	0.06	0.24	0.06	11	0.09	0.3	—
	Comparison	28	0.14	0.36		18	0	0		5	0.2	0.45	
SR of child sexual abuse (ages 10–18)	Intervention	31	0.03	0.18	0.03	18	0	0	—	11	0	0	—
	Comparison	28	0.00	0.00		18	0	0		5	0	0	
SR of child witnessing violence (ages 10–18)	Intervention	31	0.61	0.95	–0.14	18	0.56	0.98	0.28	11	0.45	0.82	—
	Comparison	28	0.75	1.24		18	0.28	0.57		5	0	0	

Outcome		Baseline				Six Months				12 Months			
		N	Mean	SD	Difference	*N*	Mean	SD	Difference	*N*	Mean	SD	Difference
CR of caregiver total number of traumatic experiences	Intervention	149	0.55	1.05	–0.06	93	0.14	0.41	0.01	26	0.19	0.63	–0.08
	Comparison	196	0.61	1.12		120	0.13	0.4		26	0.27	0.72	
CR of caregiver experience of any DV	Intervention	149	0.19	0.39	–0.03	93	0.04	0.2	0.01	26	0.12	0.33	—
	Comparison	196	0.21	0.41		120	0.03	0.18		26	0.12	0.33	
CR of caregiver experience of any non-DV trauma	Intervention	149	0.13	0.34	0.01	93	0.08	0.27	0.01	26	0.00	0.00	–0.04
	Comparison	196	0.13	0.33		120	0.07	0.25		26	0.04	0.20	
CR of caregiver depression	Intervention	149	8.83	6.09	1.14	93	6.40	5.36	0.07	26	8.08	6.65	1.27
	Comparison	196	7.68	6.00		120	6.33	5.39		26	6.81	6.15	
CR of caregiver PTSD	Intervention	149	1.69	1.61	0.22	93	1.09	1.4	–0.15	26	1.38	1.65	0.38
	Comparison	196	1.47	1.51		120	1.23	1.43		26	1.00	1.44	

NOTE: * = the significant difference between groups remains after adjustment for multiple comparisons. — = Cell is too small to show. Data are not shown for outcomes when the cell size is fewer than five for either group. Comparisons were not tested when the group size was fewer than ten for either group.
[a] A significant paired *t*-test of differences between groups ($p < 0.05$). In cells containing *0.05*, we have rounded the value to 0.05, but it is still less than 0.05.