This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

TECHNICAL REPORT

National Evaluation of Safe Start Promising Approaches

Results Appendix E: Dayton, Ohio

In Jaycox, L. H., L. J. Hickman, D. Schultz, D. Barnes-Proby, C. M. Setodji, A. Kofner, R. Harris, J. D. Acosta, and T. Francois, *National Evaluation of Safe Start Promising Approaches: Assessing Program Outcomes,* Santa Monica, Calif.: RAND Corporation, TR-991-1-DOJ, 2011

Sponsored by the U.S. Department of Justice's Office of Juvenile Justice and Delinquency Prevention

HEALTH and
INFRASTRUCTURE, SAFETY, AND ENVIRONMENT

This research was sponsored by the U.S. Department of Justice's Office of Juvenile Justice and Delinquency Prevention and was conducted under the auspices of the Safety and Justice Program within RAND Infrastructure, Safety, and Environment and under RAND Health's Health Promotion and Disease Prevention Program.

Library of Congress Control Number: 2011935596

ISBN: 978-0-8330-5822-5

Published 2011 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

DAYTON, OHIO, SAFE START OUTCOMES REPORT

ABSTRACT

The Dayton Safe Start Program implemented a therapeutic and case management intervention to improve outcomes for children ages 5 and younger who had been exposed to domestic violence. The program consisted of Child-Parent Psychotherapy (CPP) sessions and ongoing case management for children and their mothers. A full description of the program can be found in *National Evaluation of Safe Start Promising Approaches: Assessing Program Implementation* (Schultz et al., 2010). The evaluation consisted of a randomized controlled trial of the Dayton intervention, with randomization occurring at the family level. Dayton program staff enrolled 55 families in the study, with less than one-half of them (45 percent) retained for the six-month assessment. At baseline, caregivers reported that the children, with an average age of 1.7 years, had been exposed to an average of 2.6 types of violence in their lives. A substantial minority (40 percent) of enrolled families reported baseline parental stress levels that fell in the "clinical" range. Overall, between 50 and 80 percent of the intervention group families in the six-month analysis sample received each of the service offerings. Eighty percent of these families received CPP, 53 percent had multidisciplinary team involvement, and 47 percent received case management. Results of the study showed that the mean scores for the outcomes assessed moved in the expected direction. However, results of the intent-to-treat analyses showed no differences between groups over time on any of the primary, secondary, or tertiary outcome measures with adequate sample sizes to make these comparisons.

Overall, the sample size limitations mean that no conclusions can be drawn about the effectiveness of the Dayton intervention as implemented. The Dayton Safe Start model requires further testing with an adequate sample size to statistically assess whether the intervention positively affects outcomes for very young children exposed to domestic violence.

INTRODUCTION

The Dayton Safe Start program is located in Dayton, Ohio, which is within Montgomery County. At the time of the original proposal, Montgomery County, the fourth-largest county in Ohio, had significantly higher violent crime rates;

poorer social, health, and economic indicators; and disproportionately higher rates of child abuse than the rest of the state. Based on information from the Montgomery County Court of Common Pleas, between 2000 and 2002 there were 788 domestic violence cases with charges filed against a defendant, with the majority of these cases occurring in the city of Dayton (Artemis Center for Domestic Violence Alternatives, 2004). In addition, research indicated high levels of child and adolescent exposure to violence in Dayton and Montgomery County (Artemis Center for Domestic Violence Alternatives, 2004).

Prior to Safe Start, there were limited resources in the Dayton community for young children exposed to violence. To address this need, the Dayton Safe Start project came together as a collaborative endeavor with the Artemis Center as the lead agency, along with Brighter Futures (a home-based nurse visitation program) and the Young Children's Assessment and Treatment Services (YCATS) program of Samaritan Behavioral Health, Inc. While each partner had worked separately with the other organizations in the past, this was the first time that all parties had worked in collaboration. The objectives of this partnership were to increase the community's capacity to provide mental health services to infants and young children who had been exposed to violence, specifically domestic violence, and to offer those services in conjunction with services for their mothers. Dayton's Safe Start program involved Artemis offering CPP and case management in the client's home to families referred from the Brighter Futures nurse home visiting program. In earlier evaluations of CPP, CPP showed medium intervention effects on posttraumatic stress disorder (PTSD) symptoms and behavior problems (0.63 and 0.64, respectively; Lieberman, Van Horn, and Ghosh Ippen, 2005). However, this particular mix of services had not been evaluated previously. YCATS provided training and clinical supervision to Artemis staff providing Safe Start services and collected data for the evaluation.

DAYTON SAFE START

- **Intervention type:** CPP and case management
- **Intervention length:** 6–12 months (CPP); up to the end of the two-year evaluation data collection (case management)
- **Intervention setting:** In-home
- **Target population**: Children exposed to domestic violence
- **Age range:** 0–5
- **Primary referral source:** Nurse home visiting program

The outcomes evaluation detailed here presents data relevant to the question of whether the Dayton Safe Start program, as implemented within this project, improves outcomes for children exposed to violence.

INTERVENTION

The Dayton Safe Start program involved two main components: intensive, long-term CPP and case management through an advocate. The intervention period lasted approximately one year. All of the services were provided in the client's home. The program elements are described briefly in the following paragraphs. For a full description of the Dayton intervention as it was delivered, see Schultz et al. (2010).

The therapy component involved parent-child dyadic therapy using the Lieberman model. CPP is a relationship-based intervention designed for use with children up to age 6. It can be used with any child whose relationship to his or her parent or other primary caregiver is impacted by negative circumstances, including family violence. CPP integrates psychodynamic, attachment, trauma, cognitive-behavioral, and social learning theories (NCTSN, 2008). There are two components in CPP: assessment and treatment, with information gained during the assessment used to inform the intervention component. In the intervention component, child-parent interactions are the focus of six intervention modalities aimed at restoring a sense of mastery, security, and growth and promoting congruence between bodily sensations, feelings, and thinking on the part of both child and parent and in their relationship with one another (NCTSN, 2008).

The Dayton site delivered weekly, one-hour CPP sessions in the home. In implementing CPP, the Safe Start child therapist worked on identifying and talking about the abuse using CPP techniques to help the parent understand the child's behavior and symptoms. Therapy continued until 75 percent of the treatment goals were met, usually for about six months. If 75 percent of the treatment goals were not met, there was a possibility of sessions for six more months, as long as adequate progress had been made and there had been regular participation in the sessions.

The case management component was made available for each family throughout its therapy and beyond that for up to the two years of data collection for the evaluation. The therapist conducted an initial home visit in conjunction with the advocate who provided the case management, if possible. The case management activities included assistance with housing, employment, and transportation issues. The advocate also incorporated domestic violence education into the case management.

Efforts to monitor the quality of the program included initial clinician training for both the child therapist and advocate that involved discussion of the books by one of the model developers and related articles. The therapist had access to these materials, as well as an infant mental health handbook. A clinical supervisor from YCATS served as both a supervisor and mentor for the therapist, meeting weekly to review the session protocols and analyze how the sessions went. The Safe Start project director provided supervision to the advocate through weekly meetings.

METHOD

Design Overview

The design of this study was a randomized controlled trial, with randomization occurring at the family level and eligible children recruited after families were referred to the program. In addition to receiving usual support services from the Brighter Futures nurse home visiting program (e.g., in-home visits, referrals for therapy and advocacy services), the intervention group received CPP for up to 12 months and case management for up to two years. Families in the control group continued to receive the usual support services as well. For both groups, child outcomes and contextual information were assessed

at baseline, six, 12, 18, and 24 months. Study enrollment took place between July 2006 and March 2009.

Evaluation Eligibility Criteria

Initially, Dayton Safe Start relied solely on the Brighter Futures nurse home visiting program for determining eligibility and making referrals to the program. Eligibility was determined during the Brighter Futures intake process, which included a questionnaire on the mother's relationships and domestic violence. The protocol called for Brighter Futures nurses to administer the questionnaire and to refer the case to Safe Start if they detected domestic violence. In addition, families were required to be proficient in English. Partway through implementation, Dayton added several community-based agencies as referral sources. For these referrals, Artemis determined eligibility after receiving the referral.

When more than one child in the eligible age range (ages 0–5) was eligible for the program by virtue of exposure to violence, the youngest child with the closest birth date to the date of enrollment was selected as the target child for the evaluation.

Randomization Procedures

On enrollment into the study, the children were randomized into intervention or control groups using a block randomization procedure that allowed for approximately the same number of children in the intervention and control groups (see Chapter Four of the main document [http://www.rand.org/pubs/technical_reports/TR991-1.html]). Because of the possibility that the impact of the intervention could differ by child age, the sample was stratified into two groups. One group of children was recruited from birth up to 2 years of age, and the second group of children was between 3 and 5 years old.

Measures

The measures used in this study are described fully in Chapter Two of the main document (see http://www.rand.org/pubs/technical_reports/TR991-1.html). The measures were uniform across the national evaluation but prioritized within each site as to the relevance to the intervention under study. Based on the nature of the Dayton Safe Start intervention, the outcomes were prioritized as shown in Table 1.

Table 1
Prioritized Outcome Measures for Dayton Safe Start

Primary Outcome Measures			
Domain	***Source/Measure***	***Age of Child***	***Respondent***
PTSD Symptoms	Trauma Symptom Checklist for Young Children	3–5 years	Caregiver
Behavior/Conduct Problems	BITSEA and Behavior Problem Index	1–5 years	Caregiver
Social-Emotional Competence	ASQ	0–2 years	Caregiver
Social-Emotional Competence	BITSEA and SSRS (Assertion and Self-Control)	1–5 years	Caregiver
Caregiver-Child Relationship	Parenting Stress Index	All	Caregiver
Violence Exposure	Juvenile Victimization Questionnaire Caregiver Victimization Questionnaire	All	Caregiver
Secondary Outcome Measures			
Domain	***Source/Measure***	***Age of Child***	***Respondent***
Social-Emotional Competence	SSRS (Cooperation)	3–5 years	Caregiver
Tertiary Outcome Measures			
Domain	***Source/Measure***	***Age of Child***	***Respondent***
School Readiness/Performance	Woodcock-Johnson III	3–5 years	Child
Background and Contextual Factors	Everyday Stressors Index	All	Caregiver

NOTE: ASQ = Ages and Stages Questionnaire, BERS-2 = Behavior and Emotional Rating Scales—2, BITSEA = Brief Infant-Toddler Social and Emotional Assessment, SSRS = Social Skills Rating System.

Enrollment and Retention

Dayton Safe Start received almost all of its referrals from the Brighter Futures Nurse Home Visiting Program. Toward the end of implementation, Dayton expanded its referral sources to include internal referrals from Artemis and outside referrals from community-based agencies serving the target population. Once a referral was received, the project director assigned the case to a data collector so that the baseline assessment could be scheduled and completed in the client's home. After the assessment, the project director implemented the random assignment procedures and informed the referring party about the results.

According to data submitted on its Quarterly Activity Reports, Dayton Safe Start enrolled 67 percent of the families referred to the program. The most common reasons that families did not enroll included caregiver-related issues, such as lack of interest (67 percent) and lack of time (7 percent), inability to locate (7 percent), and other issues (20 percent).

In Table 2, we present the number and percentage of all enrollees who were eligible for participation at each data collection time point. Dayton program staff enrolled 55 families in the study and completed six-month assessments for 45 percent of caregivers and 67 percent of children. The resulting six-month analysis sample consists of 15 families in the control group (52-percent retention) and ten families in the intervention group (38-percent retention). For subsequent assessments, Dayton retained only 13 to 31 percent of the families overall.

Table 2
Retention of Dayton Enrollees Eligible to Participate in Assessments at Each Time Point

	Caregiver Assessment				Child Assessment			
	Six Months	12 Months	18 Months	24 Months	Six Months	12 Months	18 Months	24 Months
Intervention								
Received	15	6	7	7	2	1	2	4
Expected*	29	17	14	13	4	7	7	9
Retention Rate	52%	35%	50%	54%	50%	14%	29%	44%
Control								
Received	10	4	1	1	2	3	0	1
Expected*	26	16	13	13	2	12	9	9
Retention Rate	38%	25%	8%	8%	100%	25%	0%	11%
Overall								
Retention Rate	45%	30%	30%	31%	67%	21%	13%	28%

* The number of expected assessments for longer-term assessments differs from the number who entered the study because the field period for collecting data in this study ended in the fall of 2009, before all families entered the window of time for assessments at 12, 18, or 24 months.

NOTE: Dayton's low retention at the six-month follow-up assessment increases the potential for biased results. This degree of attrition may be related to treatment factors that lead to selection bias. For example, if families in more distress are more likely to leave the study and be lost to follow-up, then the results can be misleading.

Special Issues

Dayton struggled with enrollment throughout the project. The Brighter Futures nurses worked closely with the target population, but this relationship-building took time, making it difficult for them to introduce the topic of domestic violence. In order to refer more consistently, Safe Start program staff felt that the nurses needed to learn how to identify and assess for domestic violence. Some

were more willing and comfortable with domestic violence issues than others. For a more in-depth discussion, see Schultz et al. (2010).

Analysis Plan and Power Calculations

First, we conducted descriptive analyses to summarize the sample characteristics: age, gender, race or ethnicity, the family income level, and the child's violence exposure at baseline. Because this was a randomized experimental design, we did not expect any major differences between the two groups at baseline. However, to test this possibility, we compared child and caregiver characteristics of the intervention and control groups using t-tests and chi-square tests.

To assess the effect of the Safe Start intervention, we primarily examined differences between children in the intervention and control groups at six months. 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 control groups divided by the common standard error. 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 (Cohen, 1988). With only 25 children (or caregivers) observed at both baseline and six months (15 in intervention and ten in control), at the nominal 0.05 significance level we have a 7.5-percent chance to detect a small intervention effect, a 21.7-percent chance to detect a medium intervention effect, and a 46.7-percent chance to detect a large intervention effect in this study. CPP has been demonstrated to have medium effects on PTSD symptoms and behavior problems (0.63 and 0.64, respectively) in prior studies (Lieberman, Van Horn, and Ghosh Ippen, 2005). Unfortunately, with such a small sample size, we only have a 21.7-percent chance of statistically detecting the medium effect size observed in prior studies. Further, these data will only be able to detect a medium intervention effect at the nominal 80-percent power level. Statistical power was dampened by several factors other than overall sample size. 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. The low power in this study must be kept in mind when interpreting the results.

We examined differences between the intervention and control 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 Four of the main document (see http://www.rand.org/pubs/technical_reports/TR991-1.html), comparisons of a control group only to those who complete services (or receive a predetermined amount of services) is likely to bias results. That is, those that do not engage in services or drop out prior to completion may be systematically different than those who remain. Ideally, analyses would take into account the type and amount of services received to account for dosage variability. However, there were not enough families in this site's sample in order to proceed with this type of analysis. Thus, the findings presented here on the entire intervention sample may obscure important subgroup differences by service dose received.

To examine differences in outcomes between the intervention and control 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 five or more. We compare means within groups across time using t-tests, compare groups with t-tests or chi-square tests at each assessment point, and examine difference in differences to compare the two groups on mean changes between baseline and the six-month assessment (when the sample size is at least ten per group). The sample size in the six-month analysis sample does not allow for further modeling of the differences in differences over time.

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 False Discovery Rate (FDR) method (Benjamini and Hochberg, 1995). Our assessments of statistical significance were based on applying the FDR procedure separately to all of the primary, secondary, and tertiary outcome tests in this report using an FDR of 0.05. For instance, with 15 model test statistics conducted among the primary outcomes, this led to adopting a statistical significance cutoff of 0.018 for the within intervention

group results and 0.023 for the within control group results. In the discussion of results, we have also identified nonsignificant trends in the data, defined as those tests with p-values of less than 0.05 but not exceeding the threshold established using the FDR method to adjust for multiple significance tests. While these trends may suggest a practical difference that would be statistically significant with a larger sample size, 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.

RESULTS

Baseline Descriptive Statistics

For the descriptive statistics, we provide the characteristics for the full enrolled sample at baseline. As seen in Table 3, the baseline sample was composed of 53 percent males, with an average age of 1.7 years. Forty-nine percent of the sample was black, with some white (26 percent), Hispanic (4 percent), and other race/ethnicity children (22 percent). The vast majority (92 percent) of families had family incomes of less than $30,000, with 42 percent of them having family incomes of less than $5,000. According to the caregiver reports, children in the baseline sample had been exposed to an average of 2.7 types of violence in their lives prior to the baseline assessment. All of the caregivers were the parent or guardian of the child. As noted in the table, there were no differences for these characteristics between the intervention and control groups at baseline.

In the sample of families retained at six months, the demographics were similar to those at baseline, with slightly more females (52 percent) and whites (36 percent). Again, there were no differences at baseline between groups in the sample retained at six months (data not shown).

Table 3
Dayton Safe Start Sample Characteristics for Families in the Baseline Assessment Sample

	Combined		Intervention		Control		Test for Comparison P-Value
Child Characteristics	***N***	***Mean***	***N***	***Mean***	***N***	***Mean***	
Age	55	1.7	29	1.8	26	1.6	0.68
CR Violence Exposure	55	2.7	29	2.5	26	2.9	0.48
Gender	***N***	***%***	***N***	***%***	***N***	***%***	
Male	29	52.7	14	48.3	15	57.7	0.49
Female	26	47.3	15	51.7	11	42.3	
Race/Ethnicity	***N***	***%***	***N***	***%***	***N***	***%***	
White	14	25.5	8	27.6	6	23.1	
Black	27	49.1	16	55.2	11	42.3	
Hispanic	2	3.6	0	0.0	2	7.7	
Other	12	21.8	5	17.2	7	26.9	
Caregiver Characteristics	***N***	***%***	***N***	***%***	***N***	***%***	
Family Income Level							
Less than $5,000	21	42.0	12	44.4	9	39.1	
$5,000–$10,000	13	26.0	10	37.0	3	13.0	
$10,001–$15,000	5	10.0	3	11.1	2	8.7	
$15,001–$20,000	3	6.0	0	0.0	3	13.0	
$20,001–$30,000	4	8.0	2	7.4	2	8.7	
More than $30,000	4	8.0	0	0.0	4	17.4	
Relationship to Child							
Parent or Guardian	55	100	29	100	26	100	
Other Relationship	0	0	0	0	0	0	

NOTES: CR = Caregiver Report. Percentages may not total 100 percent because of rounding.

Next, we examined the Dayton sample at baseline on the caregiver-child relationship outcome to understand the level of severity on this measure among families entering the study. Because most of the children in the Dayton sample were less than 3 years old, relatively few were assessed using the PTSD measure, so those baseline scores are not presented. For the caregiver-child relationship, 40 percent of the sample had total stress levels that fell in the clinical range, with 41 percent for boys and 38 percent for girls. For the different subscales, 38 percent of the sample had clinical levels on the parental distress subscale, 25 percent had clinical levels on the parent-child dysfunctional interaction subscale, and 33 percent had clinical levels on the difficult child subscale.

Table 4
Baseline Assessment Estimates for Dayton Safe Start Families

CR Total Parenting Stress for Ages 0–12	Combined		Boys		Girls	
	N	%	*N*	%	*N*	%
Parental Distress—Clinical	21	38	12	41	9	35
Parent-Child Dysfunctional Interaction—Clinical	14	25	6	21	8	31
Difficult Child—Clinical	18	33	9	31	9	35
Total Stress—Clinical	22	40	12	41	10	38

NOTE: CR = Caregiver Report.

We also examined differences between the intervention and control group at baseline for Dayton's primary and tertiary outcomes (see this report's appendix). Primary outcomes include PTSD symptoms, behavior problems, social-emotional competence, caregiver-child relationship, and violence exposure. There were no differences observed between the intervention and control groups at baseline for any of the primary outcomes with a large enough sample size (Table A.1). Dayton only had one secondary outcome, which was one aspect of the social-emotional competence domain (cooperation) for 3–12-year-old children. Since the sample size was fewer than five per group at the baseline assessment, we did not compare the two groups on this outcome. Dayton's tertiary outcomes included only the background and contextual factor domain and the school readiness/performance domain (Table A.2). No differences were observed at baseline on the tertiary outcomes in the background and contextual factors domain. None of the school performance/readiness measures had an adequate sample size to compare means at baseline.

Uptake, Dosage, and Process of Care

As described fully in the Safe Start process evaluation report (Schultz et al., 2010), the Dayton Safe Start intervention services included CPP, case management, and multidisciplinary team meetings. Family-level service data were recorded by the program on the follow-up Family Status Sheet and submitted at six-month intervals following initial enrollment (see Chapter Two of the main document [http://www.rand.org/pubs/technical_reports/TR991-1.html] for a description). Tables 5a and 5b below show the type and amount of services received by the families assigned to the intervention group. The data displayed include services received by summing all time points reported by the program, with a maximum of 24 months of service provision.

As shown in Table 5a, 59 percent of these families received CPP, with an average of 19.5 sessions per family. About two-thirds received case management services, and 55 percent had multidisciplinary team involvement. Dayton reported information on the reason that the services ended for only seven of the 29 intervention group families.

Table 5a
Services Received by Dayton Safe Start Intervention Families (Baseline Assessment Sample)

Service	Number with Service	Percentage with Service*	Range	Distribution	Mean	Median
Dyadic Therapy (CPP)	17	59%	1–76	1–5 29% 6–10 18% 11–20 18% >20 35%	19.5	5.1
Case Management	18	62%	1–55	1–5 33% 6–10 22% 11–20 17% >20 28%	15.4	3.8
Multidisciplinary Team	16	55%	1–44	1–5 63% 6–10 6% 11–20 13% >20 19%	9.6	3.2

* The denominator is the 29 intervention group families with a follow-up Family Status Sheet at the six-month assessment point.
NOTE: Percentages may not total 100 percent because of rounding.

Table 5b shows the services received by that subgroup of intervention group families who participated in the six-month follow-up research assessment. These are the 15 families included in the intervention group in the six-month analysis sample for the Dayton program. Table 5b shows the services they received within the six-month period between baseline and the six-month assessment. A majority (80 percent) of these families received CPP, with an average of 12.9 sessions between baseline and the six-month assessment. Almost one-half (47 percent) received case management services, and more than one-half (53 percent) had multidisciplinary team involvement. Service ending data were reported on only one of the 15 families shown in Table 5b.

Table 5b
Six-Month Services Received by Dayton Safe Start Intervention Families in the Six-Month Assessment Sample

Service	Number with Service	Percentage with Service*	Range	Distribution	Mean	Median
Dyadic Therapy (CPP)	12	80%	1–37	1–5 33% 6–10 17% 11–20 25% >20 25%	12.9	3.2
Case Management	7	47%	1–38	1–5 43% 6–10 29% >20 29%	11.4	5.1
Multidisciplinary Team	8	53%	1–24	1–5 63% 11–20 13% >20 25%	8.5	3.6

* The denominator is the 15 intervention group families in the six-month assessment sample.
NOTE: Percentages may not total 100 percent because of rounding.

Outcomes Analysis

We begin by comparing the intervention and control groups' mean scores on primary, secondary, and tertiary outcomes at each follow-up assessment point (six, 12, 18 and 24 months) when the sample size allowed. We then look at changes in mean scores over time. For these analyses, we first tested whether there were statistically significant changes in mean scores within the intervention group and within the control group separately. Then, at the six-month assessment point, we compared the mean score change of the two groups to determine if there were statistically significant differences in mean changes. In these analyses, we used an intent-to-treat approach that included all families in the intervention group, regardless of the level of service they received. We were unable to compare groups after the six-month assessment point because of inadequate sample sizes (i.e., fewer than ten per group).

Comparison of Means Between Groups

A summary of differences between the intervention and control group at each follow-up assessment point for Dayton's primary, secondary, and tertiary outcomes is depicted in this report's appendix. Very few of the outcome measures had enough data to compare the groups at any of the follow-up assessment points. There were no statistically significant differences between the groups at any follow-up assessment point on any of the outcome measures that could be tested.

Mean Differences over Time

Table 6 shows differences over time for Dayton's primary outcomes. The second column of numbers in Table 6 shows the mean change between the baseline and six-month score for each individual family. For the intervention group, the analyses revealed a statistically significant within-group difference from baseline to six months for the total number of caregiver traumatic experiences. Between baseline and the six-month follow-up, caregivers in both groups reported significantly fewer total child victimization experiences, child witnessing violence experiences, caregiver non–domestic violence experiences, and caregiver domestic violence experiences. The decreases in the child's and caregiver's victimization were expected because of different reference periods for the baseline and six-month assessments. Within the control group, there was also a significant decrease in the total parenting stress scale at six months. Within the intervention group there was also one nonsignificant downward trend on the parental distress scale. However, because of the multiple significance tests being conducted, this trend did not reach statistical significance, and this may be due to chance. At the 12-month assessment, caregivers in both groups continued to report significantly fewer domestic violence experiences and non–domestic violence experiences. These were the only two outcomes with an adequate sample size for the significance test. At the 18- and 24-month assessment points, the sample size fell below the threshold of at least ten cases per group for significance tests.

The third column in Table 6 show the results of the comparison of the intervention group's mean change in scores from baseline to six months to the control group's mean change in scores using the statistical test of differences in differences. The results revealed no statistical differences between the intervention and control groups for any of the primary outcomes with a large-enough sample size to report the results. Sample sizes were inadequate to proceed with comparisons in subsequent time periods.

Dayton only had one secondary outcome, which was one aspect of the social-emotional competence domain (cooperation) for 3–12-year-old children. For this outcome, the sample size was fewer than five per group at six months, so no changes over time are reported here.

Table 6
Changes in Means for Primary Outcome Variables Between Baseline and Six-Month Assessment

Primary Outcome	Group	N	Within-Family Mean Changes[a]	Group-Level Comparison of Mean Changes (Unadjusted Model)[b]
Behavior/Conduct Problems				
CR Child Behavior Problems for Ages 1–18	Intervention Control	12 8	–0.10	
Social-Emotional Competence				
CR Child Assertion for Ages 1–12	Intervention Control	12 8	0.12	
CR Child Self-Control for Ages 1–12	Intervention Control	12 8	–0.08	
Caregiver-Child Relationship				
CR Parental Distress for Ages 0–12	Intervention Control	15 10	–5.40 # –4.10	–1.30
CR Parent-Child Dysfunction for Ages 0–12	Intervention Control	15 10	–0.87 –2.50	1.63
CR Difficult Child for Ages 0–12	Intervention Control	15 10	–1.13 –2.30	1.17
CR Total Parental Stress for Ages 0–12	Intervention Control	15 10	–7.40 –8.90 *	1.50
Violence Exposure				
CR Total Child Victimization Experiences for Ages 0–12	Intervention Control	15 10	–1.40 * –2.20 *	0.80
CR Child Maltreatment for Ages 0–12	Intervention Control	15 10	–0.27 –0.60	0.33
CR Child Assault for Ages 0–12	Intervention Control	15 10	–0.13 –0.40	0.27
CR Child Sexual Abuse for Ages 0–12	Intervention Control	14 10	0.00 0.00	0.00
CR Child Witnessing Violence for Ages 0–12	Intervention Control	15 10	–1.27 * –1.30 *	0.03
CR Caregiver Total Number of Traumatic Experiences	Intervention Control	15 10	–0.67 * –0.60	–0.07
CR Caregiver Experience of Any Non-DV Traumas[c]	Intervention Control	15 10	–0.27 * –0.60 *	0.33

Table 6—continued

Primary Outcome	Group	N	Within-Family Mean Changes[a]	Group-Level Comparison of Mean Changes (Unadjusted Model)[b]
CR Caregiver Experience of Any Domestic Violence[c]	Intervention	15	–0.53 *	0.27
	Control	10	–0.80 *	

[a] This column reflects within-family mean changes between the baseline and six-month scores for each group separately. * indicates a significant paired t-test of differences over time.
[b] This column reflects the group-level comparison of within-family mean changes from baseline to six months. * indicates a significant t-test of group differences.
[c] The outcome is a categorical variable, and the unadjusted within-family mean change and the group-level comparison are changes in proportion.
NOTES: CR = Caregiver Report. # indicates a nonsignificant trend in the t-test ($p<0.05$ but does not meet the FDR correction threshold). 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.

Table 7 shows the difference over time for Dayton's tertiary outcomes in the background and contextual factor domain. The sample sizes for school readiness/performance were inadequate to examine differences over time. In the background and contextual factor domain, no significant differences were observed between the groups at this time point. Sample sizes were inadequate to proceed with comparisons in any of the subsequent time periods.

Table 7
Changes in Means for Tertiary Outcome Variables Between Baseline and Six-Month Assessment

Tertiary Outcome	Group	N	Within-Family Mean Changes[a]	Group-Level Comparison of Mean Changes (Unadjusted Model)[b]
Background and Contextual Factors				
CR Caregiver Resource Problems	Intervention	15	–1.13	–1.33
	Control	10	0.20	
CR Caregiver Personal Problems	Intervention	15	–0.93	0.67
	Control	10	–1.60	

[a] This column reflects within-family mean changes between the baseline and six-month scores for each group separately. * indicates a significant paired t-test of differences over time.
[b] This column reflects the group-level comparison of within-family mean changes from baseline to six months. * indicates a significant t-test of group differences.
NOTES: CR = Caregiver Report. 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.

CONCLUSIONS

Via its implementation of CPP, Dayton's Safe Start program was able to provide an established therapeutic intervention to very young children with exposure to domestic violence in their homes. The evaluation consisted of a randomized controlled trial of the intervention. Dayton program staff enrolled 55 families in the study, with less than one-half of them (45 percent) retained for the six-month assessment. The difficulties in establishing a steady flow of referrals from the primary referring agency resulted in enrollment considerably below expectations. Most of the participants in the study were minorities and impoverished. The participants in the study had substantial violence exposure, with caregivers reporting that these very young children (average age 1.7 years) had been exposed to an average of 2.6 instances of violence in their lives prior to the baseline assessment. At baseline, families enrolled in the study were experiencing parental stress, with the caregiver for 40 percent of the families reporting levels of parental stress that fell in the clinical range.

Families in Dayton's intervention group received CPP, case management, or multidisciplinary team services. Eighty percent of the families in the six-month analysis sample received CPP, 53 percent had multidisciplinary team involvement, and 47 percent received case management. On average, these intervention families received 12.9 sessions of CPP. Dayton's advocate averaged 11.4 case management contacts per family, and the multidisciplinary team met an average of 8.5 times per family. Our process evaluation found that the intervention setting was challenging because the home-based setting of service delivery came with distractions that made it difficult for the therapist to establish trust with the family and begin to work on the caregiver-child relationship (see Schultz et al., 2010).

Given the number of participants in this study, we had only a 32-percent chance to detect a medium effect of size 0.63 at six months, when intervention services were still ongoing. We expected a medium intervention effect, given what had been commonly observed in other studies of CPP. Thus, the evaluation design did not have adequate statistical power to detect an intervention effect, should one exist.

Despite mean scores in the intervention and control groups in the expected directions, the intent-to-treat analyses found no differences between groups over time. Compared to similar children who did not receive the Safe Start intervention, the

results do not indicate that the Dayton Safe Start program improved outcomes for children. Overall, the sample size limitations mean that no conclusions can be drawn about the effectiveness of the Dayton intervention as implemented on child and family outcomes. A primary explanation is the small sample size. Tests of some measures could not be conducted at all, while with other tests the statistical power was inadequate to detect anything less that a large intervention effect. Given that the expected intervention effect size was medium, a larger sample size might have allowed for the detection of a statistically significant intervention effect. The evaluation ended early because of funding constraints when the appropriation for Safe Start was curtailed, which may have affected the sample size.

By selecting an established therapeutic intervention, the partners in Dayton increased the local capacity to treat very young children who witness domestic violence. Overall, a small retained sample at each wave greatly hampered our ability to comment on the effectiveness of this approach in improving outcomes for children. The Dayton Safe Start model requires further testing with adequate sample sizes to statistically assess whether the intervention positively affects outcomes for very young children who have been exposed to domestic violence.

REFERENCES

Artemis Center for Domestic Violence Alternatives, *Funding Proposal to the Office of Juvenile Justice and Delinquency Prevention—CFDA [Catalog of Federal Domestic Assistance]: Safe Start: Promising Approaches for Children Exposed to Violence*, Dayton, Ohio: Artemis Center for Domestic Violence Alternatives, 2004.

Benjamini, Y., and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society, Series B*, Vol. 57, 1995, pp. 289–300.

Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc., 1988.

Lieberman, A. F., P. Van Horn, and C. Ghosh Ippen, "Toward Evidence-Based Treatment: Child-Parent Psychotherapy with Preschoolers Exposed to Marital Violence," *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 44, 2005, pp. 72–79.

National Child Traumatic Stress Network, "CPP: Child Parent Psychotherapy," Raleigh, N.C., 2008. As of July 19, 2011: http://www.nctsnet.org/nctsn_assets/pdfs/promising_practices/cpp_general.pdf

NCTSN—*see* National Child Traumatic Stress Network.

Schultz, D., L. H. Jaycox, L. J. Hickman, A. Chandra, D. Barnes-Proby, J. Acosta, A. Beckman, T. Francois, and L. Honess-Morealle, *National Evaluation of Safe Start Promising Approaches: Assessing Program Implementation*, Santa Monica, Calif.: RAND Corporation, TR-750-DOJ, 2010. As of July 17, 2011: http://www.rand.org/pubs/technical_reports/TR750.html

DAYTON OUTCOMES APPENDIX

Table A.1
Comparison of Means for Dayton Primary Outcome Variables over Time

Primary Outcome		Baseline		Six Months		12 Months		18 Months		24 Months	
		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
PTSD Symptoms											
CR Child PTSD Symptoms for Ages 3–10	Intervention	5	44.00	4		0		4		6	38.83
	Control	3		2		3		1		1	
Behavior/Conduct Problems											
CR Child Behavior Problems for Ages 1–18	Intervention	22	0.27	12	0.18	6	0.00	7	0.04	7	–0.01
	Control	19	0.25	8	–0.26	4		1		1	
Social-Emotional Competence											
CR Child Personal-Social Competence for Ages 0–2	Intervention	7	45.86	0		0		0		0	
	Control	5	47.00	0		0		0		0	
CR Child Assertion for Ages 1–12	Intervention	22	0.14	12	0.05	6	0.76	7	–0.30	7	–0.19
	Control	19	0.11	8	0.30	4		1		1	
CR Child Self-Control for Ages 1–12	Intervention	22	0.00	12	–0.09	6	0.40	7	–0.08	7	–0.02
	Control	19	–0.10	8	0.01	4		1		1	
Caregiver-Child Relationship											
CR Parent Distress for Ages 0–12	Intervention	29	31.45	15	27.00	6	29.50	7	28.00	7	25.86
	Control	26	32.12	10	28.80	4		1		1	

Table A.1—continued

Primary Outcome		Baseline		Six Months		12 Months		18 Months		24 Months	
		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
CR Parent-Child Dysfunction for Ages 0–12	Intervention	29	21.62	15	21.33	6	23.17	7	24.43	7	21.86
	Control	26	22.92	10	22.50	4		1		1	
CR Difficult Child for Ages 0–12	Intervention	29	30.07	15	29.93	6	33.00	7	33.14	7	32.57
	Control	26	30.38	10	29.10	4		1		1	
CR Total Parenting Stress for Ages 0–12	Intervention	29	83.14	15	78.27	6	85.67	7	85.67	7	80.29
	Control	26	85.42	10	80.40	4		1		0	
Violence Exposure											
CR Total Child Victimization Experiences for Ages 0–12	Intervention	29	2.48	15	1.07	6	2.17	7	2.00	7	0.86
	Control	26	2.92	10	1.40	4	0.75	1		1	
CR Child Maltreatment for Ages 0–12	Intervention	28	0.39	15	0.27	6	0.50	7	0.43	7	0.29
	Control	26	0.69	10	0.50	4	0.25	1		1	
CR Child Assault for Ages 0–12	Intervention	29	0.24	15	0.07	6	0.33	7	0.57	7	0.14
	Control	26	0.38	10	0.10	4	0.00	1		1	
CR Child Sexual Abuse for Ages 0–12	Intervention	29	0.03	14	0.00	6	0.00	7	0.14	7	0.00
	Control	26	0.00	10	0.00	4	0.00	1		1	

Table A.1—continued

Primary Outcome		Baseline		Six Months		12 Months		18 Months		24 Months	
		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
CR Child Witnessing Violence for Ages 0–12	Intervention	27	1.78	15	0.53	5	1.60	7	1.00	6	0.33
	Control	25	1.84	10	0.80	4		1		1	
CR Caregiver Total Number of Traumatic Experiences	Intervention	29	0.62	15	0.07	6	0.33	7	0.57	7	0.29
	Control	26	0.46	10	0.30	4		1		1	
CR Caregiver Experience of Any Non-DV Trauma	Intervention	29	0.41	15	0.13	6	0.00	7	0.29	7	0.14
	Control	26	0.46	10	0.20	4		1		1	
CR Caregiver Experience of Any DV	Intervention	29	0.93	15	0.40	6	0.50	7	0.43	7	0.29
	Control	26	0.88	10	0.10	4		1		1	

* indicates statistically significant (p-value<FDR significance criterion); # indicates nonsignificant trend (p<0.05 and >FDR significance criterion).
NOTES: CR = Caregiver Report; DV = domestic violence. 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.

Table A.2
Comparison of Means for Dayton Tertiary Outcome Variables over Time

Tertiary Outcome		Baseline		Six Months		12 Months		18 Months		24 Months	
		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
Background and Contextual Factors											
CR Caregiver Resource Problems	Intervention	29	15.72	15	14.73	6	13.83	7	15.29	7	14.43
	Control	26	17.27	10	18.40	4		1		1	
CR Caregiver Personal Problems	Intervention	29	26.93	15	26.00	6	21.83	7	26.43	7	21.00
	Control	26	25.81	10	23.70	4		1		1	

* indicates statistically significant (p-value<FDR significance criterion); # indicates nonsignificant trend (p<0.05 and >FDR significance criterion).
NOTES: CR = Caregiver Report. 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.