



HEALTH and INFRASTRUCTURE, SAFETY, AND ENVIRONMENT

CHILDREN AND FAMILIES
EDUCATION AND THE ARTS
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INFRASTRUCTURE AND
TRANSPORTATION
INTERNATIONAL AFFAIRS
LAW AND BUSINESS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
TERRORISM AND
HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: [Jump to Page 1](#) ▼

Support RAND

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Health](#)

[RAND Infrastructure, Safety, and Environment](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

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 G: Kalamazoo County, Michigan

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

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2011 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (<http://www.rand.org/publications/permissions.html>).

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

KALAMAZOO COUNTY, MICHIGAN, SAFE START OUTCOMES REPORT

ABSTRACT

The Kalamazoo County Safe Start program developed an intervention for Head Start classrooms to improve outcomes for children (ages 3–5) who had been exposed to violence. The program included a classroom curriculum, training and consultation to Head Start staff, and groups for parents. 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 of this program consisted of a randomized controlled trial of the intervention, with randomization occurring at the classroom level. A total of 436 families were enrolled in the study, and 352 (81 percent) were retained for the six-month assessment. The participants in the study were functioning relatively well, with most of them scoring in the normal range on our child outcome measures, despite past exposure to violence. Among the families retained at six months, the majority of students in intervention classrooms were served with intervention elements as planned (95 percent), and about a third of parents participated in the optional parent groups. Given the sample size in this study, there was an 80-percent chance to detect effects of sizes of 0.30–0.41 at six months, depending on clustering, and to detect medium to large intervention effects at the subsequent follow-ups (0.41–0.68). We expected a small intervention effect, given that the intervention occurred at the classroom level and was compared with the regular Head Start curriculum. In the PTSD, social-emotional competence, and academic achievement/readiness domains, significant improvements in the intervention group were noted, but they were not different than changes observed in the control group. In the intent-to-treat analysis, no statistically significant differences between groups over time were detected. Similarly, there were no statistically significant differences noted among those that received a low, medium, or high dose of the intervention and comparable control families, or between participants in the parent groups and comparable controls. Although it is possible that the program as implemented in this study was not effective in improving outcomes for children in the intervention group relative to controls, there are several other possible explanations for the results presented here. The lack of difference between groups over time could reflect a sample size limitation and thus a lack of power to observe a small intervention effect, lack of room to improve for children who were not experiencing many symptoms or problems at baseline, increased recognition of symptoms as a function of the

intervention, or a lack of a meaningful difference between the intervention and the control arms on the particular outcomes measured. This study demonstrates the promise of working with children in this setting, with high participation rates and successful implementation of the curricular intervention.

INTRODUCTION

The Kalamazoo Safe Start program is located in Kalamazoo County, Michigan, where 16 percent of the population was below the poverty line in 2007. The Kalamazoo Safe Start program provided classroom curriculum, training and consultation to Head Start Staff, and groups for parents focused on improving outcomes for children exposed to violence in the County's Head Start program, because most students enrolled in the program were identified as low income or impoverished (Western Michigan University, 2004), and because research indicates a relationship between poverty and increased risk of exposure to violence and child abuse and neglect (Drake and Pandey, 1996; Slack et al., 2004). Moreover, the Kalamazoo Safe Start team noted a major gap in existing services for children exposed to violence since many of their developmental, emotional, or behavioral problems were not detected until the children entered school, with little or no opportunity for intervention earlier on (Western Michigan University, 2004). Thus, the main thrust of the Kalamazoo Safe Start program was to develop an intervention for pre-school-aged children enrolled in the Head Start program that would explicitly link behavior problems with violence exposure (Western Michigan University, 2004).

The team's Safe Start project builds on earlier work in which they developed and implemented a similar program for school-aged children (called the School Intervention Project) as part of a Substance Abuse and Mental Health Services Administration grant through the National Child Traumatic Stress Initiative. However, this program has not been formally evaluated. They also drew on concepts from the Classroom-Based Intervention (CBI; Macy, Bary, and Noam, 2003), which is designed to help older children aged 7–19 cope with disaster or an ongoing threat of terrorism. Thus, the Kalamazoo Safe Start project represents, to our knowledge, the first classroom-based intervention geared toward preschoolers exposed to violence and tailored specifically for the Head Start environment.

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

KALAMAZOO SAFE START

- **Intervention type:** Head Start School Intervention Program (a classroom intervention), teacher training, and parent training program
- **Intervention length:** 26 weeks
- **Intervention setting:** Head Start classrooms
- **Target population:** Children exposed to violence
- **Age range:** 3–5
- **Primary referral source:** Head Start

INTERVENTION

The Kalamazoo Safe Start program included a classroom-based intervention using a 26-week Head Start School Intervention Program (HSSIP) curriculum that was delivered by a team consisting of a teacher, aide, and bus driver. In addition, the intervention provided training and weekly consultation to the team members and an optional curriculum-based group for parents. These elements are described briefly in the following paragraphs. For a full description of the Kalamazoo intervention as it was delivered, see Schultz et al. (2010).

The intervention's manualized HSSIP school curriculum consisted of six core elements, called (1) feeling safe, (2) making and keeping friends, (3) calming mind and body, (4) feeling good and learning, (5) making meaning of experiences, and (6) literacy. Within the general framework provided by the six core elements, there were five specific "units" that described to teachers the concrete activities that they were to engage in with the students in the classroom. The five units within the curriculum were (1) learning about school, (2) learning about self, (3) learning about community, (4) learning about friendship, and (5) saying goodbye. Each unit contained structured activities intended to build both social/emotional skills and literacy. The activities were presented to the children as optional, offered as one possible activity that students could choose to participate in or not. Thus, not all students in the intervention classrooms took part in all HSSIP activities.

The intervention training for Head Start teachers and staff included a series of training slides and handouts designed to help Head Start staff understand the link between violence exposure and changes in behavior, emotions, physiology, and learning. In addition to a two-day training at the beginning of the school year delivered by Safe Start staff, ongoing consultation with school interventionists was planned to

occur weekly or biweekly in teams of teachers and aides who worked in each classroom. Central to this approach was the goal of improving the management of the child and the understanding that Head Start staff members have of their students—in essence, as stated by Safe Start staff, to go from seeing their difficult behavior as “being willfully disobedient” to seeing it as “striving for survival needs.”

The parent groups were designed to meet bimonthly for 12 90-minute meetings drawing on evidence-based approaches, such as Parent Child Interaction Therapy (Chaffin et al., 2004; Hood and Eyberg, 2003), The Incredible Years (Webster-Stratton, 1998; Webster-Stratton, 1994), the Sanctuary (Bloom, 2005), START (Benamati, 2002), and TARGET (Ford and Russo, 2006). They aimed to provide specific and age-appropriate information about psychological, emotional, behavioral, social, and academic problems associated with exposure and organized in collaboration with Kalamazoo Head Start staff.

Efforts to monitor the quality of the program included delivery of each of the intervention components by the developers of the component, with assistance by other staff. HSSIP was documented in a detailed manual, and the developers were assisted by graduate assistants. The program also used session adherence logs and attendance records to record activities, staff, and dates conducted. Supervision occurred live, following activities, and on a weekly basis. Parent group meetings were documented by handouts given to parents, log of attendance and incentives, and agendas for each meeting. Teacher training and consultation was observed and documented according to the slides used in the training.

METHOD

Design Overview

The design of this study is a cluster-randomized controlled trial, with randomization occurring at the Head Start classroom level and children recruited to participate from within the randomized classrooms. In classrooms randomized to the intervention, the regular Head Start curriculum was augmented by the intervention described above, whereas in the classrooms randomized to control, the Head Start curriculum alone was delivered. Child outcomes and contextual information were assessed at baseline, six, 12, 18, and 24 months. Study enrollment took place between September 2006 and October 2009 (spanning three consecutive school years).

Evaluation Eligibility Criteria

Consent forms were sent home to all students in participating classrooms. For those with parental consent, the following criteria were used to determine if a child was eligible to participate in the study: (1) child is enrolled in Kalamazoo County Head Start; (2) child speaks English or Spanish; (3) child is between the ages of 3 and 5 years; and (4) parent/caregiver endorses one indicator of child exposure to violence, where indicators included whether the child had seen another child harmed or threatened by an adult, had seen two or more adults fighting, had seen somebody arrested, had heard gunshots in the neighborhood, had heard an adult threaten to harm another adult, or feared being beaten up or shot. All children in the classroom could participate in the activities, not only the ones that had been exposed to violence. The only exclusion criterion was that siblings of program participants could not be enrolled in the study.

Randomization Procedures

Classrooms were randomized to the intervention or control condition in three waves, prior to the beginning of the school year in 2006, 2007, and 2008. In this case, a “classroom” consists of a three-person team (teacher, aide, and bus driver) that works with a group of students, and the classroom includes both a morning group of students and an afternoon group of students headed by the same team. The study was conducted over three years, with recruitment and baseline assessments occurring each year in the fall of a new school year, the intervention or regular Head Start services occurring over 26 weeks from December through June, and assessments occurring at six, 12, 18, and 24 months. Measures at each assessment are described in Chapter Two of the main document (see http://www.rand.org/pubs/technical_reports/TR991-1.html).

Each year, the classroom makeup and staffing were determined by Kalamazoo County Head Start, and then the list of assignments was given to the research team to allocate to intervention classrooms and control classrooms. In the first year of the study, seven classrooms (14 groups of students) were identified to participate in two locations. Randomization (within the site) resulted in four classrooms randomized to the intervention condition and three classrooms randomized to the control condition. In the second and third years of the study, assignment of classrooms to condition was limited, as staff members who had been trained in the first year to deliver the intervention were necessarily constrained to continue as staff in intervention classrooms, whereas staff who worked in the control classrooms and new staff could be assigned to either treatment or control classrooms. In Year 2, randomization of new classrooms and adherence to the first-year randomization were conducted for 11 classrooms in three

sites, resulting in five intervention classrooms and six control classrooms. In Year 3, the study took place in 14 classrooms in five locations, with seven assigned as intervention classrooms and seven assigned as control classrooms. Complete information on the randomization scheme is found in Table 1.

Table 1
Randomization of Classrooms in Kalamazoo Safe Start Study

School Year	Location	Total Classrooms	Intervention	Control
2006–2007	Site 1	4	2	2
	Site 2	10	6	4
2007–2008	Site 2	10	4	6
	Site 3	6	4	2
	Site 4	6	2	4
2008–2009	Site 2	10	4	6
	Site 4	6	4	2
	Site 5	4	2	2
	Site 6	4	2	2
	Site 7	4	2	2
Site Totals	Site 1	4	2	2
	Site 2	30	14	16
	Site 3	6	4	2
	Site 4	12	6	6
	Site 5	4	2	2
	Site 6	4	2	2
	Site 7	4	2	2
Grand Total		64	32	32

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. Given the nature of the Kalamazoo Safe Start intervention, the outcomes were prioritized as shown in Table 2.

Enrollment and Retention

Kalamazoo County Safe Start recruited Head Start classrooms to participate in the program. Recruitment for the project began with Head Start teachers, who handed out the consent forms with an introductory letter to parents at the beginning of the school year, collected back signed forms, and gave them to the Safe Start staff members. The random assignment procedures were conducted at the classroom level.

Table 2
Prioritized Outcome Measures for Kalamazoo Safe Start

Primary Outcome Measures			
<i>Domain</i>	<i>Source/Measure</i>	<i>Age of Child</i>	<i>Respondent</i>
PTSD Symptoms	Trauma Symptom Checklist for Young Children	All	Caregiver
Behavior/Conduct Problems	Behavior Problem Index	All	Caregiver
Social-Emotional Competence	BITSEA and SSRS (Assertion and Self-Control)	All	Caregiver
Social-Emotional Competence	SSRS (Cooperation)	All	Caregiver
School Readiness/Performance	Woodcock-Johnson III	All	Child
Secondary Outcome Measures			
<i>Domain</i>	<i>Measure</i>	<i>Age of Child</i>	<i>Respondent</i>
Caregiver-Child Relationship	Parenting Stress Index	All	Caregiver
Tertiary Outcome Measures			
<i>Domain</i>	<i>Measure</i>	<i>Age of Child</i>	<i>Respondent</i>
Violence Exposure	Juvenile Victimization Questionnaire	All	Caregiver
Violence Exposure	Caregiver Victimization Questionnaire	All	Caregiver

NOTE: BITSEA = Brief Infant-Toddler Social and Emotional Assessment, SSRS = Social Skills Rating System.

According to data submitted on its Quarterly Activity Reports, Kalamazoo Safe Start enrolled 61 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 or no time (32 percent), refusal to be randomized (16 percent), and other reasons that families cited for not enrolling in the program (41 percent).

In Table 3, we present the number and percentage of all enrollees who were eligible for participation at each data collection time point and the assessments that were available for analysis (at six months, two caregiver and child packet sets were not included because the dates fell outside the expanded window for data collection, and one child packet was not included because the lag between caregiver and child assessments was too long; see Chapter Four of the main document [http://www.rand.org/pubs/technical_reports/TR991-1.html]). Overall, retention rates were high, with more than 80 percent retained at six months in both the intervention and control groups and more than 60 percent retained at the other follow-up assessments.

Table 3
Retention of Enrollees Eligible to Participate in Assessments at Each Time Point

	Caregiver Assessment				Child Assessment			
	Six Months	12 Months	18 Months	24 Months	6 Months	12 Months	18 Months	24 Months
Intervention								
Received	188	99	92	46	185	88	86	45
Expected*	229	139	139	64	217	139	139	64
Retention Rate	82%	71%	66%	72%	85%	63%	62%	70%
Control								
Received	164	90	92	30	150	85	89	30
Expected*	207	123	123	36	189	121	121	36
Retention Rate	79%	73%	75%	83%	79%	70%	74%	83%
Overall								
Retention Rate	81%	72%	70%	76%	83%	67%	67%	75%

* 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 18 or 24 months.

Special Issues

In the Kalamazoo Safe Start outcomes evaluation, we encountered issues with students moving from one classroom to another after randomization. For those that moved from an intervention classroom to a control classroom, we removed data collected after the move. For those that moved from an intervention classroom to a control classroom, we preserved their data but account for the change by the number of days present in the intervention classroom. 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 baseline characteristics: age, gender, race or ethnicity, family income, child’s violence exposure at baseline, and the outcome variables. Because randomization occurred at the classroom level, there was a possibility of differences between the two groups (intervention and control) at baseline. We tested for differences in child and caregiver characteristics between intervention and control group children using t-tests and chi-square tests, controlling for clustering within classrooms.

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 that 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 352 children observed at both baseline and six months (188 in the intervention and 164 in the control group), and in the absence of any classroom cluster effect, we can expect only a 46.3-percent chance to detect a really small effect. However, because participants in the same classroom are likely to be similarly impacted by the treatment, accounting for intra-class correlation (ICC) reduces power to a 39.2-percent, 34.0-percent, and 27.2-percent chance to detect a really small effect for ICCs of sizes 0.05, 0.1, and 0.2, respectively.

The Kalamazoo Safe Start intervention occurred at the classroom level with little involvement from parents, and thus we expected a small effect size. Given the sample size in this study, there is an 80-percent chance to detect an effect of size 0.30 without clustering, or 0.32, 0.36, and 0.41, respectively (for ICCs of 0.05, 0.1, and 0.2), at six months. For the other follow-ups—12, 18, and 24 months with 189, 184, and 76 participants, respectively, remaining in the study—in the absence of any correlation within classrooms, we have an 80-percent chance to observe improvements in the outcomes as small as 0.410, 0.415, and 0.666, respectively, all of them small-to-medium effect sizes with the exception of the 24-month data. At the subsequent follow-ups and an ICC as high as 0.2, improvements as small as 0.484, 0.488, and 0.679 at 12, 18, and 24 months, respectively, can be detected with the nominal 80-percent chance. 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 statistical power in this study must be kept in mind in interpreting results.

We examined differences between the intervention and control groups using an intent-to-treat approach, which includes in analyses all 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 who receive a predetermined amount of services) is likely to bias results. That is, those who 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. We explore this issue of “dose” of intervention as described in the following paragraphs.

To examine differences between the intervention and control groups using the intent-to-treat approach described in the previous paragraphs, we present baseline and follow-up estimates of primary, secondary, and tertiary outcomes for both groups. We compare groups via chi-square or t-tests at each time point, compare means within groups across time, and examine difference in differences to comparing the two groups on changes over time between baseline and the six- and 12-month assessments. At each wave, we conducted multiple linear regressions on the continuous outcomes and linear probability regressions on the dichotomous different outcomes to test for the difference in difference via main effects and the interaction between intervention status and time after controlling for baseline characteristics (child race, gender, age, family income, and child’s exposure to violence at baseline). These baseline characteristics were selected to correct for any potential imbalance in the groups by relevant demographic characteristics. All of these regression analyses also account for clustering of children within classrooms, as the treatment was given at the classroom level. Even though data were collected in Kalamazoo for up to 24 months, special attention will be paid to the six-month follow-up regression, where the program was expected to have the greatest impact on participants, since the six-month follow-up occurred at the end of the school year in which the intervention took place.

To examine outcomes related to the as-treated sample, those families that took part in the intervention services offered, we examined the outcome means for families that took part in the intervention services offered, broken down into groups that received a low dose of the intervention, a medium dose, and a high dose, and a variable related to participation in the parent groups. Since children with more need may receive more services, we would expect that this selection scheme might 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 control 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 control groups for each dosage group at follow-up. Note that in this analysis, the full control 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.

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 example, with nine model test statistics conducted among the primary outcomes, this led to adopting a statistical significance cutoff of 0.006 in the covariate-adjusted difference in difference results. With only four secondary outcomes with enough sample sizes to allow for modeling, the FDR significance level adopted was 0.013, while such significance level was set at 0.006 for the tertiary outcomes (for eight outcomes). 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 shown in Table 4, the baseline sample was composed of 52 percent females, with an average age of 3.9 years (age range 2–4 years). The sample was ethnically and racially diverse, with 21 percent of caregivers describing their child as white, and 46 percent black. Most caregivers reported low family incomes. Caregivers reported that the children had been exposed to an average of 2.8 types of violence in their lives prior to the baseline assessment. As can be seen in the table, randomization at the classroom level did not entirely align the racial/ethnic composition of the intervention and control classrooms in this study, with apparently more nonwhite children in the intervention group as compared with the control group. We also examined differences between the groups among those families retained in the six-month sample, with the rationale that the most robust findings related to differences

between the intervention and control groups would apply to the six-month sample, assessed at the end of the school year, just after the intervention ended. Here the demographics were similar to those at baseline, and the difference between intervention and controls in terms of gender and race/ethnicity was still observed, but there were no differences in terms of family income (data not shown).

Table 4
Kalamazoo Safe Start Sample Characteristics for Families in the Baseline Assessment Sample

	Combined		Intervention		Control		Test for Comparison P-Values
<i>Child Characteristics</i>	<i>N</i>	<i>Mean</i>	<i>N</i>	<i>Mean</i>	<i>N</i>	<i>Mean</i>	
Age	436	3.9	229	3.8	207	3.9	0.60
CR Violence Exposure	436	2.8	229	2.7	107	2.8	0.70
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	
<i>Gender</i>							
Male	208	47.7	119	52.0	89	43.0	0.06
Female	228	52.3	110	48.0	118	57.0	
<i>Race/Ethnicity</i>							
Hispanic	45	10.3	17	7.4	28	13.5	0.01
White	90	20.6	37	16.2	53	25.6	
Black	202	46.3	118	51.5	84	40.6	
Other	99	22.7	57	24.9	42	20.3	
<i>Caregiver Characteristics</i>							
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	
<i>Family Income Level</i>							
Less than \$5,000	62	15.5	40	19.3	22	11.3	0.10
\$5,000–\$10,000	90	22.4	50	24.2	40	20.6	
\$10,001–\$15,000	86	21.4	46	22.2	40	20.6	
\$15,001–\$20,000	69	17.2	29	14.0	40	20.6	
\$20,001–\$30,000	51	12.7	22	10.6	29	14.9	
More than \$30,000	43	10.7	20	9.7	23	11.9	
<i>Relationship to Child</i>							
Parent-Guardian	421	97.5	221	97.8	200	97.1	0.64
Other Relationship	11	2.5	5	2.2	6	2.9	

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

We also examine the Kalamazoo sample at baseline on two outcomes (PTSD symptoms and parenting stress) to describe the level of severity on these indexes among families entering the project (Table 5). Most caregivers reported child PTSD symptoms that fell in the normal range, on average, for both boys and girls. Ninety-three percent of caregivers of boys reported child PTSD symptoms in the normal range, as did 83 percent of caregivers of girls. For the caregiver-child relationship, 30 percent of the sample had total stress levels that fell in the clinical range, with 30 percent for boys and 29 percent for girls. For the different subscales, 21 percent of the sample had clinical levels on the parental distress subscale, 22 percent had clinical levels on the parent-child dysfunctional interaction subscale, and 30 percent had clinical levels on the difficult child subscale.

Table 5
Baseline Assessment Estimates for Kalamazoo Safe Start Families

	Combined		Boys		Girls	
	N	%	N	%	N	%
CR PTSD Symptoms for Ages 3–10						
Normal	365	88	186	93	179	83
Borderline	16	4	5	3	11	5
Significant	33	8	8	4	25	12
CR Total Parenting Stress for Ages 0–12						
Parental Distress—Clinical	92	21	49	24	43	19
Parent-Child Dysfunctional Interaction—Clinical	94	22	49	24	45	20
Difficult Child—Clinical	128	30	61	30	67	30
Total Stress—Clinical	128	30	63	30	65	29

NOTE: CR = Caregiver Report.

Finally, we examined differences between the intervention and control group at baseline for Kalamazoo’s primary, secondary, and tertiary outcomes (see this report’s appendix). Primary outcomes include PTSD symptoms, behavior problems, social-emotional competence, and school readiness/performance (see Table A.1, first column). No statistically significant differences between groups were observed.

Secondary outcomes include the caregiver-child relationship. There were no statistically significant differences observed between the intervention and control group at baseline on secondary outcomes (see Table A.2) or tertiary outcomes (child and caregiver violence exposure; see Table A.3).

Uptake, Dosage, and Process of Care

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 6a and 6b below show the type and amount of services received by the families assigned to the intervention group. As described fully in the Safe Start process evaluation report (Schultz et al., 2010), Kalamazoo intervention services delivered to children occurred at the classroom level, and thus exposure to the classroom curriculum is equivalent to the number of days present in the classroom. Presence in the classroom does not necessarily mean that the child participated in the Safe Start curriculum exercises, but it is likely that he or she did, according to our process evaluation observations. Moreover, each day that the child was present in the classroom represents time interacting with the Head Start staff who had received initial training on the intervention and were receiving ongoing consultation to work more effectively in the classroom with children exposed to violence. Adult attendance in the caregiver group therapy sessions was optional.

Table 6a presents the results for 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 6a include services received by summing all time points reported by the program, with a maximum of 24 months of service provision. Service data for at least one follow-up wave were available on all of the 229 initial Kalamazoo intervention group enrollees.

As seen in Table 6a, the vast majority of students in the intervention classrooms received the curricular intervention (95 percent) and teacher consultation (96 percent), with an average of 49 days of Safe Start curriculum during the first school year (the first six months of Kalamazoo's Safe Start project) and an average of 14.5 of the weekly consultation meetings with staff. A smaller number of adults took part in the caregiver group therapy (33 percent), with an average of 3.5 sessions per caregiver. The program reported information on the reason that the school intervention services ended for 209 of the 229 intervention families. Only about 3 percent of families ended services earlier than anticipated (either changing schools or leaving the school); others completed the program as planned.

Table 6a
Services Received by Kalamazoo Safe Start Intervention Families (Baseline Sample)

Service	Number with Service	Percentage with Service*	Range	Distribution	Mean	Median
Classroom Curriculum	217	95%	2-121	2-15 8% 16-26 15% 27-36 12% 37-46 14% 47-57 17% 58-70 7% 71-82 10% 83-96 12% >96 4%	48.9	46.1
Consultation Meetings	219	96%	3-42	3-6 19% 8-11 17% 12-15 45% >17 19%	14.5	13.8
Parent Group	76	33%	1-17	1-3 66% 4-6 18% 7-9 9% >9 7%	3.5	1.7

* The denominator is the 229 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 6b shows the services received by that subgroup of intervention group families who participated in the six-month follow-up research assessment. These are the 188 families included in the intervention group in the outcome analyses sample for the Kalamazoo program. Table 6b shows the services they received within the six-month period between baseline and the six-month assessment.

As shown in Table 6b, families who completed the six-month assessment received very similar amounts of services as the full baseline sample, with the vast majority receiving the curriculum and staff consultation, and about a third participating in the parent groups. We received information on completion of classroom services on 165 of the 188 families that participated in the six-month research assessment. Of those, all but one family completed the intervention as planned, with that family dropping out of services prior to completion.

Table 6b
Six-Month Services Received by Kalamazoo Safe Start Intervention Families (Six-Month Analysis Sample)

Service	Number with Service	Percentage with Service*	Range	Distribution	Mean	Median
Classroom Curriculum	178	95%	2–92	2–15 9% 16–26 14% 27–36 13% 37–46 20% 47–57 19% 59–70 7% 71–82 11% 83–92 8%	45.6	43.7
Consultation Meetings	179	95%	3–27	3–6 18% 8–11 17% 12–15 46% >17 19%	13.6	13.4
Parent Group	61	32%	1–17	1–3 61% 4–6 21% 7–9 10% >9 8%	3.8	2.1

* The denominator is the 188 intervention group families with a follow-up Family Status Sheet at the six-month assessment point who participated in the six-month research assessment.

NOTE: Percentages may not total 100 percent because of rounding.

Outcomes Analysis

We begin by examining differences in primary, secondary, and tertiary outcomes between the intervention and control groups at each follow-up assessment point. We then analyze changes in mean scores over time both within the intervention and control groups and between the groups. For these analyses, we use an intent-to-treat approach that included all families randomized to the intervention, regardless of the level of service they received. Finally, we present descriptive data on families that received services, as compared with similar controls, on primary outcomes only.

Comparison of Means Between Groups

A summary of differences between the intervention and control group at each assessment point for Kalamazoo’s primary, secondary, and tertiary outcomes is depicted in this report’s appendix (Tables A.1, A.2, and A.3).

Primary outcomes include PTSD symptoms, behavior problems, social-emotional competence, and school readiness/performance. At the six-month follow-up assessment, there were no statistically significant differences on any of the primary outcome measures. However, an observable nonsignificant trend was noted, in which groups differed on caregiver report of child assertion, with caregivers in intervention classrooms reporting less assertion than those in control classrooms, similar to the

baseline assessment difference. Similarly, at the 18-month follow-up assessment, children in the intervention group performed significantly better on the passage comprehension test, as compared with children in the control group. We cannot rule out that either of these differences between the groups may be due to chance, however, because of the multiple significance tests being conducted (i.e., the group difference did not exceed the more stringent statistical criterion set using the FDR method). There were no statistically significant differences between the intervention and control groups at the 12-month or 24-month assessment points (see Table A.1).

For secondary outcomes (caregiver-child relationship), no statistically significant differences between groups were observed at any time point. However, one observable nonsignificant trend was apparent between the two groups on the subscale of parental distress within the parenting stress measure, but not on the total score, at the six-month follow-up assessment (see Table A.2). Caregivers in the intervention group reported less parenting distress than those in the control group. However, because of the multiple significance tests being conducted, this trend did not reach statistical significance and thus may be due to chance.

For tertiary outcomes (child and caregiver violence exposure), again no statistically significant differences between groups were observed at any time point (see Table A.3). However, an observable nonsignificant trend showed that at the 12-month follow-up assessment, caregivers in the intervention group reported that their children witnessed more violence than those in the control group. However, because of the multiple significance tests being conducted, this trend did not reach statistical significance and thus may be due to chance.

Mean Differences over Time

Table 7 shows differences over time for Kalamazoo's primary outcomes, comparing changes for each individual family between baseline and six months. In the second column of numbers in Table 7, the mean change between six-month scores and baseline scores (within family) is shown for each group. The comparison here is whether there was significant change on the outcomes for the families in each group separately (rather than a comparison of one group with the other). At six months, statistically significant changes in scores within groups were observed on several primary outcome variables, including reductions in PTSD symptoms (intervention only), increases in reports of child self-control (both groups) and cooperation (intervention only), decreases in scores on the passage comprehension test (both groups), and increases in scores on the letter-word identification and applied problems tests (intervention only).

The statistical test of differences in differences in unadjusted models (third column in Table 7) compares the amount of mean change for the intervention group families with the amount of mean change for the control groups between the baseline and six-month assessments. These comparisons revealed no statistically significant differences in change between the intervention and control groups, nor did the adjusted models (fourth column in Table 7).

For the 12-month assessment (data not shown), there were statistically significant within-group changes observed for improved child assertion (intervention only), increases in two measures of school readiness/performance (both groups), and a decrease in the third measure of school readiness/performance (both groups). However, no difference in differences or adjusted models showed differences that exceeded the threshold for significance using the FDR significance level. At the 18- and 24-month follow-up assessments, similar changes were observed over time within each group, but no difference in differences or adjusted models showed differences that exceeded the threshold for significance using the FDR significance level (data not shown).

Table 8 shows differences over time for Kalamazoo's secondary outcomes. In the second column of numbers in Table 8, the mean change between six-month scores and baseline scores is shown for each group. No statistically significant changes in scores were observed on secondary outcomes. However, observable nonsignificant trends were noted on two secondary outcomes, including reductions in the caregiver's report of parent distress (intervention only) and increases in the caregiver's report of parent-child dysfunction (control only), both subscales of the parenting stress measure. However, because of the multiple significance tests being conducted, this trend did not reach statistical significance and thus may be due to chance. The statistical test of differences in differences in unadjusted models (third column in Table 8) revealed no statistically significant differences between intervention and control groups, nor did the adjusted models (fourth column in Table 8). For the 12-month and 24-month follow-ups, no significant changes over were noted, and difference in differences and adjusted models did not show any differences between groups. At 18 months, two significant changes over time were observed within groups, with the control group showing improvements on two subscales of the parenting stress index. However, no difference in differences or adjusted models showed differences that exceeded the threshold for significance using the FDR significance level.

Table 9 shows differences over time for Kalamazoo's tertiary outcomes. In the second column of numbers in Table 9, the mean change between six-month scores and baseline scores is shown for each group. These show reductions in the reports of child

and caregiver victimization between the measure of lifetime exposure at baseline and the measure of victimization during the period between baseline and the six-month assessment, as would be expected, in both groups. The statistical test of differences in differences in unadjusted and adjusted models revealed no differences within groups over time or between intervention and control groups at any time point examined. For the 12-, 18-, and 24-month assessments, findings within groups over time were similar, but, again, no differences between groups in terms of change over time emerged.

Table 7
Changes in Means for Primary Outcome Variables Between Baseline and Six-Month Assessment and Group-Level Comparison of Mean Changes

Primary Outcome		N	Within-Family Mean Changes ^a	Group-Level Comparison of Mean Changes (Unadjusted Model) ^b	Group-Level Comparison of Mean Changes (Adjusted Model) ^c
PTSD Symptoms					
CR Child PTSD Symptoms for Ages 3–10	Intervention	184	-1.11 *	-0.37	-0.92
	Control	160	-0.74		
Behavior/Conduct Problems					
CR Child Externalizing Behavior Problems for Ages 3–18	Intervention	169	-0.14	-0.22	-1.01
	Control	147	0.07		
CR Child Internalizing Behavior Problems for Ages 3–18	Intervention	168	-0.11	-0.11	-0.45
	Control	148	0.00		
Social-Emotional Competence					
CR Child Assertion for Ages 1–12	Intervention	188	0.08	0.01	0.05
	Control	164	0.07		
CR Child Self-Control for Ages 1–12	Intervention	188	0.18 *	0.04	0.15
	Control	164	0.14 *		
CR Child Cooperation for Ages 3–12	Intervention	158	0.69 *	0.14	0.44
	Control	130	0.55		
School Readiness/Performance					
Letter Word Identification for Ages 3–18	Intervention	123	3.32 *	1.12	1.70
	Control	109	2.19		
Passage Comprehension for Ages 3–18	Intervention	118	-3.37 *	-0.64	-1.20
	Control	113	-2.73 *		

Table 7—continued

Primary Outcome		N	Within-Family Mean Changes ^a	Group-Level Comparison of Mean Changes (Unadjusted Model) ^b	Group-Level Comparison of Mean Changes (Adjusted Model) ^c
Applied Problems for Ages 3–18	Intervention	93	3.95 *	2.23	-0.50
	Control	89	1.72		

^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 This column reflects the estimate of the difference between the two groups' within-family mean change from baseline to six months, controlling for age, gender, ethnicity, income, and violence exposure at baseline. * indicates a significant test for the estimate.

NOTES: CR = Caregiver Report; SR = Child Self-Report. # indicates a nonsignificant trend in the t-test ($p < 0.05$ but does not meet the FDR correction threshold).

Table 8
Changes in Means for Secondary Outcome Variables Between Baseline and Six-Month Assessment and
Group-Level Comparison of Mean Changes

Secondary Outcome		N	Within-Family Mean Changes ^a	Group-Level Comparison of Mean Changes (Unadjusted Model) ^b	Group-Level Comparison of Mean Changes (Adjusted Model) ^c
Caregiver-Child Relationship					
CR Parental Distress for Ages 0–12	Intervention	185	–1.34 #	–0.81	–0.45
	Control	159	–0.53		
CR Parent-Child Dysfunction for Ages 0–12	Intervention	187	0.40	–0.63	–0.93
	Control	164	1.03 #		
CR Difficult Child for ages 0–12	Intervention	186	–0.05	–0.10	–1.14
	Control	164	0.04		
CR Total Parental Stress for Ages 0–12	Intervention	183	–1.00	–1.40	–2.46
	Control	159	0.40		

^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 This column reflects the estimate of the difference between the two groups' within-family mean change from baseline to six months, controlling for age, gender, ethnicity, income, and violence exposure at baseline. * indicates a significant test for the estimate.

NOTES: CR = Caregiver Report; SR = Child Self-Report. # indicates a nonsignificant trend in the t-test ($p < 0.05$ but does not meet the FDR correction threshold).

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

Tertiary Outcome		N	Within-Family Mean Changes ^a	Group-Level Comparison of Mean Changes (Unadjusted Model) ^b	Group-Level Comparison of Mean Changes (Adjusted Model) ^c
Violence Exposure					
CR Total Child Victimization Experiences for Ages 0–12	Intervention	186	-1.27 *	0.16	0.16
	Control	164	-1.43 *		
CR Child Maltreatment for Ages 0–12	Intervention	183	-0.24 *	0.01	0.00
	Control	161	-0.25 *		
CR Child Assault for Ages 0–12	Intervention	184	-0.22 *	0.14	0.17
	Control	161	-0.36 *		
CR Child Sexual Abuse for Ages 0–12	Intervention	185	-0.02	0.01	-0.01
	Control	162	-0.02		
CR Child Witnessing Violence for Ages 0–12	Intervention	177	-0.86 *	-0.11	-0.07
	Control	153	-0.75 *		
CR Caregiver Total Number of Traumatic Experiences	Intervention	186	-0.15 *	0.04	0.05
	Control	163	-0.18 *		
CR Caregiver Experience of Any Non-DV Traumas ^d	Intervention	188	-0.09 *	-0.03	-0.06
	Control	164	-0.06 *		
CR Caregiver Experience of Any Domestic Violence ^d	Intervention	188	-0.14 *	0.01	0.02
	Control	164	-0.15 *		

^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 This column reflects the estimate of the difference between the two groups' within-family mean change from baseline to six months, controlling for age, gender, ethnicity, income, and violence exposure at baseline. * indicates a significant test for the estimate.

^d The outcome is a categorical variable, and the unadjusted within-family mean change and the group-level comparison are changes in proportion, while the covariate-adjusted group-level comparison is the difference in proportions obtained from a linear probability model.

NOTES: CR = Caregiver Report; DV = domestic violence; SR = Child Self-Report. # indicates a nonsignificant trend in the t-test ($p < 0.05$ but does not meet the FDR correction threshold).

Safe Start Service Dosage and Changes in Primary Outcomes

To examine any effects of Safe Start service dosage on outcomes, we divided the Kalamazoo intervention families into three dosage groups, as families tend to receive different amounts of services. Using the data reported by the site, we created two dosage variables for Kalamazoo Safe Start. The first is a three-level variable capturing the number of classroom curriculum consultation meetings combined, with a low dosage being 35 contacts or fewer, a medium dosage being 36 to 60 contacts, and a high dosage being more than 60 contacts. About a third of the sample fell into each of these groupings. We also created a dichotomous variable for participation in the parent groups, with about a third of the sample participating and the remainder not. These variables were used in the subsequent outcomes analysis to examine the impact of the intervention according to the services received.

In a classroom-based intervention in which participation in the program is relatively equal across children, regardless of need, we did not expect differences between the different dosage groups, and, indeed, no large differences were noted. We expected that families that participated in the parent groups might differ from families that did not. For both analyses, we equated the individuals in each group as much as possible by using the propensity score-matching method to pair families in each dosage group with families with similar needs in the control group. The analyses examined the difference in mean change scores between the intervention and control groups for each dosage group. Note that in this analysis, the full control group is used in the matching of each of the dosage levels.

Table A.4 shows the results of the propensity score-matching analyses for Kalamazoo's primary outcomes. The statistical test comparing the two groups on changes in means scores between baseline and six months did not reveal any statistically significant differences between the intervention and control group for any of the dosage levels. Table A.5 shows the results comparing those who participated in at least one parent group with those who did not. Here also, the statistical test of differences in differences did not reveal any statistically significant differences between the intervention and control group for participants in the parent group.

CONCLUSIONS

This classroom-based intervention was fielded in Head Start classrooms and evaluated via a cluster-randomized controlled trial. The Kalamazoo Safe Start program team was able to recruit and enroll a large number of participants in the study and retained over 80 percent of them for the six-month assessment, demonstrating

feasibility in reaching this population and in recruiting them to participate in research. The participants in the study were functioning relatively well, with most of them scoring in the normal range on our child outcome measures, despite past exposure to violence. While this level of functioning at baseline is relatively good as compared with clinical samples, caregivers reported more child and caregiver problems than would be found in the general population. Among the families retained at six months, the majority of students in intervention classrooms were served with intervention elements as planned (95 percent), and about a third of parents participated in the optional parent groups. However, this participation was not considered by the program to represent a necessary component for intervention effectiveness (see Schultz et al., 2010).

Given the sample size in this study, there was an 80-percent chance to detect an effect of sizes of 0.30 to 0.41 at six months, depending on clustering, and to detect medium-to-large intervention effects at the subsequent follow-ups (0.41 to 0.68). We expected a small intervention effect, given that the intervention occurred at the classroom level and was compared with the regular Head Start curriculum. In the intent-to-treat analysis, despite mean scores in the intervention groups in the expected directions at six months and some significant improvements noted among those in the intervention group, no differences between groups over time were detected. Similarly, there were no statistically significant differences noted among those that received a low, medium, or high dose of the intervention and comparable control families, or between participants in the parent groups and comparable controls.

Overall, the results did not indicate that the Kalamazoo Safe Start program improved outcomes for children over time, relative to the group of similar children who did not receive the full Safe Start intervention. However, despite the rigorous randomized design, firm conclusions about the effectiveness of the Kalamazoo Safe Start program cannot be drawn based on these results. Although it is possible that the program as implemented in this study was not effective in improving outcomes for children in the intervention group relative to controls, there are several other possible explanations for the results presented here. The lack of difference between groups over time could reflect a lack of power to observe a small effect of the program on the measured outcomes. That is, a larger sample size might have allowed for the detection of statistically significant changes associated with the Safe Start services. Caregivers in this program reported relatively low levels of child symptoms at baseline, and this may have made it difficult to demonstrate significant changes in children over time, since there may not have been “room” to improve for children who were not experiencing many symptoms or problems at baseline. It is possible that interventions focused on

violence could heighten caregiver sensitivity to and recognition of child symptoms and thus any intervention effect may be obscured by caregiver reports of increased or level symptoms. The inability to detect significant differences between the groups may also have been due to the particular outcomes measured. That is, the Kalamazoo Safe Start program may have improved the lives of children and families in ways that were not measured (or were not measured adequately) in this study.

In sum, although the analyses of data did not clearly reveal impacts of the program at the child and family levels, the program demonstrates the promise of working with children in this setting, with high participation rates and successful implementation of the majority of program elements to children who participated. Although no statistical differences were detected between groups over time, the children in this study were functioning relatively well despite exposure to violence, and thus program impacts, if they exist, may take a preventative form and may not be apparent during this two-year study.

REFERENCES

- Benamati, J., *Systemic Training to Assist in the Recovery from Trauma (START) for Residential Staff (Curriculum)*, 2002.
- 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.
- Bloom, S. L., "Introduction to Special Section—Creating Sanctuary for Kids: Helping Children to Heal From Violence," *Therapeutic Community: The International Journal for Therapeutic and Supportive Organizations*, Vol. 26, No. 1, 2005, pp. 57–63.
- Chaffin, M., J. F. Silovsky, B. Funderburk, L. A. Valle, E. V. Brestan, T. Balachova,, et al., "Parent-Child Interaction Therapy with Physically Abusive Parents: Efficacy for Reducing Future Abuse Reports," *Journal of Consulting and Clinical Psychology*, Vol. 72, No. 3, 2004, pp. 500–510.
- Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc., 1988.
- Drake, B., and S. Pandey, "Understanding the Relationship Between Neighborhood Poverty and Specific Types of Child Maltreatment," *Child Abuse & Neglect*, Vol. 20, No. 11, 1996, pp. 1003–1018.
- Ford, J. D., and E. Russo, "Trauma-Focused, Present-Centered, Emotional Self-Regulation Approach to Integrated Treatment for Posttraumatic Stress and Addiction: Trauma Adaptive Recovery Group Education and Therapy (TARGET)," *American Journal of Psychotherapy*, Vol. 60, 2006, pp. 335–350.
- Hood, K. K., and S. M. Eyeberg, "Outcomes of Parent-Child Interaction Therapy: Mother's Reports of Maintenance Three to Six Years after Treatment," *Journal of Clinical Child and Adolescent Psychology*, Vol. 32, 2003, pp. 419–429.
- Macy, R. D., S. Bary, and G. G. Noam, "Youth Facing Threat and Terror Supporting Preparedness and Resilience," in *New Directions for Youth Development*, Jossey-Bass, 2003, pp. 51–79.
- 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
- Slack, K. S., J. L. Holl, M. McDaniel, J. Yoo, and K. Bolger, "Understanding the Risks of Child Neglect: An Exploration of Poverty and Parenting Characteristics," *Child Maltreatment*, Vol. 9, No. 4, November 2004, pp. 395–408.
- Webster-Stratton, C., "Parent Training with Low Income Families: Promoting Parental Engagement Through a Collaborative Approach," in J. R. Lutzker, ed., *Handbook of Child Abuse Research and Treatment: Issues in Clinical Child Psychology*, New York: Plenum Press, 1998, pp. 183–210.
- Webster-Stratton, C., "Advancing Videotape Parent Training: A Comparison Study," *Journal of Consulting and Clinical Psychology*, Vol. 62, No. 3, 1994, pp. 583–593.
- Western Michigan University, *Funding Proposal to the Office of Juvenile Justice and Delinquency Prevention—CFDA Title: Safe Start Promising Approaches for Children Exposed to Violence*, Kalamazoo, Mich., 2004.

KALAMAZOO OUTCOMES APPENDIX

Table A.1
Comparison of Means for Kalamazoo Primary Outcome Variables

		Baseline		Six Months		12 Months		18 Months		24 Months	
Primary Outcome		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
PTSD Symptoms											
CR Child PTSD Symptoms for Ages 3–10	Intervention	228	34.58	185	33.10	99	33.18	91	32.62	45	31.96
	Control	203	34.22	163	33.80	90	33.27	92	32.35	29	32.03
Behavior/Conduct Problems											
CR Child Externalizing Behavior Problems for Ages 3–18	Intervention	217	24.52	180	23.86	98	25.32	89	24.47	44	24.32
	Control	191	23.94	159	24.62	87	24.43	87	24.09	29	24.38
CR Child Internalizing Behavior Problems for Ages 3–18	Intervention	219	13.41	169	13.09	84	13.37	80	13.20	37	13.19
	Control	200	13.27	151	13.35	81	13.23	83	13.00	24	13.46
Social-Emotional Competence											
CR Child Assertion for Ages 1–12	Intervention	229	0.21	#	188	0.30	#	99	0.36	92	0.52
	Control	207	0.40		164	0.45		90	0.50	92	0.51
CR Child Self-Control for Ages 1–12	Intervention	229	0.16		188	0.38		99	0.30	92	0.44
	Control	207	0.26		164	0.35		90	0.41	92	0.43
CR Child Cooperation for Ages 3–12	Intervention	213	12.08	169	12.86	74	12.88	68	13.10	38	13.53
	Control	186	12.67	142	13.12	78	12.86	72	13.21	28	12.11

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	
School Readiness/Performance												
Letter Word Identification for Ages 3–18	Intervention	172	-4.74	161	-1.52	86	0.37	78	2.32	39	4.05	
	Control	158	-3.20	136	-1.46	82	-0.40	83	0.14	29	7.03	
Passage Comprehension for Ages 3–18	Intervention	171	11.96	162	9.28	87	5.07	77	2.48	#	42	-0.31
	Control	170	11.94	133	9.18	85	3.72	86	-0.31		30	0.87
Applied Problems for Ages 3–18	Intervention	135	-5.58	158	-3.13	83	-0.42	78	0.10		44	0.66
	Control	132	-3.54	133	-1.55	82	-0.87	86	1.02		28	4.39

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

Secondary Outcome		Baseline		Six Months			12 Months		18 Months		24 Months	
		N	Mean	N	Mean		N	Mean	N	Mean	N	Mean
Caregiver-Child Relationship												
CR Parent Distress for Ages 0–12	Intervention	228	26.79	185	25.30	#	99	26.23	91	25.24	43	26.49
	Control	205	26.76	160	26.89		89	26.34	91	24.85	27	26.30
CR Parent-Child Dysfunction for Ages 0–12	Intervention	229	20.61	187	20.58		99	20.22	92	20.34	45	21.60
	Control	207	20.06	164	21.31		90	20.22	92	20.57	30	19.83
CR Difficult Child for Ages 0–12	Intervention	228	29.13	187	28.53		99	28.45	92	27.50	46	28.57
	Control	207	28.35	164	28.83		90	27.69	92	27.08	30	28.27
CR Total Parenting Stress for Ages 0–12	Intervention	227	76.52	184	74.26		99	74.91	91	73.15	43	77.51
	Control	205	75.19	160	77.02		89	74.17	91	72.26	27	75.15

NOTES: CR = Caregiver Report. * indicates statistically significant (p-value<FDR significance criterion); # indicates nonsignificant trend (p<0.05 and >FDR significance criterion). 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.3
Comparison of Means for Kalamazoo Tertiary Outcome Variables

Tertiary Outcome		Baseline		Six Months		12 Months		18 Months		24 Months	
		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
Violence Exposure											
CR Total Child Victimization Experiences for Ages 0–12	Intervention	229	2.72	186	1.54	97	1.60	92	1.33	46	1.13
	Control	207	2.80	164	1.49	90	1.31	92	1.40	30	1.57
CR Child Maltreatment for Ages 0–12	Intervention	226	0.31	184	0.08	96	0.10	90	0.04	46	0.09
	Control	204	0.34	162	0.12	89	0.07	89	0.10	30	0.03
CR Child Assault for Ages 0–12	Intervention	228	0.80	185	0.64	98	0.56	92	0.61	46	0.46
	Control	207	0.97	161	0.65	90	0.60	91	0.63	30	0.73
CR Child Sexual Abuse for Ages 0–12	Intervention	228	0.03	186	0.01	95	0.04	92	0.02	46	0.00
	Control	207	0.06	162	0.04	89	0.00	92	0.01	30	0.00
CR Child Witnessing Violence for Ages 0–12	Intervention	222	1.38	182	0.55	96	0.66	88	0.39	42	0.33
	Control	195	1.16	158	0.47	86	0.41	88	0.36	28	0.46
CR Caregiver Total Number of Traumatic Experiences	Intervention	229	0.32	186	0.17	98	0.14	92	0.13	45	0.11
	Control	206	0.30	164	0.13	90	0.22	91	0.07	30	0.10
CR Caregiver Experience of Any Non-DV Trauma	Intervention	229	0.17	188	0.06	99	0.06	92	0.07	46	0.02
	Control	207	0.15	164	0.10	90	0.09	92	0.11	30	0.17

Table A.3—continued

Tertiary Outcome		Baseline		Six Months		12 Months		18 Months		24 Months	
		N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
CR Caregiver Experience of Any Domestic Violence	Intervention	229	0.20	188	0.06	99	0.03	92	0.04	46	0.00
	Control	207	0.20	164	0.05	90	0.08	92	0.07	30	0.07

NOTES: CR = Caregiver Report; DV = domestic violence. * indicates statistically significant (p-value < FDR significance criterion); # indicates nonsignificant trend (p < 0.05 and > FDR significance criterion). 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.4
Changes in Means by Dosage Group for Kalamazoo Primary Outcome Variables Between Baseline and Six-Month Assessment

Primary Outcome	Dosage	Group	N	Baseline Mean	Six-Month Mean	Difference
PTSD Symptoms						
CR Child PTSD Symptoms for Ages 3–10	Low	Intervention	60	34.93	33.20	-1.17
		Control	60	34.87	34.30	
	Medium	Intervention	60	32.12	31.98	-0.45
		Control	60	32.12	32.43	
	High	Intervention	51	34.33	33.35	-1.10
		Control	54	34.20	34.52	
Behavior/Conduct Problems						
CR Child Externalizing Behavior Problems for Ages 3–18	Low	Intervention	50	23.89	23.40	-0.30
		Control	56	24.04	24.00	
	Medium	Intervention	56	24.04	23.55	-0.78
		Control	57	24.09	24.37	
	High	Intervention	52	24.13	24.48	-0.12
		Control	53	23.49	24.08	
CR Child Internalizing Behavior Problems for Ages 3–18	Low	Intervention	50	13.41	12.94	-0.22
		Control	54	13.41	13.20	
	Medium	Intervention	51	12.70	12.69	-0.40
		Control	57	12.70	13.14	
	High	Intervention	56	13.39	13.36	0.30
		Control	57	13.42	13.11	
Social-Emotional Competence						
CR Child Assertion for Ages 1–12	Low	Intervention	57	0.21	0.30	-0.11
		Control	57	0.22	0.41	
	Medium	Intervention	61	0.36	0.38	-0.09
		Control	61	0.35	0.45	
	High	Intervention	46	0.25	0.36	-0.02
		Control	46	0.26	0.39	
CR Child Self-Control for Ages 1–12	Low	Intervention	59	0.15	0.37	0.07
		Control	59	0.12	0.27	
	Medium	Intervention	61	0.29	0.46	0.09
		Control	61	0.30	0.38	
	High	Intervention	57	0.17	0.34	0.14
		Control	57	0.14	0.16	
CR Child Cooperation for Ages 3–12	Low	Intervention	43	12.31	13.14	-0.28
		Control	51	12.31	13.31	
	Medium	Intervention	52	12.40	13.06	0.09
		Control	57	12.70	13.26	
	High	Intervention	50	12.00	12.60	-0.42
		Control	54	12.00	12.94	

Table A.4—continued

Primary Outcome	Dosage	Group	N	Baseline Mean	Six-Month Mean	Difference
School Readiness/Performance						
Letter Word Identification for Ages 3–18	Low	Intervention	40	5.63	-2.50	-5.55*
		Control	48	5.88	3.35	
	Medium	Intervention	47	2.80	0.23	0.62
		Control	51	3.67	0.67	
	High	Intervention	30	4.79	5.43	5.86
		Control	33	4.61	-0.45	
Passage Comprehension for Ages 3–18	Low	Intervention	39	-13.13	-6.56	3.36
		Control	47	-13.15	-10.40	
	Medium	Intervention	41	-11.79	-10.54	-1.11
		Control	47	-11.72	-8.62	
	High	Intervention	23	-12.96	-12.22	-3.09
		Control	25	-12.48	-8.52	
Applied Problems for Ages 3–18	Low	Intervention	31	3.05	-0.29	-2.90
		Control	38	3.34	3.21	
	Medium	Intervention	30	3.94	0.87	-1.56
		Control	33	4.03	1.76	
	High	Intervention	12	6.88	1.17	0.21
		Control	16	6.88	3.25	

NOTES: CR = Caregiver Report. * indicates statistically significant (p-value<0.05). 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.5
Changes in Means by Participation in Parent Groups for Kalamazoo Primary
Outcome Variables Between Baseline and Six-Month Assessment

Primary Outcome	Parent Group	Group	N	Baseline Mean	Six-Month Mean	Difference
PTSD Symptoms						
CR Child PTSD Symptoms for Ages 3-10	No	Intervention	111	33.91	32.41	-0.82
		Control	113	33.92	33.36	
	Yes	Intervention	61	33.87	33.62	0.25
		Control	62	34.10	33.55	
Behavior/Conduct Problems						
CR Child Externalizing Behavior Problems for Ages 3-18	No	Intervention	101	24.14	23.27	-0.95
		Control	109	24.20	24.39	
	Yes	Intervention	57	23.91	24.89	0.67
		Control	57	23.61	23.93	
CR Child Internalizing Behavior Problems for Ages 3-18	No	Intervention	101	13.28	12.94	-0.30
		Control	109	13.17	13.21	
	Yes	Intervention	57	13.03	13.19	-0.23
		Control	60	13.03	13.37	
Social-Emotional Competence						
CR Child Assertion for Ages 1-12	No	Intervention	101	0.29	0.38	0.05
		Control	101	0.30	0.33	
	Yes	Intervention	55	0.32	0.33	-0.06
		Control	55	0.33	0.39	
CR Child Self-Control for Ages 1-12	No	Intervention	114	0.16	0.43	0.09
		Control	114	0.15	0.34	
	Yes	Intervention	63	0.27	0.30	-0.06
		Control	63	0.30	0.39	
CR Child Cooperation for Ages 3-12	No	Intervention	81	12.44	13.05	0.10
		Control	88	12.44	12.88	
	Yes	Intervention	54	12.27	12.83	0.16
		Control	60	12.25	12.73	
School Readiness/Performance						
Letter Word Identification for Ages 3-18	No	Intervention	63	5.06	1.75	0.14
		Control	72	4.99	1.75	
	Yes	Intervention	42	2.36	-1.00	-2.81
		Control	47	3.06	2.49	
Passage Comprehension for Ages 3-18	No	Intervention	70	-11.58	-8.94	-0.70
		Control	83	-11.88	-8.41	
	Yes	Intervention	38	-12.11	-9.08	0.64
		Control	44	-11.75	-8.55	
Applied Problems for Ages 3-18	No	Intervention	39	4.71	0.41	-1.87
		Control	45	4.71	1.89	
	Yes	Intervention	22	2.43	-0.68	-0.08
		Control	28	2.43	0.14	

NOTES: CR = Caregiver Report. * indicates statistically significant (p-value<0.05). 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.