



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 L: San Diego County, California

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

SAN DIEGO COUNTY, CALIFORNIA, SAFE START OUTCOMES REPORT

ABSTRACT

The San Diego County Safe Start program involved a large community-level effort to increase awareness and capacity to serve the needs of children (ages 3–12) exposed to violence. The focus of the outcome evaluation was on the direct services portion of the program, which included an extensive assessment that was implemented in combination with a trauma-focused therapy, advocacy, and case coordination. A full description of the entire 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 intervention, with randomization occurring at the family level. San Diego’s Safe Start staff enrolled 104 families in the study, with 93 percent of them retained for the six-month assessment. At baseline, caregivers reported that children had been exposed to an average of 2.5 types of violence in their lifetimes and that 21 percent of children had posttraumatic stress disorder (PTSD) symptoms in the “significant” range. Overall, 40 percent of caregivers reported total parenting stress levels in the “clinical” range. San Diego’s primary services consisted of child advocacy and Trauma-Focused Cognitive Behavioral Therapy (TF-CBT) therapy sessions. Among the 51 intervention group families in the retained sample, 63 percent received at least one TF-CBT session, with an average of seven sessions per family. Nearly all in the retained sample received advocacy sessions (96 percent), ranging between one and 20 sessions during the six-month period. Within the intervention group, some significant improvements in scores were noted for the caregiver report of child PTSD symptoms and behavior problems. However, there were no significant differences between the intervention and control groups on any group-level test of change on the primary, secondary, or tertiary outcomes in six-month and 12-month analyses. Similarly, no differences in primary outcomes were detected among families receiving different dosages of services, relative to comparable control families. Overall, the sample size limitations mean that no conclusions can be drawn about the effectiveness of the San Diego Safe Start intervention as implemented

on child and family outcomes. Further testing is needed with an adequate sample in order to statistically assess the potential impacts of the intervention.

INTRODUCTION

The San Diego Safe Start program is located in San Diego County, California. In its 2004 application to the U.S. Department of Justice's Office of Juvenile Justice and Delinquency Prevention requesting Safe Start funding, San Diego County reported that each year there were more than 20,000 domestic violence-related 911 calls. Domestic violence fatality reviews conducted in the County between 1997 and 2003 found that 64 percent of the adult victims had at least one minor child, and in 23 percent of the cases a child witnessed the murder of his or her parent. Exposure to domestic violence was also a factor in many child welfare referrals. Since these cases were typically classified as "emotional abuse," however, the exact number of domestic violence referrals could not be determined. Overall, emotional abuse cases made up about 20 percent of the new child welfare dependency petitions before county juvenile courts (County of San Diego, 2004).

Despite a high level of existing coordination around domestic violence in San Diego, the county conducted a self-assessment and identified gaps in services available for children exposed to domestic violence. In particular, the county wanted to focus on families who had a child served within the county's Child Welfare Services agency and developed Safe Start to offer a family-level intervention within the context of Child Welfare Services. The intervention included TF-CBT (Cohen et. al., 2003), which was conducted in conjunction with an extensive child and caregiver assessment using the Trauma Assessment Pathway (TAP) model (Taylor et al., 2005). TF-CBT has been found to reduce symptoms associated with trauma in children in clinical trials (King et al., 2000; Cohen et al., 2004; Cohen, Mannarino, and Knudsen, 2005), but it has yet to be evaluated within a child welfare population or when paired with child advocacy and case coordination.

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

SAN DIEGO SAFE START

- **Intervention type:** TF-CBT, domestic violence–focused child advocacy, and case coordination
- **Intervention length:** Approximately 20 weekly sessions completed over six months for therapy services; six months or more for the advocacy and case coordination
- **Intervention setting:** Office-based for therapy services and case coordination; in-home or office-based for child advocacy
- **Target population:** Children exposed to domestic violence identified within a county child welfare population
- **Age range:** 3–12 years
- **Primary referral source:** Two regions of San Diego County’s Child Welfare Services

INTERVENTION

The San Diego Safe Start intervention focused on families with children involved with the county’s Child Welfare Services agency. The focus of the outcome evaluation was on the direct services portion of the program, which involved three main components: (1) TF-CBT with the TAP model, (2) child advocacy, and (3) case coordination. These elements are described briefly in the following paragraphs. For a full description of the San Diego intervention as it was delivered, see Schultz et al. (2010).

TF-CBT is a therapeutic intervention used for children, adolescents, and their parents or primary caregivers who have developed clinical levels of PTSD as the result of a traumatic event, such as child sexual or physical abuse, loss of a loved one, and exposure to domestic, school, or community violence (Cohen et al., 2006). TF-CBT involves therapy sessions in which children work to build emotional skills training and, later, with the help of trained therapists, children begin to confront the experience that initialized the PTSD symptoms (Cohen et al., 2003). As part of this process, the therapy includes

- education about trauma and common reactions
- help with parenting and behavior problems

- relaxation/stress management training
- learning about feelings and ways to express them
- developing creative ways for children to gradually discuss their traumatic experiences
- changing any unhelpful thoughts about the trauma
- family sessions to help the family talk together about the trauma
- learning and practicing safety skills.

The work with parents (or other primary caregivers) is generally focused on increasing their understanding of (1) the impacts of trauma on children's behavior and overall well-being and (2) appropriate strategies for supporting the process of addressing the impacts of trauma.

In the Safe Start program, TF-CBT was used in combination with an extensive assessment of the child and primary caregiver, called the TAP model (Taylor et al., 2005). The TAP model used a multifaceted assessment process to assist clinicians in gaining a more in-depth understanding of the child; his or her developmental level and traumatic experience; and the family, community and cultural system in which the child lives. The TAP assessment itself was expected to take up to three one-hour sessions to complete. Following that, the TF-CBT intervention, consisting of approximately 20 sessions, was delivered weekly at a therapist's office. If the goals of therapy had not been achieved after 20 sessions, children could continue with additional sessions.

Child advocacy services were provided in a client's home or other location, depending on the needs of the child and family. The advocate worked with staff in the child welfare offices as needed. The advocacy services were specifically focused on the child and family's domestic violence-related needs. They also provided emotional support and accompanied the family to court and appointments with agency service providers. The advocacy component varied in length, depending on the level of family, need but was expected to extend approximately six months.

The San Diego Safe Start program also provided regular case coordination meetings (called "clinical case reviews") among all three key service providers who were independently, yet simultaneously, working with the child and family: the child welfare case manager, the Safe Start therapist, and the Safe Start advocate. The purpose of the case coordination meetings was for the service

providers to discuss their individual perspectives of the child and family's status, needs, and progress. The meetings served as forums for the three key service providers to jointly discuss any case difficulties, resolve any disagreements in service needs or provision, and jointly develop plans to help families achieve success in meeting goals in each of the three service domains: child welfare, advocacy, and therapy. The project held these case coordination meetings monthly.

Efforts to monitor the therapist's implementation of the Safe Start intervention included a prerequisite TF-CBT online training course, three four-hour training sessions on TAP and TF-CBT, and participation in at least eight additional Safe Start training and clinical support sessions. Clinical consultants were also available to therapists on an as-needed basis by phone or email. In addition, the advocates were trained domestic violence advocates. Oversight of the child advocates was provided by their agency supervisor in regular meetings and in case-specific consultation on an as-needed basis. Monthly case management and review meetings including Child Welfare and Advocacy line staff were implemented early in the project to add an additional layer of quality assurance for all Safe Start cases.

METHOD

Design Overview

The design of this study is a randomized controlled trial, with randomization occurring at the family level. In addition to the usual services provided by Child Welfare Services, children in the treatment group received TF-CBT, child advocacy, and case coordination. Children assigned to the control group received the usual services of San Diego County's Child Welfare Services agency. These services varied by the needs of the family and may have included case management, financial assistance, and individual or family therapy services. No control group children, however, received TF-CBT from a Safe Start therapist. Child outcomes and contextual information were assessed at baseline, six, 12, 18, and 24 months. Study enrollment took place between January 2007 and March 2009, follow-up assessments continued until October 2009, and collection of 12-month assessments continued through May 2010. A previous version of this report covering the period of July 2006 through October 2009 was completed

earlier. This report updates the earlier report by adding data collected during the seven-month extension.

Evaluation Eligibility Criteria

When a Child Welfare Services supervisor in one of the participating offices identified domestic violence as one of the identified risk factors in a case, the family was flagged for Safe Start recruitment. A point person in each office would use the following criteria to determine program eligibility:

- At least one child in the family was between the ages of 3 and 12.
- Domestic violence exposure had occurred within six months of the family's referral to Child Welfare Services, as documented in the Child Welfare Services case plan.
- Child Welfare family maintenance cases were opened voluntarily or court-ordered.
- The family was able to effectively communicate in either English or Spanish.
- A nonoffending caregiver was available to participate in the child's treatment.
- The child and parent/primary caregiver were free of significant cognitive impairment or severe mental illness.

If more than one child was within the program's eligible age range, the child with the most recent birthday was selected as the target child.

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 to equalize the ages of children within the two groups. The first age group consisted of children between the ages of 3 and 6, and the second contained children between the ages of 7 and 12. In the case of three families who

had been enrolled and participated in initial assessments, it was later discovered that they were ineligible for the study. Thus, these three families were removed from all data and analyses.

Families that were randomized to the control group received the usual services provided by Child Welfare Services. The level and content of these services varied depending on the Child Welfare Services' standard assessment of family and child need. The services may include case management, financial support and assistance, and individual and group therapy sessions. If therapy was deemed needed for a given family, it would consist of sessions of any one of the various models delivered by therapists under contract with San Diego County. TF-CBT therapy and child advocacy services were not available to control families (see Schultz et al., 2010).

Families randomized to the intervention group received the usual services of Child Welfare Services, but all were referred to TF-CBT sessions with a specially trained Safe Start therapist. Intervention group families also received child advocacy services and were the focus of case coordination meetings among service providers.

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 San Diego Safe Start intervention, the outcomes were prioritized as shown in Table 1.

Table 1
Prioritized Outcome Measures for San Diego 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	3–10 years	Caregiver
PTSD Symptoms	Trauma Symptom Checklist for Children	8–12 years	Child
Behavior/Conduct Problems	Behavior Problem Index	All	Caregiver
Violence Exposure	Juvenile Victimization Questionnaire	All	Caregiver
Violence Exposure	Juvenile Victimization Questionnaire	11–12 years	Child
Violence Exposure	Caregiver Victimization Questionnaire	All	Caregiver
Secondary Outcome Measures			
<i>Domain</i>	<i>Measure</i>	<i>Age of Child</i>	<i>Respondent</i>
Depressive Symptoms	Children’s Depression Inventory	8–12 years	Child
Behavior/Conduct Problems	Delinquency Items	11–12 years	Child
Social-Emotional Competence	BERS-2 (School Functioning, Affective Strengths)	6–12 years	Caregiver
Social-Emotional Competence	BERS-2 (School Functioning, Affective Strengths)	11–12 years	Child
Social-Emotional Competence	BITSEA and SSRS (Assertion and Self-Control)	All	Caregiver
Social-Emotional Competence	SSRS (Cooperation)	All	Caregiver
Caregiver-Child Relationship	BERS-2 (Family Involvement)	6–12 years	Caregiver
Caregiver-Child Relationship	BERS-2 (Family Involvement)	11–12 years	Child
Caregiver-Child Relationship	Parenting Stress Index	All	Caregiver
School Readiness/Performance	Woodcock-Johnson III	All	Child
Tertiary Outcome Measures			
<i>Domain</i>	<i>Measure</i>	<i>Age of Child</i>	<i>Respondent</i>
Background and Contextual Factors	Everyday Stressors Index	All	Caregiver

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

Enrollment and Retention

Recruitment in the San Diego Safe Start project was a two-phase process. First, eligible families needed to give consent to Child Welfare workers to be contacted by researchers to hear about the study. Second, families needed to give consent to the research team to participate in the study. Enrollment did not occur until after families had consented to participate and had completed the baseline

research assessment. After the assessment, the random assignment procedures were implemented by the research director, and the Child Welfare agency and family were informed about the results.

Data about referrals and recruitment were taken from the Quarterly Activity Reports submitted by the San Diego site every quarter through the fall of 2009. The site provided referral information at the child welfare worker recruitment stage (Phase 1) for just the first six quarters of the full ten-quarter recruiting period. Referrals and recruiting data for the research team (Phase 2) were available over all ten quarters. Thus, we know about recruitment among all families only for only about 60 percent of the full study fielding period.

For the first six quarters of the study, 18 percent of the families (68 of 385) who were approached by child welfare workers about study participation agreed to be contacted by the research team. Of those families who declined to be contacted by the research team, a lack of time was the primary reason offered (43 percent). Of those who agreed to be contacted about the study, 75 percent (51 of 68) of the families were ultimately enrolled in the evaluation. In short, 13 percent of the 385 families were recruited in Phase 1 of the site's recruiting process.

Phase 2 of the recruiting process, conducted by the research team, resulted in enrollment of 82 percent of the 127 families referred from Phase 1. As noted previously, these data cover the entire ten-quarter recruiting period of the study. The most common reasons for families declining in Phase 2 were caregiver concern about a lack of time or lack of caregiver interest in participation.

In Table 2, we present the number and percentage of all enrollees who were eligible for participation at each data collection time point. San Diego's program staff enrolled 94 families in the study and completed a six-month research assessment for 93 percent of caregivers and 91 percent of children. At 12 months, the retention rate was 85 percent for caregivers and 84 percent for children. For subsequent research assessments, the site retained from 91 to 100 percent of the families; however, this represented a very small number of actual assessments (see Table 2).

Table 2
Retention of 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	51	49	9	1	51	49	9	1
Expected*	54	54	10	1	54	54	10	1
Retention Rate	94%	91%	90%	100%	94%	91%	90%	100%
Control								
Received	46	39	11	1	44	38	11	1
Expected*	50	50	11	1	50	50	11	1
Retention Rate	92%	78%	100%	100%	88%	76%	100%	100%
Overall								
Retention Rate	93%	85%	95%	100%	91%	84%	95%	100%

* 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 before all families entered the window of time for assessments at 18 or 24 months.

Special Issues

The pace of recruitment was slower than expected over the course of the study. The program made several adjustments over the course of the evaluation. Among these, the maximum age range was expanded during the first year of the project from 8 years to 12 years, and additional training was provided to child welfare workers on the referral process and identification of eligible cases. These changes did serve to increase the pace of enrollment. Once families were enrolled, the site was able to maintain a very high rate of retention in the study and participation in follow-up assessments. 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, family income level, and the child's violence exposure. We also compared the two groups on primary, secondary, and tertiary outcomes at baseline. Because this was a randomized experimental design, we did not expect any meaningful differences between the two groups at

baseline. However, to be certain, we tested for differences in child and caregiver characteristics and outcomes between intervention and control group children 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).

The TF-CBT intervention used in the San Diego program has shown moderate to large intervention effects when implemented in clinical trials, but there is some uncertainty as to whether the same intervention effects should be expected in a community setting or child welfare population. Therefore, it seems more reasonable to expect a small to medium effect. Of the 104 children enrolled at baseline (54 intervention and 50 control), 97 of them were observed at both baseline and six months (51 intervention and 46 control). With this sample size, at the nominal 0.05 significance level we can expect only a 16-percent chance to detect a small effect and a 68-percent chance to detect a medium effect, but we will have an 80-percent chance (the usual nominal) to detect a slightly larger-than-medium effect of size 0.576, according to Cohen's 1988 effect size classification.

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 the 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), comparison 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 who drop out prior to completion may be systematically different than those who remain. We also conducted analyses that take into account the type and amount of services received to account for dosage variability.

For the intent-to-treat analyses, we examined differences between the intervention and control groups, presenting baseline and follow-up estimates of primary, secondary, and tertiary outcomes for both groups when the sample size is greater than or equal to five. We compare means within groups across time using t-tests, compare groups via chi-square or t-tests at each time point, and examine difference in differences to compare the two groups on mean changes over time between baseline and follow-up assessments (when the sample size is at least ten per group). Where the sample size is over 20 in each group, we conducted multiple linear regressions on the different continuous outcomes and linear probability regressions on the dichotomous outcomes to test for the difference in difference via main effects and the interaction between intervention status and time after controlling for baseline characteristics (parent and child age, gender, race, and exposure to violence). These baseline characteristics were selected to correct for any potential imbalance in the groups by relevant demographic characteristics.

To assess outcomes related to the as-treated sample (i.e., 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. Since children with more need are likely to receive more services, we would expect this selection scheme to possibly present an unforeseeable bias, with families more in need receiving more services. To account for this selection bias related to service dosage, we used the propensity score-matching method to pair families in each dosage group with families with similar needs in the 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 the six-month 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 ($p < 0.05$), 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 (as reported in Tables 6–8) using an FDR of 0.05. With 14 model test statistics conducted among the primary outcomes, this led to adopting a statistical significance cutoff of 0.004 in the covariate-adjusted difference in difference results. With only ten secondary outcomes with enough sample sizes to allow for modeling, the FDR significance level adopted was 0.005, while such significance level was set at 0.025 for the tertiary outcomes (only two outcomes and with adequate data to examine only unadjusted difference in difference 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's adjustment 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 3, the baseline sample was composed of 57 percent females. The majority of the children were age 6 and younger (58 percent), with 32 percent age 7, 8, or 9. The children in the sample were predominately Hispanic (57 percent), with the remainder mostly other race/ethnicity children (27 percent). The majority (84 percent) of children had

family incomes under \$30,000 per year. According to the caregiver reports, children in the baseline sample had been exposed to an average of 2.5 types of violence in their lives prior to the baseline assessment. All caregivers completing the baseline assessments were a birth or adoptive parent of the target child. There were no statistically significant differences between the intervention and control groups on the distribution of these background characteristics. Because of the very high retention rate at the six-month follow-up, the demographics remained essentially unchanged in the sample available for six-month outcome analyses (data not shown).

Table 3
San Diego Safe Start Families Sample Characteristics in the Baseline Assessment Sample

	Combined		Intervention		Control		Test for Comparison P-Value
<i>Child Characteristics</i>	<i>N</i>	<i>Mean</i>	<i>N</i>	<i>Mean</i>	<i>N</i>	<i>Mean</i>	
Age	104	6.5	54	6.7	50	6.4	0.62
CR Violence Exposure	102	2.4	63	2.5	49	2.4	0.83
SR Violence Exposure	7	2.7	4	4.3	3	0.7	0.07
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	
<i>Gender</i>							
Male	45	43	26	48	19	38	0.30
Female	59	57	28	52	31	62	
<i>Race/Ethnicity</i>							
Hispanic	59	57	32	59	27	54	0.94
White	8	8	4	7	4	8	
Black	9	9	4	7	5	10	
Other	28	27	14	26	14	28	
<i>Caregiver Characteristics</i>							
<i>Family Income Level</i>							
Less than \$5,000	20	20	9	18	11	23	0.53
\$5,000–\$10,000	18	18	8	16	10	21	
\$10,001–\$15,000	20	20	10	20	10	21	
\$15,001–\$20,000	13	13	10	20	3	6	
\$20,001–\$30,000	12	12	6	12	6	13	
More than \$30,000	16	16	8	16	8	17	
<i>Relationship to Child</i>							
Parent or Guardian	104	100	54	100	50	100	
Other Relationship	0	0		0		0	

NOTES: CR = Caregiver Report; SR = Child Self-Report. Percentages may not total 100 percent because of rounding.

Next, we examined the San Diego sample at baseline on two outcomes (PTSD symptoms and parenting stress) to understand the level of severity on these indexes among families entering the study. As shown in Table 4, at baseline, caregivers reported PTSD symptoms that fell in the significant range for 21 percent of the sample overall, 11 percent of boys and 28 percent of girls. For the caregiver-child relationship, 41 percent of the sample had total stress levels that fell in the clinical range, with 36 percent for boys and 44 percent for girls.

Table 4
Baseline Assessment Estimates for San Diego Safe Start Families

	Combined		Boys		Girls	
	N	%	N	%	N	%
CR PTSD Symptoms for Ages 3–10						
Normal	58	67%	30	81%	28	56%
Borderline	11	13%	3	8%	8	16%
Significant	18	21%	4	11%	14	28%
CR Total Parenting Stress for Ages 0–12						
Parental Distress–Clinical	31	30%	11	25%	20	34%
Parent-Child Dysfunctional Interaction–Clinical	35	34%	14	32%	21	36%
Difficult Child–Clinical	38	37%	17	39%	21	36%
Total Stress–Clinical	42	41%	16	36%	26	44%

NOTE: CR = Caregiver Report.

Finally, we examined differences between the intervention and control groups at baseline for San Diego’s primary, secondary, and tertiary outcomes (see this report’s appendix). Primary outcomes included PTSD symptoms, behavior problems, and child violence exposure. No statistically significant differences were observed between the groups overall on these measures (see Table A.1, first column). No statistically significant differences or trends were observed between the intervention and control groups in relation to secondary or tertiary outcome variables at baseline (see Tables A.2 and A.3, first column).

Uptake, Dosage, and Process of Care

As described fully in the process evaluation report (Schultz et al., 2010), the San Diego Safe Start intervention focused on families with children involved with the county Child Welfare Services agency. The direct services portion of the program involved three main components: (1) TF-CBT, (2) child advocacy, and (3) case coordination meetings between the service providers. Tables 5a and 5b show the type and amount of services received by the families assigned to the intervention group.

Table 5a sums total services, with a maximum of 24 months of service provision. Sixty-one percent of the 54 families enrolled at baseline received an average of 12 TF-CBT sessions, ranging between one and 29 sessions. The program reported the reason that TF-CBT sessions ended for 24 of the 33 participating families. Of those, 12 (50 percent) reportedly completed the TF-CBT

sessions successfully, and 12 (50 percent) families did not successfully complete services for a variety of reasons, including loss of program contact, family’s elective end of services, or other discontinuation. Nearly all (96 percent) of the intervention group families received at least one advocacy session. On average, families received 11 advocacy sessions, with 31 percent receiving five or fewer sessions and 29 percent receiving 16 or more. The San Diego program reported information on the reason that the advocacy services ended for 77 percent of all enrolled families who received the services (40 of 52). Among those, 45 percent successfully completed services, and the remainder disengaged from services on their own (45 percent), or services were terminated by the program (10 percent). At least one case consultation meeting was held for 50 percent of the intervention group families. For 44 percent of these families, there was a single meeting, but two meetings were held in 26 percent of the cases. The program reported that service ending information was not meaningful for case coordination meetings, as these were not dictated by family need but instead were more tied to overall program milestones.

Table 5a
Services Received by San Diego Safe Start Intervention Group Families
(Baseline Sample)

Service	Number with Service	Percentage with Service*	Range	Distribution	Mean	Median
TF-CBT Sessions	33	61%	1–29	1–5 36% 6–10 18% 11–15 15% 16+ 30%	12	9
Advocacy	52	96%	1–38	1–5 31% 6–10 29% 11–15 12% 16+ 29%	11	9
Case Coordination Meetings	27	50%	1–9	1 44% 2 26% 3–4 15% 5+ 15%	3	1

* The denominator is the 54 intervention group families who were initially enrolled in the intervention group for whom one or more follow-up Family Status Sheets were submitted.
 NOTE: Percentages may not total 100 percent because of rounding.

Table 5b shows the services received between the baseline and six-month assessment for only the 51 intervention families who participated in at least the

six-month follow-up research assessment. In other words, Table 5b shows the first six months of services for families in the study's six-month retained sample. TF-CBT sessions were delivered to 63 percent of the retained sample families. They participated in an average of seven sessions each, with half of the families receiving between one and five sessions. Nearly all in the retained sample received advocacy sessions, ranging between one and 20 sessions during the six-month period. Case coordination meetings occurred on behalf of 49 percent of intervention group families in the six-month retained sample. Of those, one meeting was held for 56 percent of the families, but 28 percent of the families were the focus of three or four meetings during the six-month window.

During the first six months of participation, the San Diego program reported the reason advocacy services ended for nine (18 percent) families in the six-month retained sample who received this service. Of those, services ended after being satisfactorily completed for 44 percent of families, but families discontinued services on their own in 44 percent of cases or were ended by the program in the remaining 11 percent of cases. In the first six months of participation, the program reported that the TF-CBT service ended for 32 percent of the families receiving this service. Of those, 10 percent (one of ten) reportedly ended TF-CBT sessions after successful completion. The remaining 90 percent ended sessions before successful completion for reasons including loss of contact with the family and family choice to discontinue.

Table 5b
Six-Month Service Received by San Diego Safe Start Intervention Families in the Six-Month Analysis Sample

Service	Number with Service	Percentage with Service*	Range	Distribution	Mean	Median
Advocacy	49	96%	1–20	1–5 33% 6–10 39% 11–15 14% 16+ 14%	9	8
TF-CBT Sessions	32	63%	1–21	1–5 50% 6–10 19% 11–15 22% 16+ 9%	7	5
Case Coordination Meetings	25	49%	1–4	1 56% 2 16% 3 16% 4 12%	2	1

* The denominator is the 51 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.

Outcome 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 and 12 months). Results for the 18- and 24-month assessment points are not discussed, as the sample size fell below the minimum threshold of at least ten cases per group. We first tested whether there were statistically significant changes in mean scores within the intervention group and the control group (see the appendix). Then, at each assessment point, we compared the mean score change of the two groups to determine if there were statistically significant differences in mean changes, using an intent-to-treat approach that included all families in the intervention group, regardless of the level of service they received (see Tables 6, 7, and 8). In these sections, data are displayed only for individual measures where there was an adequate number of families for statistical testing (i.e., ten or more families per group). Finally, we present descriptive data on families that received services, as compared to similar controls, on primary outcomes only.

Comparison of Means Between Groups

A summary of differences between the intervention and control group at each follow-up assessment point for San Diego's primary, secondary, and tertiary outcomes are depicted in the appendix. Table A.1 shows the results for the primary outcomes, which included PTSD symptoms, behavior problems, and child violence exposure. There were no statistically significant differences between the overall group means at the six-month or 12-month assessment point. Adequate data were not available for comparing group means at other time points.

Table A.2 shows San Diego's secondary outcomes, consisting of depressive symptoms, social-emotional competence, caregiver-child relationship, and school readiness/performance. Like the primary outcomes, there were no statistically significant differences between the groups at any follow-up assessment point (where enough data were available for testing).

San Diego's tertiary outcomes consisted only of the background and contextual factor domain. The measures were caregiver resource problems and caregiver personal problems. As shown in Table A.3, there were no statistically significant differences at any follow-up assessment point between the intervention and control group on these measures.

Mean Differences over Time

Table 6 shows differences over time for San Diego'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. 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). For the intervention group, the analyses revealed several statistically significant within-group differences from baseline to six months, including decreases in the caregiver's report of PTSD symptoms and both measures of child behavior problems. There was also a significant decline within the control group on child self-reported PTSD symptoms. There were a number of significant declines on the caregiver-reported violence exposure measures within both the treatment group and the control group. Specifically, there were significant declines for both groups on total types of child victimization experiences, the child witnessing violence, and caregiver reports of experience of domestic violence. There was also a significant decline within the intervention group on child maltreatment

reports and within the control group on child assault. Overall, these significant within-group changes on the violence exposure measures are not surprising because the time periods compared in the assessments are not equivalent (i.e., baseline measures ask about lifetime or 12-month experiences relative to the six-month period covered in the follow-up assessment). Several nonsignificant trends emerged in the violence exposure domain, with intervention group caregivers reporting a nonsignificant downward trend in child assault, and control group caregivers reporting a nonsignificant downward trend in child maltreatment and non-domestic violence related traumas. However, because of the multiple significance tests being conducted, these trends did not reach statistical significance and thus may be due to chance.

For the 12-month follow-up assessment (data not shown), the within-group mean changes were consistent with those at six months, with statistically significant change found for both groups on total types of child victimization experiences, the child witnessing violence, and caregiver experience of domestic violence. At 12 months, there were also statistically significant decreases within the intervention group in externalizing behavior, internalizing behavior, child assault experiences, and child maltreatment experiences. Within the control group, there was also a nonsignificant downward trend in child assault experiences at 12 months.

For the 18-month follow-up assessment (data not shown), only the control group had an adequate sample size to conduct significance tests of within-group mean changes. Within the control group, there were statistically significant decreases in the caregiver's report of child PTSD symptoms, total types of child victimization experiences, child witnessing violence experiences, and child assault experiences.

The third column in Table 6 shows 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 analyses revealed no statistically significant differences in the amount of change between the two groups on any of the primary measures. The final column of Table 6 reports the results for the primary measures in models that adjust for age, gender, ethnicity, income, and violence exposure at baseline. No statistically significant differences between the two groups emerged in these analyses. For the 12-month follow-up assessment, the group comparison of mean

change over time revealed no statistically significant differences for any measure (data not shown). Likewise, there were no significant findings in statistical tests of adjusted models at 12 months for any measures (primary, secondary, or tertiary).

Table 7 shows differences over time for San Diego's secondary outcomes. Between baseline and the six-month follow-up, there was one statistically significant within-group difference. Specifically, on one caregiver-child relationship measure, the intervention group caregivers reported a significant six-month decline in parental distress. There was an observable non-significant decline on this measure within the control group as well. There were also non-significant downward trend within both the control group and the intervention group on total parental stress. At the 12-month assessment point, there were some statistically significant within-group changes. For both the intervention group and the control group, there were significant downward trends in reports of parental distress, difficult child, and total parental stress at 12 months, relative to baseline. For the 18-month follow-up assessment (data not shown), only the control group had an adequate sample size to conduct significance tests of within-group mean changes. Within the control group, there were observable nonsignificant decreases in reports of parental distress and total stress and increases in reports of child emotional control.

Table 8 presents the results for the San Diego program's tertiary measures. At the six-month assessment point, control group caregivers reported a significant decline in personal problems. There was also an observable nonsignificant downward trend in personal problems for intervention group caregivers. At the 12-month assessment point, caregivers in both the intervention and control groups reported a significant decline in personal problems, relative to baseline. Caregivers in the intervention group also reported significantly fewer resources problems at 12 months. For the 18-month follow-up assessment (data not shown), only the control group had an adequate sample size to conduct significance tests of within-group mean changes. Caregivers in the control group continued to report significantly fewer personal problems at 18 months. There were no other statistically significant results or trends in the data on within-family changes.

For secondary and tertiary outcomes, the comparison of change in the group-level mean scores from baseline to six months revealed no statistically

significant differences between groups in the unadjusted or adjusted models (third column of Tables 7 and 8). Similarly, no statistically significant differences in mean change scores were observed in unadjusted or adjusted comparisons by the 12-month assessment point.

Table 6
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	47	-3.74 *	-2.40	-2.54
	Control	41	-1.34		
SR Child PTSD Symptoms for Ages 8–12	Intervention	18	-1.50	0.75	0.55
	Control	12	-2.25 *		
Behavior/Conduct Problems					
CR Child Externalizing Behavior Problems for Ages 3–18	Intervention	42	-2.45 *	-0.94	-0.39
	Control	39	-1.51		
CR Child Internalizing Behavior Problems for Ages 3–18	Intervention	50	-1.36 *	-0.84	-0.80
	Control	42	-0.52		
Violence Exposure					
CR Total Child Victimization Experiences for Ages 0–12	Intervention	51	-1.55 *	-0.10	0.00
	Control	45	-1.44 *		
CR Child Maltreatment for Ages 0–12	Intervention	51	-0.35 *	-0.14	-0.08
	Control	46	-0.22 #		
CR Child Assault for Ages 0–12	Intervention	51	-0.27 #	0.05	0.13
	Control	46	-0.33 *		
CR Child Sexual Abuse for Ages 0–12	Intervention	51	0.04	0.06	0.06
	Control	45	-0.02		
CR Child Witnessing Violence for Ages 0–12	Intervention	46	-1.04 *	-0.04	-0.07
	Control	37	-1.00 *		
CR Caregiver Total Number of Traumatic Experiences	Intervention	51	-0.06	-0.08	-0.06
	Control	46	0.02		
CR Caregiver Experience of Any Non-DV Traumas ^d	Intervention	51	0.00	0.09	0.11
	Control	46	-0.09 #		

Table 6—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
CR Caregiver Experience of Any Domestic Violence ^d	Intervention	51	-0.71 *	0.01	-0.02
	Control	46	-0.72 *		

^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.

^d This 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). Data are not shown for outcomes when the cell size is fewer than five for the group. P-values for the t-test results are not shown when the group size is fewer than ten for either group. Adjusted model results are not shown when the group size is fewer than 20 for either group.

Table 7

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
Depressive Symptoms					
SR Child Depressive Symptoms for Ages 8–18	Intervention	17	-0.82	-1.13	-2.81
	Control	13	0.31		
Social-Emotional Competence					
CR Child Affective Strengths for Ages 6–12	Intervention	32	0.44	1.10	0.35
	Control	27	-0.67		
CR Child School Functioning for Ages 6–12	Intervention	32	0.00	0.38	1.04
	Control	26	-0.38		
CR Child Assertion for Ages 1–12	Intervention	49	-0.16	-0.11	-0.15
	Control	46	-0.05		
CR Child Self-Control for Ages 1–12	Intervention	51	0.01	-0.15	-0.30
	Control	46	0.16		
CR Child Cooperation for Ages 3–12	Intervention	45	-0.02	-0.34	-0.33
	Control	38	0.32		
Caregiver-Child Relationship					
CR Parental Distress for Ages 0–12	Intervention	51	-4.22 *	-1.39	-0.04
	Control	46	-2.83 #		
CR Parent-Child Dysfunction for Ages 0–12	Intervention	51	-0.02	1.76	2.85
	Control	46	-1.78		
CR Difficult Child for Ages 0–12	Intervention	51	-1.67	0.18	2.15
	Control	46	-1.85		
CR Total Parental Stress for Ages 0–12	Intervention	51	-5.90 #	0.55	4.96
	Control	46	-6.46 #		

Table 7—continued

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
CR Family Involvement for Ages 6–12	Intervention	32	0.75	0.86	0.07
	Control	27	-0.11		
School Readiness/Performance					
Letter Word Identification for Ages 3–18	Intervention	41	-2.12	-1.33	0.85
	Control	39	-0.79		
Passage Comprehension for Ages 3–18	Intervention	42	-3.60	-1.60	-0.80
	Control	42	-2.00		
Applied Problems for Ages 3–18	Intervention	45	0.31	-1.14	-1.15
	Control	40	1.45		

^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). Data are not shown for outcomes when the cell size is fewer than five for the group. P-values for the t-test results are not shown when the group size is fewer than ten for either group. Adjusted model results are not shown when the group size is fewer than 20 for either group.

Table 8
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
Background and Contextual Factors					
CR Caregiver Resource Problems	Intervention	51	-0.76	0.37	0.80
	Control	46	-1.13		
CR Caregiver Personal Problems	Intervention	51	-1.59	0.67	1.47
	Control	46	-2.26 *		

^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). Data are not shown for outcomes when the cell size is fewer than five for the group. T-tests were not conducted when the group size was fewer than ten for either group. Adjusted model results are not shown when the group size is fewer than 20 for either group.

Safe Start Service Dosage and Changes in Primary Outcomes

To examine any intervention effects of the dosage of Safe Start services on the intervention's primary outcomes, we divided the intervention families into three equal groups depending on the number of total Safe Start service sessions received in the first six months. We calculated this total by summing for each family the number of advocacy sessions, therapy sessions, and case coordination meetings that took place. The Safe Start intervention represented a package of services, rather than a single service of a specific type. Thus, we defined the overall Safe Start dose as a sum of all services together. Because the sample size already substantially reduced statistical power in these analyses, we were unable to further explore whether there may have been different impacts for the different services offered by the site.

Approximately one-third (32 percent) of the families received between one and nine total service sessions in the six-month period and were thus defined as the low-dosage group. Another third (34 percent) of families received between ten and 16 total service sessions and were designated the medium-dosage group. The remaining 34 percent of families were designated the high-dosage group, receiving between 18 and 33 sessions.

Since children and families with more need are likely to receive more services, we would expect this selection bias, with needier families 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 baseline scores on the outcome measure of interest. The analyses examined the difference in mean score changes 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.

The results of the propensity score-matching analyses for San Diego's primary outcomes revealed no statistically significant differences between the intervention and control groups for any of the dosage levels with an adequate group size. These findings show that there appeared to be little difference overall between the dosage groups on baseline symptoms or problems. On statistical tests of the change in scores between baseline and six months for each family, no statistically significant differences by dosage group were detected (see Table A.4). We also analyzed changes in scores between baseline and 12 months for

each family and found no statistically significant differences by dosage group (see Table A.5).

CONCLUSIONS

San Diego's Safe Start Program provided child advocacy and TF-CBT, with enhanced case coordination between service providers, to families in a child welfare population. The evaluation consisted of a randomized controlled trial of the intervention. San Diego's Safe Start staff enrolled 104 families in the study, with 93 percent of them retained for the six-month assessment. At baseline, caregivers reported that children had been exposed to an average of 2.5 types of violence in their lifetimes and that 21 percent of children had PTSD symptoms in the significant range. Overall, 41 percent of caregivers reported total parenting stress levels in the clinical range. Nearly all San Diego Safe Start enrolled intervention group families (97 percent) received at least one session with a child advocate, with an average of ten sessions per family. At least one TF-CBT session occurred for 61 percent of the intervention group families who were initially enrolled, averaging ten per family over the course of the program.

Among the 51 intervention group families in the retained sample, 63 percent received at least one TF-CBT session, with an average of seven sessions per family. Nearly all in the analytic sample received advocacy sessions (96 percent), ranging between one and 20 sessions during the first six-month period.

For this intervention implementing TF-CBT in the context of a child welfare population, we had anticipated a small-to-medium intervention effect. Based on the number of participants in the study, however, we did not likely have enough statistical power to detect an effect of that size. Instead, we had an 80-percent chance to detect a slightly larger than medium effect (of size 0.576).

In the intent-to-treat analyses, there were no significant differences (or nonsignificant trends) between the intervention and control groups on any group-level test of change on the primary, secondary, or tertiary outcomes. This held true for the 12-month analyses. Analyses at later time points could not be conducted because of inadequate sample sizes. Finally, examination of the influence of different dosages of Safe Start services found no statistically significant differences in primary outcomes related to the amount of services received by intervention group families.

Overall, however, the sample size limitations mean that no conclusions can be drawn about the effectiveness of the San Diego Safe Start 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 than a large intervention effect. Given that the expected intervention effect size was small to 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.

In addition to the direct service component assessed here, San Diego's program also involved a component to improve system-level coordination. This component was intended to increase awareness of the needs of children exposed to domestic violence and increase capacity to provide appropriate services for this population. The national Safe Start process evaluation assessment found that this component did appear to have increased awareness, communication, and coordination among providers serving this population (Schultz et al., 2010). It may be that such systems improvements provide important benefits to families that can be documented in future research or captured with larger sample sizes than those available here.

REFERENCES

- 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.
- Cohen, J. A., E. Deblinger, A. P. Mannarino, and R. A. Steer, "A Multi-Site, Randomized Controlled Trial for Children with Sexual Abuse-Related PTSD Symptoms," *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 43, No. 4, 2004, pp. 393–402.
- Cohen, J., A. Mannarino, and E. Deblinger, *Child and Parent Trauma-Focused Cognitive Behavioral Therapy Treatment Manual*, unpublished manual, 2003.
- _____, *Treating Trauma and Traumatic Grief in Children and Adolescents*, New York City: Guilford Press, 2006.
- Cohen, J. A., A. P. Mannarino, and K. Knudsen, "Treating Sexually Abused Children: One Year Follow-Up of a Randomized Controlled Trial," *Child Abuse and Neglect*, Vol. 29, No. 2, 2005, pp. 135–145.
- County of San Diego, *Funding Proposal to the Office of Juvenile Justice and Delinquency Prevention—CFDA Title: Safe Start: Promising Approaches for Children Exposed to Violence*, County of San Diego, San Diego, Calif., 2004.
- King, N. J., B. J. Tonge, P. Mullen, N. Myerson, D. Heyne, R. Rollings, R. Martin, and T. H. Ollendick, "Treating Sexually Abused Children with Posttraumatic Stress Symptoms: A Randomized Clinical Trial," *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 39, No. 11, 2000, pp. 1347–1355.
- 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
- Taylor, N., A. Gilbert, G. Mann, and B. E. Ryan, *Assessment-Based Treatment for Traumatized Children: A Trauma Assessment Pathway*, San Diego, Calif.: Chadwick Center for Children & Families, 2005.

SAN DIEGO OUTCOMES APPENDIX

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

Primary Outcome	Group	Baseline		Six Months		12 Months		18 Months	
		N	Mean	N	Mean	N	Mean	N	Mean
PTSD Symptoms									
CR Child PTSD Symptoms for Ages 3–10	Intervention	49	40.94	48	37.40	44	37.93	9	33.33
	Control	45	38.47	42	37.12	34	35.74	11	34.45
SR Child PTSD Symptoms for Ages 8–12	Intervention	19	9.11	18	7.50	21	7.05	1	
	Control	15	7.93	12	5.83	17	7.18	5	5.40
Behavior/Conduct Problems									
CR Child Externalizing Behavior Problems for Ages 3–18	Intervention	46	26.15	46	23.78	45	24.07	9	24.00
	Control	45	24.56	44	23.39	37	24.08	10	25.00
CR Child Internalizing Behavior Problems for Ages 3–18	Intervention	54	15.26	50	13.88	49	13.71	9	13.00
	Control	46	14.30	46	13.85	39	13.64	11	14.73
Violence Exposure									
CR Total Child Victimization Experiences for Ages 0–12	Intervention	53	2.45	51	0.92	49	0.45	9	0.56
	Control	49	2.37	46	0.91	38	0.71	11	0.27
CR Child Maltreatment for Ages 0–12	Intervention	53	0.45	51	0.12	49	0.08	9	0.00
	Control	50	0.44	46	0.17	38	0.16	11	0.09
CR Child Assault for Ages 0–12	Intervention	53	0.45	51	0.20	49	0.12	9	0.22
	Control	50	0.48	46	0.15	38	0.21	11	0.00
CR Child Sexual Abuse for Ages 0–12	Intervention	53	0.00	51	0.04	49	0.00	9	0.00
	Control	49	0.02	46	0.00	38	0.00	11	0.00
CR Child Witnessing Violence for Ages 0–12	Intervention	50	1.42	49	0.35	48	0.10	8	0.13
	Control	43	1.33	40	0.28	37	0.16	10	0.00

Table A.1—continued

Primary Outcome	Group	Baseline		Six Months		12 Months		18 Months	
		N	Mean	N	Mean	N	Mean	N	Mean
SR Total Child Victimization Experiences for Ages 11–18	Intervention	4		3		5	0.60	0	
	Control	3		3		5	0.40	0	
SR Child Maltreatment for Ages 11–18	Intervention	4		3		5	0.00	0	
	Control	3		3		5	0.20	0	
SR Child Assault for Ages 11–18	Intervention	4		3		5	0.00	0	
	Control	3		3		5	0.00	0	
SR Child Sexual Abuse for Ages 11–18	Intervention	4		3		5	0.00	0	
	Control	3		3		5	0.00	0	
SR Child Witnessing Violence for Ages 11–18	Intervention	4		3		5	0.20	0	
	Control	3		3		5	0.20	0	
CR Caregiver Total Number of Traumatic Experiences	Intervention	54	0.15	51	0.10	49	0.06	9	0.00
	Control	50	0.14	46	0.13	39	0.05	11	0.09
CR Caregiver Experience of Any Non-DV Trauma	Intervention	54	0.06	51	0.06	49	0.02	9	0.00
	Control	50	0.08	46	0.00	39	0.03	11	0.09
CR Caregiver Experience of Any DV	Intervention	54	0.80	51	0.10	49	0.10	9	0.11
	Control	50	0.90	46	0.17	39	0.05	11	0.00

NOTES: CR = Caregiver Report; DV = domestic violence; SR = Child Self-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 the group. Comparisons were not tested when the group size was fewer than ten for either group.

Table A.2
Comparison of Means for San Diego Secondary Outcome Variables over Time

Secondary Outcome	Group	Baseline		Six Months		12 Months		18 Months	
		N	Mean	N	Mean	N	Mean	N	Mean
Depressive Symptoms									
SR Child Depressive Symptoms for Ages 8–18	Intervention	18	7.06	18	5.89	21	5.19	1	
	Control	15	7.40	13	7.62	16	7.06	4	
Behavior/Conduct Problems									
SR Teen Delinquency for Ages 11–18	Intervention	4		3		5	0.20	0	
	Control	3		3		5	0.00	0	
Social-Emotional Competence									
CR Child Affective Strengths for Ages 6–12	Intervention	33	16.12	32	16.44	36	16.00	7	18.29
	Control	31	16.58	27	16.33	28	16.64	9	16.78
SR Child Affective Strengths for Ages 11–18	Intervention	4		3		5	14.20	0	
	Control	3		3		5	16.80	0	
CR Child School Functioning for Ages 6–12	Intervention	33	21.06	32	21.41	35	20.40	7	21.43
	Control	31	20.35	26	20.12	28	21.25	9	19.67
SR Child School Functioning for Ages 11–18	Intervention	4		3		5	21.80	0	
	Control	3		3		5	23.20	0	
CR Child Cooperation for Ages 3–12	Intervention	49	11.78	49	11.67	49	0.18	9	13.11
	Control	44	11.57	41	11.54	38	0.21	11	11.18
CR Child Assertion for Ages 1–12	Intervention	52	–0.01	50	–0.17	49	0.23	9	0.36
	Control	50	0.07	46	0.05	38	0.22	11	0.12
CR Child Self-Control for Ages 1–12	Intervention	53	0.21	51	0.20	47	11.85	9	0.24
	Control	50	0.26	46	0.45	36	11.22	11	0.53
Caregiver-Child Relationship									
CR Parent Distress for Ages 0–12	Intervention	53	29.92	51	25.92	49	24.78	9	22.44
	Control	50	28.78	46	25.63	38	25.58	11	22.82

Table A.2—continued

Secondary Outcome	Group	Baseline		Six Months		12 Months		18 Months		
		N	Mean	N	Mean	N	Mean	N	Mean	
CR Parent-Child Dysfunction for Ages 0–12	Intervention	53	22.74	51	22.76	49	21.51	9	17.78	
	Control	50	23.34	46	21.30	38	21.74	11	21.45	
CR Difficult Child for Ages 0–12	Intervention	53	30.32	51	28.82	49	26.94	9	24.44	
	Control	50	30.08	46	27.85	38	26.84	11	25.73	
CR Total Parenting Stress for Ages 0–12	Intervention	53	82.98	51	77.51	49	73.22	9	64.67	
	Control	50	82.20	46	74.78	38	74.16	11	70.00	
CR Family Involvement for Ages 6–12	Intervention	33	22.73	32	23.44	36	23.28	7	25.00	
	Control	31	23.77	27	23.96	28	24.79	9	23.11	
SR Family Involvement for Ages 11–18	Intervention	4		3		5	21.20	0		
	Control	3		3		5	24.40	0		
School Readiness/Performance										
Letter Word Identification for Ages 3–18	Intervention	49	4.57	45	3.29	47	5.51	9	-1.78	
	Control	46	4.15	41	1.85	38	1.76	11	6.73	
Passage Comprehension for Ages 3–18	Intervention	48	0.92	45	-1.56	47	-1.45	9	-2.00	
	Control	48	-1.63	43	-3.74	38	-3.08	11	-1.55	
Applied Problems for Ages 3–18	Intervention	50	1.84	#	47	1.98	49	0.88	9	-1.67
	Control	47	-4.51		43	-3.23	38	-3.11	11	1.64

NOTES: CR = Caregiver Report; SR = Child Self-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 the group. Comparisons were not tested when the group size was fewer than ten for either group.

Table A.3
Comparison of Means for San Diego Tertiary Outcome Variables over Time

Tertiary Outcome	Group	Baseline		Six Months		12 Months		18 Months	
		N	Mean	N	Mean	N	Mean	N	Mean
Background and Contextual Factors									
CR Caregiver Resource Problems	Intervention	54	14.26	51	13.65	49	12.82	9	12.11
	Control	50	14.78	46	13.39	39	13.18	11	11.09
CR Caregiver Personal Problems	Intervention	54	23.04	51	21.75	49	21.00	9	19.89
	Control	50	23.70	46	21.24	39	20.59	11	18.55

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 the 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 San Diego 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	13	34.15	33.08	-1.15
		Control	13	34.62	34.69	
	Medium	Intervention	9	33.33	33.11	
		Control	9	33.67	35.89	
	High	Intervention	10	38.70	40.50	2.50
		Control	10	38.30	37.60	
Behavior/Conduct Problems						
CR Child Externalizing Behavior for Ages 3–18	Low	Intervention	12	24.50	22.42	-0.75
		Control	12	23.83	22.50	
	Medium	Intervention	5	24.60	22.60	
		Control	5	22.20	22.00	
	High	Intervention	9	22.40	24.67	
		Control	10	22.20	20.40	
CR Child Internalizing Behavior for Ages 3–18	Low	Intervention	8	13.00	12.50	
		Control	8	13.25	12.38	
	Medium	Intervention	8	14.89	13.50	
		Control	9	14.78	14.67	
	High	Intervention	10	14.50	14.80	0.70
		Control	10	14.40	14.00	
Violence Exposure						
CR Total Child Victimization Experiences for Ages 0–12	Low	Intervention	16	1.81	0.25	-0.75
		Control	16	1.81	1.00	
	Medium	Intervention	15	2.40	0.73	-0.40
		Control	15	2.40	1.13	
	High	Intervention	16	2.38	1.88	0.56
		Control	16	2.38	1.31	
CR Child Maltreatment for Ages 0–12	Low	Intervention	15	0.13	0.00	-0.13
		Control	15	0.20	0.20	
	Medium	Intervention	11	0.18	0.09	0.09
		Control	11	0.18	0.00	
	High	Intervention	14	0.43	0.29	0.21
		Control	14	0.43	0.07	
CR Child Assault for Ages 0–12	Low	Intervention	15	0.20	0.07	-0.13
		Control	15	0.13	0.13	
	Medium	Intervention	12	0.25	0.00	0.00
		Control	12	0.25	0.00	
	High	Intervention	15	0.40	0.47	0.60
		Control	15	0.67	0.13	

Table A.4—continued

Primary Outcome	Dosage	Group	N	Baseline Mean	Six-Month Mean	Difference
CR Child Sexual Abuse for Ages 0–12	Low	Intervention	16	0.00	0.00	0.00
		Control	16	0.00	0.00	
	Medium	Intervention	15	0.00	0.13	0.13
		Control	15	0.00	0.00	
	High	Intervention	16	0.00	0.00	0.00
		Control	16	0.00	0.00	
CR Child Witnessing Violence for Ages 0–12	Low	Intervention	14	1.00	0.07	0.21
		Control	14	1.29	0.14	
	Medium	Intervention	12	1.23	0.17	0.16
		Control	13	1.23	0.15	
	High	Intervention	13	1.36	1.00	0.90
		Control	14	1.50	0.14	
CR Caregiver Total Number of Traumatic Experiences	Low	Intervention	13	0.00	0.08	0.00
		Control	13	0.00	0.08	
	Medium	Intervention	13	0.08	0.08	–0.15
		Control	13	0.00	0.15	
	High	Intervention	16	0.00	0.13	0.00
		Control	16	0.00	0.13	
CR Caregiver Experience of Any Non-DV Traumas	Low	Intervention	15	0.00	0.00	0.07
		Control	15	0.07	0.00	
	Medium	Intervention	15	0.07	0.00	0.00
		Control	15	0.07	0.00	
	High	Intervention	17	0.06	0.12	0.12
		Control	17	0.06	0.00	
CR Caregiver Experience of Any Domestic Violence	Low	Intervention	13	0.77	0.00	–0.08
		Control	13	0.77	0.08	
	Medium	Intervention	14	0.86	0.07	–0.14
		Control	14	0.86	0.21	
	High	Intervention	17	0.94	0.24	0.00
		Control	17	0.94	0.24	

NOTES: CR = Caregiver Report; DV = domestic violence; SR = Child Self-Report.

* indicates statistically significant (p-value<FDR level). Data are not shown for outcomes when the cell size is fewer than five for the group. Comparisons were not tested when the group size was fewer than ten for either group.

Table A.5
Changes in Means by Dosage Group for San Diego Primary Outcome
Variables Between Baseline and 12-Month Assessment

Primary Outcome	Dosage	Group	N	Baseline Mean	12-Month Mean	Difference
PTSD Symptoms						
CR Child PTSD Symptoms for Ages 3–10	Low	Intervention	11	36.31	35.36	-0.71
		Control	13	37.46	38.54	
	Medium	Intervention	6	38.17	33.50	
		Control	6	33.50	33.50	
	High	Intervention	10	36.18	35.50	-1.93
		Control	11	36.27	38.00	
Behavior/Conduct Problems						
CR Child Externalizing Behavior for Ages 3–18	Low	Intervention	8	23.44	20.38	
		Control	9	22.78	21.00	
	Medium	Intervention	7	23.86	23.43	
		Control	7	24.29	25.29	
	High	Intervention	9	22.30	22.89	
		Control	10	22.50	21.20	
CR Child Internalizing Behavior for Ages 3–18	Low	Intervention	12	14.23	13.17	-1.09
		Control	13	14.00	13.92	
	Medium	Intervention	8	14.50	13.63	
		Control	8	13.13	13.00	
	High	Intervention	7	14.00	12.71	
		Control	8	14.75	14.00	
Violence Exposure						
CR Total Child Victimization Experiences for Ages 0–12	Low	Intervention	15	1.94	0.60	-0.21
		Control	16	1.94	0.81	
	Medium	Intervention	13	2.23	0.08	-0.46
		Control	13	2.23	0.54	
	High	Intervention	16	2.29	0.44	-0.41
		Control	17	2.29	0.76	
CR Child Maltreatment for Ages 0–12	Low	Intervention	14	0.20	0.14	-0.14
		Control	15	0.20	0.27	
	Medium	Intervention	11	0.27	0.00	-0.27
		Control	11	0.27	0.27	
	High	Intervention	14	0.47	0.07	-0.36
		Control	15	0.33	0.27	
CR Child Assault for Ages 0–12	Low	Intervention	14	0.27	0.07	-0.14
		Control	15	0.13	0.13	
	Medium	Intervention	10	0.20	0.00	0.20
		Control	10	0.40	0.00	
	High	Intervention	16	0.35	0.13	0.04
		Control	17	0.35	0.06	

Table A.5—continued

Primary Outcome	Dosage	Group	N	Baseline Mean	12-Month Mean	Difference
CR Child Sexual Abuse for Ages 0–12	Low	Intervention	15	0.00	0.00	0.00
		Control	16	0.00	0.00	0.00
	Medium	Intervention	13	0.00	0.00	0.00
		Control	13	0.00	0.00	0.00
	High	Intervention	16	0.00	0.00	0.00
		Control	17	0.00	0.00	0.00
CR Child Witnessing Violence for Ages 0–12	Low	Intervention	12	1.33	0.42	0.08
		Control	12	1.08	0.08	0.00
	Medium	Intervention	11	1.09	0.00	0.00
		Control	11	1.09	0.00	0.00
	High	Intervention	15	1.31	0.00	–0.15
		Control	16	1.31	0.13	0.00
CR Caregiver Total Number of Traumatic Experiences	Low	Intervention	13	0.00	0.15	0.08
		Control	13	0.00	0.08	0.00
	Medium	Intervention	12	0.00	0.00	0.00
		Control	12	0.00	0.00	0.00
	High	Intervention	14	0.00	0.00	0.07
		Control	15	0.07	0.00	0.00
CR Caregiver Experience of Any Non-DV Traumas	Low	Intervention	17	0.06	0.06	0.06
		Control	17	0.06	0.00	0.00
	Medium	Intervention	13	0.08	0.00	0.00
		Control	13	0.08	0.00	0.00
	High	Intervention	16	0.00	0.00	–0.06
		Control	16	0.00	0.06	0.00
CR Caregiver Experience of Any Domestic Violence	Low	Intervention	14	0.79	0.00	0.00
		Control	14	0.79	0.00	0.00
	Medium	Intervention	11	0.82	0.00	0.00
		Control	11	0.82	0.00	0.00
	High	Intervention	17	0.94	0.24	0.18
		Control	17	0.94	0.06	0.00

NOTES: CR = Caregiver Report; DV = domestic violence; SR = Child Self-Report.

* indicates statistically significant (p-value<FDR level). Data are not shown for outcomes when the cell size is fewer than five for the group. Comparisons were not tested when the group size was fewer than ten for either group.