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 M: San Mateo 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

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 MATEO COUNTY, CALIFORNIA, SAFE START OUTCOMES REPORT

ABSTRACT

The San Mateo County Safe Start program implemented a relationship-oriented, therapeutic intervention to improve outcomes for children (ages 0–7) exposed to violence who live in kinship families. The program consisted of Child-Parent Psychotherapy (CPP) sessions provided to children and their kinship caregivers in the intervention group. A full description of the program can be found in *National 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 Mateo program staff enrolled 69 kinship families in the study, with 57 percent of them retained for the six-month assessment (representing 24 intervention and 15 control families). At baseline, caregivers reported that children had been exposed to an average of nearly 4 types of violence in their lives. Kinship caregivers reported child posttraumatic stress disorder (PTSD) symptoms in the "significant" range for 40 percent of children, and 57 percent of them reported experiencing clinical levels of total parenting stress. In the six months between the baseline and six-month research assessment, 22 of the 24 families in the retained sample had received at least one session of CPP, with an average of 14 per family over the six-month period. Results of the intent-to-treat analyses showed no differences between groups over time on any of the primary, secondary, or tertiary outcome measures with adequate sample sizes to make these comparisons. Overall, however, the sample size limitations mean that no conclusions can be drawn about the effectiveness of the San Mateo Safe Start intervention as implemented on child and family outcomes. The San Mateo Safe Start model was an attempt to provide specially targeted services for kinship families caring for children exposed to violence. This is a population that has been underserved and under-studied, and therefore further development and evaluation of service models like San Mateo Safe Start specifically for this population are greatly needed.

INTRODUCTION

The San Mateo County Safe Start program operated within San Mateo County, California, located within the San Francisco Bay area. The San Mateo County Safe Start program focused on serving "kinship families," referring to families in which children

live in the care of grandparents, other relatives, or non-related extended family members. Approximately seven percent of the county's children live in such family arrangements, compared with an estimated five percent of children nationally (Edgewood Center for Children and Families, 2004). Kinship children were more likely to live in low-income households in San Mateo County than other children and as such were at elevated risk for a host of difficulties associated with poverty, such as poor health, lack of food, and difficulties in school (Edgewood Center for Children and Families, 2004). Edgewood Center Children and Families is the San Francisco Bay area's primary provider of services to kinship children and their families. In a 2004 review of its existing cases, Edgewood found that in 72 percent of their kinship families children had experienced abuse and neglect, and about 50 percent had witnessed family violence (Edgewood Center for Children and Families, 2004).

Many children in kinship care and their caregivers experience mental health challenges; however, they seldom receive services to address these issues (Ehrle and Geen, 2002). Edgewood's Kinship Support Network (KSN), the country's first comprehensive, public-private program providing services to kinship families (Edgewood Center for Children and Families, 2004), sought to respond to this service gap by pilot-testing a home-based child-caregiver psychotherapy intervention to meet the unique needs of young children exposed to violence and living in kinship families. For the Safe Start intervention, child-parent psychotherapy (CPP) was selected as the intervention model, adapted for use with kinship families, with the consultation, direct training, and ongoing clinical supervision by one of the CPP model developers, Patricia Van Horn. In earlier evaluations of CPP, Lieberman, Van Horn, and Ghosh Ippen (2005) showed medium intervention effects on PTSD symptoms and behavior problems (0.63 and 0.64, respectively). However, CPP had not previously been implemented or evaluated for effectiveness with kinship families.

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

SAN MATEO COUNTY SAFE START

- **Intervention type:** CPP
- **Intervention length:** 12 months
- **Intervention setting:** In-home, office, or other community setting
- **Target population:** Children who reside with their grandparents or other relatives (i.e., kinship care) and have been exposed to domestic violence and/or community violence and/or have experienced abuse or neglect
- **Age range:** 0–7
- **Primary referral sources:** San Mateo County Kinship Support Network

INTERVENTION

The San Mateo County Safe Start model consisted of CPP delivered by master's- and doctorate-level clinicians in the clients' home or community or at one of Edgewood's offices, averaging about one hour per week for one year. These elements are described briefly in the following paragraphs. For a full description of the San Mateo County Safe Start intervention as it was delivered, see Schultz et al. (2010).

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

While the CPP model is relatively flexible, it was designed for a parent-child dyad, mostly commonly a mother and child. In the San Mateo kinship setting, however, the primary caregiver was most often a grandparent. The site expected that it would need guidance on the appropriate strategies for adapting the CPP model to address unique kinship circumstances. To that end, it retained one of the model developers to provide initial training for the clinical staff to deliver the model and also provide ongoing, in-person weekly clinical supervision. The site's intention was to ensure that

the model was modified for the target population in a way that would not detract from the key clinical features of the model (see Schultz et al., 2010, for a more complete discussion).

In the planned delivery of the model, the therapists used the first one to three therapy sessions to establish a treatment plan that contained the issues that would serve as a particular focus of the therapy sessions and a description of the family-specific goals of treatment. In this process, the therapists also identified other needs of the family, such as individual therapy for the primary caregiver, legal assistance, or assistance obtaining food stamps, and the like. For these needs, therapists would connect families to KSN community workers or services to assist them.

Efforts to monitor the quality of the program included the Safe Start therapists and their supervisor participating in weekly two-hour clinical supervision sessions with the CPP model developer. In addition, each clinician received an in-person individual one-hour weekly session of clinical supervision with a trained member of the model developer's staff. They also received in-person clinical supervision and case consultation with their agency supervisor at other points as needed. In addition, the therapists had access to materials that described the treatment model, the Safe Start program, and the implementation plans.

METHOD

Design Overview

The design of this study was a randomized controlled trial, with randomization occurring at the family level and eligible children recruited after families were referred to the program. In addition to usual case management and support services from the KSN, the intended program design was for the intervention group to receive CPP for up to 12 months. Families assigned to the control group received usual support services as well. For both groups, child outcomes and contextual information were assessed at baseline, six, 12, 18, and 24 months. The study reported here took place between June 2006 and 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

Kinship families with children in the target age-range enrolled in the KSN and were also invited to participate in the program and study. For those interested in

participation, a six-item trauma screening inventory asking about the child's exposure to domestic violence, community violence, and abuse was administered to the caregiver. If there was a positive answer to any item, the kinship family was eligible to participate in the evaluation.

The specific eligibility criteria for the Safe Start intervention were (1) kinship families living in San Mateo County and referred by the KSN; (2) at least one child age 0–7 who had been exposed to violence, as determined by caregiver answering yes to any one of the screening questions; and (3) a primary caregiver proficient in English or Spanish and able to understand the informed consent process and the assessments.

If there were multiple eligible children, the target child to serve as the focus of the intervention was selected by the caregiver based on which child she or he felt had the greatest need.

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 three groups. One group consisted of children from birth up to 2 years of age, the second group of children was between 3 and 6 years old, and last group was children who were 7 years old.

Measures

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

Table 1
Prioritized Outcome Measures for San Mateo County Safe Start

Primary Outcome Measures			
Domain	***Source/Measure***	***Age of Child***	***Respondent***
PTSD Symptoms	Trauma Symptom Checklist for Young Children	3–7 years	Caregiver
Behavior/Conduct Problems	BITSEA and Behavior Problem Index	1–7 years	Caregiver
Social-Emotional Competence	ASQ	0–2 years	Caregiver
Social-Emotional Competence	BITSEA and SSRS (Assertion and Self-Control)	1–7 years	Caregiver
Caregiver-Child Relationship	BERS-2 (Family Involvement)	6–7 years	Caregiver
Caregiver-Child Relationship	Parenting Stress Index	All	Caregiver
Secondary Outcome Measures			
Domain	***Measure***	***Age of Child***	***Respondent***
Social-Emotional Competence	BERS-2 (School Functioning, Affective Strengths)	6–7 years	Caregiver
Cooperation	SSRS	3–7 years	Caregiver
Violence Exposure	Juvenile Victimization Questionnaire	All	Caregiver
Violence Exposure	Caregiver Victimization Questionnaire	All	Caregiver
Tertiary Outcome Measures			
Domain	***Measure***	***Age of Child***	***Respondent***
School Readiness/Performance	Woodcock-Johnson III	3–7 years	Child
Background and Contextual Factors	Everyday Stressors Index	All	Caregiver

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

Enrollment and Retention

San Mateo County Safe Start received all of its referrals from the KSN. Community workers of the KSN completed an intake process with all potentially eligible families, offered available services, and began working with the family on accessing the services in which they might be interested. After two months of involvement with the KSN, the community workers introduced the Safe Start program to the families. This period also allowed time to assist kinship caregivers with

determining and/or obtaining legal guardianship status over the children in their care. When families expressed an interest in participating in the San Mateo County Safe Start program, the community worker administered a six-item trauma screening inventory containing questions about the child's exposure to domestic violence, community violence, and abuse. If there was a positive answer to any item, the kinship family would be scheduled to meet with Edgewood research staff, which would provide a detailed explanation of the Safe Start study and services, including the random assignment of the treatment services; confirm legal guardianship; obtain consent; confirm eligibility; and conduct a baseline assessment interview. After the assessment, the data assessor implemented the random assignment procedures and informed the family and the community worker.

According to data submitted on its Quarterly Activity Reports through March 2009, San Mateo County Safe Start enrolled 83 percent of the families referred to the program. The most common reasons that families did not enroll in the study included legal guardian-related issues, such as inability to locate or obtain permission from the non-custodial parent, (53 percent), and caregiver-related issues, such as no interest (40 percent) or no time (7 percent).

In Table 2, we present the number and percentage of all enrollees who were eligible for participation at each data collection time point. San Mateo's program staff initially enrolled 69 families in the study and completed a six-month research assessment for 39 caregivers (57 percent) and 26 children (54 percent). For subsequent research assessments, the site retained from 28 to 57 percent of the families, depending on the assessment point and type.

It must be noted that the retention rate for the control group particularly was low and there was differential retention between the intervention and control groups (73 percent and 42 percent, respectively). This retention at that the six-month follow up assessment increases the potential for biased results, since there is differential retention in the intervention and the control group. Since attrition may be related to treatment factors that lead to selection bias, it can be particularly problematic when it differs across the two groups. For example, if families in more distress are more likely to leave the study overall or to drop out of one group at a higher rate than another, the results can be misleading. Unfortunately, the size of the baseline and retained sample were not adequate to allow for appropriate statistical tests of potential group differences resulting from differential retention.

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	24	16	13	13	16	12	10	12
Expected*	33	28	28	23	24	27	27	22
Retention Rate	73%	57%	46%	57%	67%	44%	37%	55%
Control								
Received	15	14	5	7	10	11	5	5
Expected*	36	31	26	23	24	31	26	23
Retention Rate	42%	45%	19%	30%	42%	35%	19%	22%
Overall								
Retention Rate	57%	51%	33%	43%	54%	40%	28%	38%

* The number of 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 and enrollment was slower than expected. This was in part caused by hesitation of staff within the KSN to introduce the Safe Start program to kinship families. There were also challenges in recruiting kinship families into the intervention because of the experimental nature of the study and the stigma of receiving mental health services. Safe Start program staff explained that elderly caregivers who often head kinship families were particularly likely to hold negative views of mental health interventions generally and to express concern about being stigmatized (e.g., considered "crazy") for participation. Attempts to address these challenges included conducting additional trainings with the community service workers to assist them in overcoming hesitation in recruiting families and providing them with ways to talk about the intervention that reduced the potential stigma surrounding the mental health component. 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).

For the intervention effect of size 0.63 commonly observed in prior studies of CPP (Lieberman, Van Horn, and Ghosh Ippen, 2005), we will have a 46.2-percent chance to detect an effect of this size, with only 35 children observed at both baseline and six months and using the nominal 0.05 significance level. On the other hand, we can expect only a 9.1-percent chance to detect a small intervention effect and a 31.6-percent chance to detect a medium effect, but we will have an 80-percent chance (using the usual 0.05 nominal) to detect a large intervention effect of 0.947. 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 those assigned to the intervention group regardless of the amount of services received. As discussed in Chapter Four of the main document (see http://www.rand.org/pubs/technical_reports/TR991-1.html), 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. Ideally, analyses would take into account the type and amount of services received to account for dosage variability. However, there were not enough families in this site's sample in order to proceed with this type of analysis. Thus, the findings presented here on the entire

intervention sample may obscure important subgroup differences by service dose received.

In the analyses of available data, we examine differences between the intervention and control groups. We present baseline and follow-up estimates of primary, secondary, and tertiary outcomes for both groups when the sample size is greater than or equal to five. We compare 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). Because the sample size did not exceed 20 in each group at any time point, we were unable to do any statistical modeling of differences in differences that includes covariates (child age, gender, race or ethnicity, family income level, and the child's violence exposure at baseline). Thus, the results reported here should be considered very preliminary, since we were unable to correct for any potential imbalance in the groups on these characteristics.

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. The FDR significance level differed for unadjusted difference in difference models because the number of statistical tests varied by outcome type. With seven test statistics conducted among the primary outcomes, this led to adopting a statistical significance cutoff of 0.007 in the unadjusted difference in difference results. With nine secondary outcomes tested, the FDR significance level adopted was 0.008. On the tertiary outcomes, two tests resulted in a significance cutoff of 0.025. In the discussion of results, we have also identified nonsignificant trends in the data, defined as those tests resulting in p-values of less than 0.05 but that did not exceed the FDR criterion for statistical significance. These trends may suggest a practical difference that would be statistically significant with a larger sample size. By the same token, however, they must be interpreted with caution, because we cannot 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 55 percent females, with an average age of 4.7 years. Most of the children (43 percent) were between 3 and 5 years old, with 32 percent age 6 or 7 and the remaining 25 percent age 2 or younger. The children in the sample were predominately Hispanic (45 percent), with fewer black (16 percent), white (9 percent), and other race/ethnicity children (30 percent). The majority (58 percent) of children had family incomes over $30,000 per year. According to the caregiver reports, children in the baseline sample had been exposed to an average of 3.7 types of violence in their lives prior to the baseline assessment. Because this is a kinship family population, nearly all caregivers (64 of 69) completing the baseline assessments were not a parent of the target child. There were no statistically significant differences between the intervention and control groups on the distribution of these background characteristics.

In the sample of 39 families retained at six months, the demographics were essentially unchanged compared with the full sample enrolled at baseline. Thus, like the baseline sample, there were no statistically significant differences at baseline between groups at six months (data not shown).

Table 3
San Mateo Safe Start Sample Characteristics for Families in Baseline Assessment Sample*

	Combined		Intervention		Control		Test for Comparison P-Value
Child Characteristics	***N***	***Mean***	***N***	***Mean***	***N***	***Mean***	
Age	69	4.7	33	4.9	36	4.6	0.51
CR Violence Exposure	57	3.7	27	3.8	30	3.6	0.79
	N	***%***	***N***	***%***	***N***	***%***	
Gender							
Male	31	44.9	14	42.4	17	47.2	0.69
Female	38	55.1	19	57.6	19	52.8	
Race/Ethnicity							
Hispanic	31	44.9	15	45.5	16	44.4	
White	6	8.7	0	0.0	6	16.7	
Black	11	15.9	5	15.2	6	16.7	
Other	21	30.4	13	39.4	8	22.2	
Caregiver Characteristics	***N***	***%***	***N***	***%***	***N***	***%***	
Family Income Level							
Less than $5,000	3	5.8	2	7.7	1	3.8	
$5,000–$10,000	3	5.8	2	7.7	1	3.8	
$10,001–$15,000	2	3.8	2	7.7	0	0.0	
$15,001–$20,000	6	11.5	3	11.5	3	11.5	
$20,001–$30,000	8	15.4	3	11.5	5	19.2	
More than $30,000	30	57.7	14	53.8	16	61.5	
Relationship to Child							
Parent-Guardian	3	4.5	2	6.1	1	2.9	0.54
Other Relationship	64	95.5	31	93.9	33	97.1	

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

Next, we examined the San Mateo 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 child PTSD symptoms that fell in the significant range for 40 percent of the sample overall, 33 percent of boys and 46 percent of girls. For the caregiver-child relationship, 55 percent of the sample had total stress levels that fell in the clinical range.

Table 4
Baseline Assessment Estimates for San Mateo Safe Start Families

	Combined		Boys		Girls	
CR PTSD Symptoms for Ages 3–7	*N*	*%*	*N*	*%*	*N*	*%*
Normal	25	58.1	14	66.7	11	50.0
Borderline	1	2.3	0.0	0.0	1	4.5
Significant	17	39.5	7	33.3	10	45.5
CR Total Parenting Stress	*N*	*%*	*N*	*%*	*N*	*%*
Parental Distress—Clinical	33	47.8	18	58.1	15	39.5
Parent-Child Dysfunctional Interaction—Clinical	38	55.1	17	54.8	21	55.3
Difficult Child—Clinical	37	54.4	21	67.7	16	43.2
Total Stress—Clinical	39	56.5	18	58.1	21	55.3

NOTE: CR = Caregiver Report.

We also examined differences between the intervention and control group at baseline for San Mateo's primary, secondary, and tertiary outcomes (see this report's appendix). There were no statistically significant differences between the intervention and control groups at baseline on any of these outcome measures (Tables A.1, A.2, and A.3, first columns).

Uptake, Dosage, and Process of Care

As described fully in the process evaluation report (Schultz et al., 2010), the San Mateo Safe Start intervention consisted of CPP therapy sessions. Tables 5a and 5b show the number of CPP sessions received by families assigned to the intervention group. Table 5a shows the services received by all families initially enrolled in the intervention group and sums across all time points reported by the program, with a maximum of 24 months of service provision. CPP sessions were provided to 85 percent of all intervention group families. The reason that the remaining families did not engage in services was not tracked for the national evaluation. The number of sessions for the participating families ranged between two and 51, with an average of 24 per family. The program reported the reason that CPP sessions ended for 14 of the 28 participating families. Of those, ten of the families were reported to have completed CPP therapy successfully, while two families dropped out or otherwise discontinued services prior to completion. Services ended for two additional families for other, unspecified reasons.

Table 5b shows the CPP sessions received just for that subgroup of intervention families who were retained in at least the six-month follow-up research assessment. In other words, this is the group of 24 families that serves as our analytic sample for examination of outcomes. In the six months between the baseline and six-month

research assessment, 22 of the 24 families in the retained sample had received at least one session of CPP, with an average of 14 sessions per family over the six-month period. The reason for the CCP therapy service ending for the analytic sample was only reported for two families in this group (and is therefore not described here).

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

Service	Number with Service	Percentage with Service*	Range	Distribution**	Mean	Median
Total CPP Therapy Sessions	28	85%	2–51	1–11 36% 12–22 28% 23–39 32% 40+ 19%	24	22

* The denominator is the 33 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
Six-Month Service Received by San Mateo Safe Start Intervention Families in the Six-Month Analysis Sample

Service	Number with Service	Percentage with Service*	Range	Distribution**	Mean	Median
Total CPP Therapy Sessions	22	92%	1–23	1–8 27% 9–14 23% 15–19 32% 20+ 18%	14	14

* The denominator is the 24 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 comparing the intervention and control groups' overall mean scores on primary, secondary, and tertiary outcomes at each follow-up assessment point (six, 12, 18, and 24 months). We then look at changes in mean scores over time for only the six- and 12-month points (as adequate data were not available for subsequent comparison). For these analyses, we first tested whether there were statistically significant changes in mean scores within the intervention group and the control group. Then, at the two assessment points, 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.

Comparison of Means Between Groups

A summary of differences between the intervention and control group at each follow-up assessment point for San Mateo's primary, secondary, and tertiary outcomes is depicted in this report's appendix. Primary outcomes include PTSD, behavior problems, social-emotional competence, and caregiver-child relationship. No statistically significant differences were observed between the groups at any follow-up assessment point on these measures (Table A.1). Similarly, there were no statistically significant differences between the groups at any follow-up assessment point on the secondary or tertiary measures (see Tables A.2 and A.3).

Mean Differences over Time

Table 6 shows differences over time for San Mateo's primary outcomes (when there were at least ten families per group assessed on the associated measure). The second column of numbers in Table 6 shows the mean change between the baseline and six-month score for each individual family. Within the intervention group only, there were two significant changes: a reduction in total parenting stress and in caregiver report of difficult child perceptions. There was also a nonsignificant trend within the intervention group toward a reduction on parental distress, but we cannot rule out that this may be due to chance because of the multiple significance tests being conducted. Within the control group, there was a significant reduction on the measure of parent distress. There was also one nonsignificant trend observed toward a reduction in behavior problems reported for control group children. However, because of the multiple significance tests being conducted, this trend did not reach statistical significance and thus may be due to chance.

By the 12-month assessment points, there was a significant change (mean increase) within the intervention group only on the child assertion measure. There was also a significant decline for control group families (compared to baseline) in caregiver-reported parental distress. At 18 months, only intervention group families were retained in adequate numbers to describe data. At this time point, there were significant within-group improvements (relative to baseline) in child assertion, parental distress, and difficult child reports. There was also a nonsignificant trend toward a reduction in total parental stress, but we cannot rule out that this difference may be due to chance because of the multiple significance tests being conducted. By the 24-month mark, there were significant improvements within the intervention group on all primary outcome measures except the child problems measure. (Data are not shown for the 12-, 18-, or 24-month time period.)

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 with the control group's mean change in scores using the statistical test of differences in differences. As Table 6 shows, there were no significant differences between the groups on any of the primary outcome measures. Likewise, there were no significant differences between the two groups on any measure at the 12-, 18-, or 24-month assessment points (data not shown). The number of observations were not sufficient to conduct analyses that adjusted for demographics.

Table 7 displays data on the secondary outcomes, where there were at least ten families assessed on the measure to allow for statistical testing. At the six-month assessment point, there were significant mean decreases for families in the intervention group on four measures of child victimization: witnessing, maltreatment, assault, and total. Within the control group, there were significant six-month declines in child witnessing violence and total child victimization. These significant declines on violence exposure measures at six months are to be expected, however, because of the different periods of observation (lifetime exposure versus six-month exposure). There was also a nonsignificant trend toward improvement within the intervention group on child cooperation and a decline in child assault within the control group. As noted above, these trends may be due to chance in light of the multiple statistical tests being performed and should be interpreted cautiously.

At the 12-month follow-up assessment, there were also significant within-group declines (relative to baseline) for both the intervention and control groups on the total child victimization and the child maltreatment measures of the violence exposure domain. There was also a significant decline within the intervention group on child assault. For the remaining two follow-up periods, there were only significant declines on child maltreatment and total child victimization at the 18-month mark (results not shown).

For secondary outcomes, the comparison of the intervention group's change in mean scores to the control group's change in mean scores from baseline to six months revealed no statistically significant differences between groups on any measure (see the third column of Table 7). Similarly, by the 12-month assessment, no statistically significant differences were detected between the groups on the amount of decline for any measure (results not shown). There were not enough control group families retained at the 18- and 24-month time points for additional comparisons.

Table 8 shows the tertiary outcomes at six months, where adequate data were available to conduct statistical tests. There was one significant within-group change,

with control group caregivers reporting a significant decline in personal problems relative to the baseline period. There were three nonsignificant trends within the intervention group. These were toward reductions in caregiver reported personal problems and the letter identification and applied problems measure of the school readiness domain. However, because of the multiple significance tests being conducted, this trend did not reach statistical significance and thus may simply be due to chance. For the remaining three time points, the caregiver personal problems measure continued to show significant declines. At the 18-month mark, the caregiver resource problem measure was also significant (in the direction of improvement relative to baseline) for the intervention group.

There were no statistically significant group-level differences in means on the tertiary outcomes at six months (see Table 8, third column) or by 12 months (results not shown).

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]
PTSD Symptoms					
CR Child PTSD Symptoms for Ages 3–10	Intervention	16	–2.88		
	Control	9			
Behavior/Conduct Problems					
CR Child Behavior Problems for Ages 1–18	Intervention	23	–0.24		0.10
	Control	15	–0.34	#	
Social-Emotional Competence					
CR Child Assertion for Ages 1–12	Intervention	22	0.11		–0.16
	Control	15	0.26		
CR Child Self-Control for Ages 1–12	Intervention	22	0.23		–0.05
	Control	15	0.29		
Caregiver-Child Relationship					
CR Parental Distress for Ages 0–12	Intervention	24	–3.46	#	0.54
	Control	15	–4.00	*	
CR Parent-Child Dysfunction for Ages 0–12	Intervention	24	–1.83		–0.43
	Control	15	–1.40		
CR Difficult Child for Ages 0–12	Intervention	24	–3.46	*	–2.53
	Control	15	–0.93		
CR Total Parental Stress for Ages 0–12	Intervention	24	–8.75	*	–2.42
	Control	15	–6.33		

[a] This column reflects within-family mean changes between the baseline and six-month scores for each group separately. * indicates a significant paired t-test of differences over time.
[b] This column reflects the group-level comparison of within-family mean changes from baseline to six months. * indicates a significant t-test of group differences.
NOTES: CR = Caregiver Report. # indicates a nonsignificant trend in the t-test ($p<0.05$ but does not meet the FDR correction threshold). Mean change estimates are not shown when the group size is fewer than ten, and comparisons are not shown when the group size is fewer than ten for either group.

Table 7
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]
Social-Emotional Competence					
CR Child Cooperation for Ages 3–12	Intervention	13	1.92	#	
	Control	8			
Violence Exposure					
CR Total Child Victimization Experiences for Ages 0–12	Intervention	21	–3.38	*	–0.30
	Control	12	–3.08	*	
CR Child Maltreatment for Ages 0–12	Intervention	20	–1.30	*	–0.07
	Control	13	–1.23	*	
CR Child Assault for Ages 0–12	Intervention	20	–0.60	*	0.09
	Control	13	–0.69	#	
CR Child Sexual Abuse for Ages 0–12	Intervention	18	0.00		0.17
	Control	12	–0.17		
CR Child Witnessing Violence for Ages 0–12	Intervention	17	–1.24	*	
	Control	9			
CR Caregiver Total Number of Traumatic Experiences	Intervention	22	–0.09		–0.09
	Control	14	0.00		
CR Caregiver Experience of Any Non-DV Traumas[c]	Intervention	24	–0.13		0.01
	Control	15	–0.13		
CR Caregiver Experience of Any Domestic Violence[c]	Intervention	24	0.00		0.00
	Control	15	0.00		

[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 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. # indicates a nonsignificant trend in the t-test ($p<0.05$ but does not meet the FDR correction threshold). Mean change estimates are not shown when the group size is fewer than ten, and comparisons are not shown when the group size is fewer than ten for either group.

Table 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]
Background and Contextual Factors					
CR Caregiver Resource Problems	Intervention Control	24 15	–0.08 –0.07		–0.02
CR Caregiver Personal Problems	Intervention Control	24 15	–1.96 –2.73	# *	0.78
School Readiness/Performance					
Letter Word Identification for Ages 3–18	Intervention Control	13 8	3.62	#	
Passage Comprehension for Ages 3–18	Intervention Control	13 8	–3.00		
Applied Problems for Ages 3–18	Intervention Control	13 8	5.54	#	

[a] This column reflects within-family mean changes between the baseline and six-month scores for each group separately. * indicates a significant paired t-test of differences over time.
[b] This column reflects the group-level comparison of within-family mean changes from baseline to six months. * indicates a significant t-test of group differences.
NOTES: CR = Caregiver Report. # indicates a nonsignificant trend in the t-test ($p<0.05$ but does not meet the FDR correction threshold). Mean change estimates are not shown when the group size is fewer than ten, and comparisons are not shown when the group size is fewer than ten for either group.

CONCLUSIONS

San Mateo's Safe Start Program provided CPP therapy to kinship families in the context of a randomized controlled trial of the intervention. San Mateo program staff enrolled 69 kinship families in the study, with 57 percent of them retained for the six-month assessment (representing 24 intervention and 15 control families). Overall, the participants in the study had substantial violence exposure, with caregivers reporting that children had been exposed to an average of nearly 4 types of violence in their lives prior to the baseline assessment. At baseline, caregivers reported significant PTSD symptoms for 40 percent of children, and 57 percent of the kinship caregivers reported clinical levels of total parenting stress. Overall, CPP sessions were provided to 85 percent of all intervention group families initially enrolled, with an average of 24 sessions per family. In the six months between the baseline and six-month research assessment, 22 of the 24 families in the retained sample had received at least one session of CPP, with an average of 14 per family over the six-month period.

In the intent-to-treat analyses, despite mean scores in the intervention and control groups in the expected directions, there were no statistically significant differences between the intervention and control groups on any measure at either the six- or 12-month time point. In other words, the analysis of data on study enrollees did not show significant improvements in outcomes for the children/kinship families who received the intervention when compared with similar children/kinship families who received the usual services available in the community.

Overall, however, the sample size limitations mean that no conclusions can be drawn about the effectiveness of the San Mateo 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 that a large intervention effect. Given that the expected intervention effect size was medium, a larger sample size might have allowed for the detection of a statistically significant intervention effect. The evaluation ended early because of funding constraints when the appropriation for Safe Start was curtailed, which may have affected the sample size.

In sum, the San Mateo Safe Start model was an attempt to provide specially targeted services for kinship families caring for children exposed to violence. This is a population that has been underserved and under-studied, and therefore further development and evaluation of service models like San Mateo Safe Start specifically for this population are greatly needed.

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.

Edgewood Center for Children and Families, *Funding Proposal to the Office of Juvenile Justice and Delinquency Prevention—CFDA Title: Safe Start: Promising Approaches for Children Exposed to Violence*, San Mateo County, Calif., 2004.

Ehrle, J., and R. Geen, "Children Cared for by Relatives: What Services Do They Need?" Washington, D.C.: The Urban Institute, *Assessing the New Federalism* Policy Brief, B-74, 2002.

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

National Child Traumatic Stress Network, "CPP: Child Parent Psychotherapy," Raleigh, N.C., 2008. As of July 21, 2011: http://www.nctsnet.org/nctsn_assets/pdfs/promising_practices/cpp_general.pdf

NCTSN—*see* National Child Traumatic Stress Network.

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

SAN MATEO OUTCOMES APPENDIX

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

	Group	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	23	46.26	18	42.50	15	38.60	12	35.17	13	31.23
	Control	24	45.21	10	44.20	12	36.67	5	34.00	7	37.71
Behavior/Conduct Problems											
CR Child Behavior Problems for Ages 1–18	Intervention	32	0.42	23	–0.03	16	–0.04	13	–0.25	13	–0.23
	Control	35	0.35	15	–0.01	14	–0.02	5	–0.26	7	0.12
Social-Emotional Competence											
CR Child Assertion for Ages 1–12	Intervention	31	–0.15	22	0.17	15	0.42	12	0.78	13	0.60
	Control	34	0.07	15	0.07	13	0.33	5	0.95	7	0.61
CR Child Self-Control for Ages 1–12	Intervention	31	–0.09	22	0.21	15	0.41	12	0.69	13	0.90
	Control	34	–0.07	15	0.06	13	0.17	5	0.87	7	0.11
Caregiver-Child Relationship											
CR Parent Distress for Ages 0–12	Intervention	33	33.18	24	28.67	16	28.31	13	25.92	13	21.77
	Control	36	31.53	15	26.13	13	24.62	5	26.80	7	27.29
CR Parent-Child Dysfunction for Ages 0–12	Intervention	33	27.27	24	24.04	16	20.31	13	22.23	13	19.69
	Control	36	23.86	15	22.20	13	20.23	5	23.00	7	22.14

Table A.1—continued

	Group	Baseline		Six Months		12 Months		18 Months		24 Months	
Primary Outcome		**N**	**Mean**	**N**	**Mean**	**N**	**Mean**	**N**	**Mean**	**N**	**Mean**
CR Difficult Child for Ages 0–12	Intervention	33	35.76	24	30.25	17	30.76	13	27.38	13	24.77
	Control	36	30.89	15	30.60	13	27.85	5	27.40	7	28.71
CR Total Parenting Stress for Ages 0–12	Intervention	33	96.21	24	82.96	16	79.06	13	75.54	13	66.23
	Control	36	86.28	15	78.93	13	72.69	5	77.20	7	78.14
CR Family Involvement for Ages 6–12	Intervention	15	21.33	9	25.11	10	23.90	9	23.78	10	24.90
	Control	14	24.21	5	23.80	7	25.43	3		4	

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 San Mateo Secondary Outcome Variables over Time

	Group	Baseline			Six Months		12 Months		18 Months		24 Months	
Secondary Outcome		**N**	**Mean**		**N**	**Mean**	**N**	**Mean**	**N**	**Mean**	**N**	**Mean**
Social-Emotional Competence												
CR Child Affective Strengths for Ages 6–12	Intervention	15	13.93		9	17.22	10	17.10	9	16.00	10	17.20
	Control	14	16.50		5	16.20	7	16.14	3		4	
CR Child School Functioning for Ages 6–12	Intervention	15	17.73		9	21.00	10	20.90	9	20.67	9	21.78
	Control	13	16.31		5	19.00	7	21.14	3		4	
CR Child Cooperation for Ages 3–12	Intervention	21	8.86	#	14	12.29	11	13.36	10	12.40	12	14.33
	Control	21	11.71		9	11.22	9	12.67	5	15.80	5	12.20
Violence Exposure												
CR Total Child Victimization Experiences for Ages 0–12	Intervention	27	3.81		24	0.42	16	0.75	13	0.15	13	7.69
	Control	30	3.57		15	0.20	14	0.21	5	0.20	7	0.00
CR Child Maltreatment for Ages 0–12	Intervention	24	1.25		24	0.04	16	0.13	13	0.00	13	0.08
	Control	31	1.35		14	0.00	14	0.00	5	0.00	7	0.00
CR Child Assault for Ages 0–12	Intervention	24	0.58		24	0.04	16	0.13	13	0.08	13	0.00
	Control	30	0.77		15	0.00	14	0.07	5	0.00	7	0.00
CR Child Sexual Abuse for Ages 0–12	Intervention	25	0.12		24	0.00	16	0.00	13	0.00	13	0.00
	Control	30	0.27		15	0.00	14	0.00	5	0.00	7	0.00

Table A.2—continued

	Group	Baseline		Six Months		12 Months		18 Months		24 Months	
Secondary Outcome		**N**	**Mean**	**N**	**Mean**	**N**	**Mean**	**N**	**Mean**	**N**	**Mean**
CR Child Witnessing Violence for Ages 0–12	Intervention	21	1.52	24	0.08	16	0.38	13	0.08	12	0.00
	Control	25	1.56	15	0.00	14	0.07	5	0.00	7	0.00
CR Caregiver Total Number of Traumatic Experiences	Intervention	33	0.12	22	0.09	15	0.13	13	0.00	13	0.08
	Control	35	0.14	15	0.07	14	0.07	5	0.00	7	0.00
CR Caregiver Experience of Any Non-DV Trauma	Intervention	33	0.09	24	0.00	16	0.00	13	0.00	13	0.00
	Control	36	0.08	15	0.07	14	0.07	5	0.00	7	0.00
CR Caregiver Experience of Any DV	Intervention	33	0.00	24	0.00	16	0.00	13	0.00	13	0.00
	Control	36	0.00	15	0.00	14	0.07	5	0.00	7	0.00

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.3
Comparison of Means for San Mateo Tertiary Outcome Variables over Time

	Group	Baseline		Six Months		12 Months			18 Months		24 Months	
Tertiary Outcome		**N**	**Mean**	**N**	**Mean**	**N**	**Mean**		**N**	**Mean**	**N**	**Mean**
Background and Contextual Factors												
CR Caregiver Resource Problems	Intervention	33	12.18	24	12.00	16	10.19	0.61	13	9.31	13	11.08
	Control	36	13.31	15	10.73	14	9.43		5	8.60	7	9.86
CR Caregiver Personal Problems	Intervention	33	24.88	24	21.54	16	20.19	0.99	13	19.46	13	18.77
	Control	36	25.17	15	21.07	14	20.21	0.99	5	23.40	7	20.71
School Readiness/Performance												
Letter Word Identification for Ages 3–18	Intervention	23	3.61	15	6.87	12	4.17	0.40	9	11.11	11	8.45
	Control	23	–0.43	10	2.50	10	7.60		5	3.00	5	0.60
Passage Comprehen-sion for Ages 3–18	Intervention	23	2.30	15	4.60	12	3.42	0.72	10	6.00	12	5.00
	Control	21	2.81	10	3.80	11	5.27		5	2.80	5	–0.20
Applied Problems for Ages 3–18	Intervention	23	–2.52	15	7.93	11	8.00		10	6.60	12	10.92
	Control	23	–0.61	10	3.80	9	5.33		5	–2.40	5	4.60

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.