Improving Outcomes for Children Exposed to Violence

Safe Start Promising Approaches

Appendix F. Philadelphia, Pennsylvania: Intervention, Study, and Results

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

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

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

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

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

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

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

Support RAND

Make a tax-deductible charitable contribution at

www.rand.org/giving/contribute

www.rand.org

Appendix F. Philadelphia, Pennsylvania: Intervention, Study, and Results

Introduction

Philadelphia Safe Start, a project of the Health Federation of Philadelphia, developed the Enhanced Home Visiting Program, which augmented standard home visiting and case management in EHS with a trauma-focused intervention for mothers of 0- to 3-year-old children.

The Philadelphia Safe Start program served nine ZIP Codes in northern Philadelphia, Pennsylvania. According to the U.S. Census Bureau's 2014 American Community Survey (U.S. Census Bureau, 2016), these ZIP Codes had a combined population of 360,164 residents, 27 percent of whom were younger than age 18 and 10 percent of whom were younger than age 5. Approximately 10 percent of the population were white, 57 percent were black, 26 percent were Hispanic, and 7 percent were Native American, Asian, or other. The 2014 median household income was $29,036, compared with the national average of $53,482, and about 33 percent of the population were living at or below the poverty threshold. Of families with children under the age of 18, 43 percent were living at or below the poverty threshold. In 2014, the violent crime rate in the larger city of Philadelphia was 591.2 per 100,000 residents, compared with the national average of 202.6 per 100,000 (City-Data.com).

The Enhanced Home Visiting Program trauma intervention, S.E.L.F. (*S.E.L.F.*, undated), is based on a group curriculum that utilizes the Sanctuary Model (Bloom, 2013). By embedding trauma screening for the caregiver and the S.E.L.F. intervention within the home visitation model of EHS, this program aimed to reduce the incidence and impact that violence has on the child, as well as to increase protective factors by strengthening the caregiver's ability to provide a safe, stable, and nurturing relationship. By doing this, the project intended to break the cycle of intergenerational transmission of trauma.

The outcome evaluation reported here presents data relevant to the question of whether the Philadelphia Safe Start program, as implemented within this project, improves outcomes for caregivers and children exposed to violence over and above EHS services (home visiting and case management).

Philadelphia Safe Start

- **Intervention components:** EHS (home visiting and case management) augmented with S.E.L.F. intervention
- **Intervention length:** Up to 26 modules delivered weekly for up to six months
- **Intervention setting:** Home
- **Target population:** All families eligible to receive services from the EHS program in several ZIP Codes of Philadelphia
- **Age range, in years:** 0–3
- **Primary referral source:** EHS

Intervention Summary

Referrals

All of Philadelphia's referrals (100 percent) came from EHS, as planned. Both existing families and new families were recruited into Enhanced Home Visiting Program. For existing families, the EHS caseworker screened caregivers for eligibility and invited them to participate. A family, whether new and existing in EHS, was eligible if the child was 28 months or younger at the time of study enrollment (meaning that the family had at least eight months of remaining eligibility in EHS). Caregivers and children were screened for trauma exposure, but this was not used to determine eligibility.

If they joined the project, a Safe Start staff member would arrange to complete the baseline assessment. For new families, caregivers were given information about the project at the time of enrollment in EHS, and, if they were interested, a Safe Start staff member attempted to join the EHS caseworker to complete the assessment at the same time as completion of EHS assessments. The Safe Start staff person scheduled independent meetings with the families if the EHS caseworker was unavailable.

Intervention Components

All families (both intervention and comparison groups) received services normally provided within EHS' home visitation child development, health and nutrition, and family support model. Services begin with a comprehensive assessment to determine service needs, goals, objectives, and priorities, followed by a support plan. Support plans involved child development and parent–child interventions provided by a multidisciplinary team, including a child and family advocate, nurse, social worker, and nutritionist. The support plan was then carried out during weekly 90-minute home visits from the child family advocate, monthly visits from the social worker, biweekly EHS center-based parent–child socialization and play groups, and quarterly educational parenting groups. Intervention services also included assisting the primary caregiver in managing family issues and dynamics; ensuring that children's health care, mental health, and academic needs are met; and helping the family navigate the myriad of social service systems to ensure that the family's basic needs (e.g., housing, education, employment, food, clothing) are met.

Families in the intervention group also had a trauma specialist added to the EHS team, who provided the S.E.L.F. intervention (*S.E.L.F.*, undated). S.E.L.F. is a psychoeducational therapeutic approach to facilitate client movement through four stages of trauma recovery:

- *safety* (attaining safety in self, relationships, environment, and beliefs)
- *emotional management* (identifying levels of affect and modulating affect in response to memories, persons, and events)
- *loss* (feeling grief, dealing with personal losses, and confronting resistance to change)
- *future* (trying out new roles and ways of relating and behaving as a survivor to ensure personal safety and help others).

It is a nonlinear intervention, meaning that modules are chosen based on need and not on any specific order. Its use in this project with individual meetings (45 to 90 minutes) at the caregiver's home differs from the standard implementation of groups in clinical and community-based organizational settings. An in-depth assessment determines the module of S.E.L.F. with which to start; then, after a module was administered, the trauma specialist followed a protocol to reassess the caregiver and choose a module for the next session.

Design Overview

The design of this study was an RCT, with randomization occurring at the family level. Both intervention and comparison families received EHS services, and the families in the intervention group also received S.E.L.F. sessions (up to 26 sessions over six months). For both groups, we assessed child outcomes and contextual information at baseline, six, and 12 months. Chapter Two summarizes and Appendix A fully describes the measures used in this study. The measures were uniform across the national evaluation but prioritized within each site as to the relevance to the intervention under study. Given the nature of the Philadelphia Safe Start intervention, the outcomes were prioritized as shown in Table F.1.

Table F.1. Philadelphia Prioritized Outcomes

Domain	Source or Measure	Child's Age, in Years	Respondent
Primary outcome measures			
Caregiver mental health	PHQ-8 depression scale	0–3	Caregiver
Secondary outcome measures			
Background and contextual	Attitudinal Barriers to Care	0–3	Caregiver
Social and emotional competence	ASQ:SE	0–3	Caregiver
Caregiver mental health	PC-PTSD	0–3	Caregiver
Tertiary outcome measures			
Background and contextual	ESI total stressors, resource problems, and personal problems	0–3	Caregiver
Behavior and conduct problems	BITSEA problem scale	1–3	Caregiver
	Child behavior problems (BPI Total Scores and BPI Externalizing subscale)	3	Caregiver
Social and emotional competence	SSIS (cooperation, assertion, and self-control)	3	Caregiver
Family functioning	FES Conflict scale	0–3	Caregiver
Child PTSD symptoms	TSCYC PTSD scale	3	Caregiver
Child depression	BPI Internalizing subscale	3	Caregiver
Violence exposure	JVQ child victimization experiences (total, child maltreatment, child assault, child sexual abuse, and child witnessing violence)	0–3	Caregiver
	Caregiver victimization (total, non-DV, and DV)	0–3	Caregiver

Study enrollment took place between January 2012 and June 2015, with follow-up assessments completed at six and 12 months after enrollment on a rolling basis. However, because of resource constraints and the desire to maximize power at six months, the site stopped collecting 12-month assessments on new enrollees partway through the study, in April 2015, meaning that follow-up assessments ended in December 2015.

Analysis Plan and Power Calculations

We conducted descriptive analyses to summarize the sample baseline characteristics: age, gender, race or ethnicity, family income level, child's violence exposure, and the outcome variables. The randomized controlled design makes it unlikely that there were differences between intervention and comparison groups, but we tested for differences in child and caregiver characteristics between intervention and comparison group children using *t*-tests and chi-squared tests to ensure that this was the case. We also examined whether those families who were lost to

follow-up differed in any systematic way from those who were retained, using *t*-tests and chi-squared tests.

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

At the time of the study planning, there was no literature evaluating the S.E.L.F. curriculum (nor its modification), so the effect size was unknown. Because of the robustness of services in both the intervention and comparison groups (EHS case management and home visiting in each), we tried to design the study so as to detect a small effect. However, with the limited capacity to do so within the Philadelphia EHS system, it was not feasible to plan a fully powered study. Table F.2 shows the enrollment by group, comparing the actual enrollment with the target enrollment needed for power, assuming an 80-percent retention rate. With total enrollment of 233, Philadelphia enrolled 37 percent of the sample size required to detect a small intervention effect. Philadelphia's overall six-month retention rate of 92 percent for caregivers in the intervention group and 81 percent for the comparison group meant that it retained a total of 196 families in the study, representing 38 percent of the retained sample required to have an 80-percent chance of detecting a small intervention effect. At 12 months, Philadelphia's retention rate, excluding the later enrollees for whom a 12-month assessment was not planned, resulted in 163 families, or 32 percent of the retained sample required to detect a small effect. Given the sample size for child outcomes, there was power to detect a medium effect of 0.40 at six months and 0.44 at 12 months, according to Cohen's 1988 effect size classification.

Table F.2. Philadelphia Required Versus Actual Enrollment for a Small Effect Size

Requirement	Intervention Group	Comparison Group	Total
Enrolled sample needed for power	319	319	638
Total enrollment	111	122	233
Percentage of needed enrollment	35	38	37
Retained sample needed for power	255	255	510
Retained sample, six months	99	97	196
Percentage of needed retention, six months	39	38	38
Retained sample, 12 months	79	84	163
Percentage of needed retention, 12 months	31	33	32

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

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

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

To examine outcomes related to the as-treated sample, those families who took part in the intervention services offered, we examined the outcome means for families who took part in the intervention services offered, broken down into groups that received a low dose of the intervention (zero to two sessions), a medium dose (three to 11 sessions), and a high dose (12 or more sessions) at six months and again at 12 months. Because children with more need are likely to receive more services, we would expect this selection scheme to possibly present an unforeseeable bias, with families more in need receiving more services. To account for this selection bias related to service dosage, we used the propensity score matching method to pair

families in each dosage group with families with similar needs in the comparison group. The matching paired families based on similar baseline scores on the outcome measure of interest. The analyses then examined the difference in differences between the intervention and comparison groups for each dosage group at follow-up. Note that, in this analysis, the full comparison group is used in the matching of each of the dosage levels. We examine only primary outcomes with this method, in recognition that it is exploratory and preliminary. Further, the power for the dosage analyses is reduced because of the smallness of the samples for the different dosage groups.

When conducting large numbers of simultaneous hypothesis tests, as we did in this study, it is important to account for the possibility that some results will achieve statistical significance simply by chance. The use of a traditional 95-percent confidence interval, for example, will result in one out of 20 comparisons achieving statistical significance as a result of random error. We therefore adjusted for false positives using the FDR method (Benjamini and Hochberg, 1995). We based our assessments of statistical significance on applying the FDR procedure separately to the primary, secondary, and tertiary outcome tests in this report (as reported in Tables F.8 and F.9) using an FDR of 0.05. In the discussion of results, we have also identified significant trends in the data, defined as those tests with *p*-values of less than 0.05 without adjusting for multiple significance tests. Although these results might suggest a practical difference that would be statistically significant with a larger sample, they must be interpreted with caution because we cannot rule out that the difference was due to chance because of the multiple significance tests being conducted.

Study Results

Enrollment and Retention

As noted above, Philadelphia enrolled a total of 233 families in the study, with 111 in the intervention group and 122 in the comparison group. In Table F.3, we present the number and percentage of all enrollees who were eligible for participation at each follow-up data collection time point. As shown, 89 percent of families enrolled in the intervention group were retained for the six-month caregiver assessment, with 80 percent retained in the comparison group. Retention for the caregiver assessment at 12 months was 83 percent for the intervention group and 86 percent for the comparison group, taking into account those families who enrolled later when a 12-month assessment was not planned.

Table F.3. Retention of Enrollees Eligible to Participate in the Caregiver Assessment at Each Time Point

Requirement	Six Months	12 Months
Intervention		
Received	99	79
Expected	111	95
Retention rate, as a percentage	89	83
Comparison		
Received	97	84
Expected	122	98
Retention rate, as a percentage	80	86

Baseline Descriptive Statistics

For the descriptive statistics, we provide the characteristics for the full sample enrolled at baseline. As shown in Table F.4, children who participated were, on average, 1.2 years old (range 0 –3), with 53 percent being male. The racial and ethnic background of families was 68 percent black, 22 percent Hispanic, 9 percent other, and 2 percent white. Families reported a range of family incomes, but the majority (69 percent) reported annual incomes below $10,000. Families reported an average exposure for the child to under one type of violence in the past six months (0.7). We observed no statistically significant differences between the intervention and comparison groups at baseline.

Table F.4. Philadelphia Safe Start Baseline Sample Characteristics

Characteristic	Combined		Intervention		Comparison		Test for Comparison *p*-Value
	N	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)	
Child							
Age	233	1.2 (0.8)	111	1.2 (0.8)	122	1.3 (0.8)	0.28
CR of violence exposure	232	0.7 (1.1)	111	0.6 (0.9)	121	0.8 (1.2)	0.18
		Percentage		**Percentage**		**Percentage**	
Child							
Gender							0.71
Male	123	53	60	54	63	52	
Female	110	47	51	46	59	48	
Race and ethnicity							0.43
Hispanic	50	22	28	26	22	18	
White	4	2	2	2	2	2	
Black	154	68	72	66	82	69	
Other	20	9	7	6	13	11	
Caregiver							
Family income level							0.80
Less than $10,000	136	69	69	72	67	66	
$10,001–20,000	27	14	12	13	15	15	
$20,001–30,000	22	11	9	9	13	13	
More than $30,000	13	7	6	6	7	7	
Relationship to child							0.95
Parent or guardian	231	99	110	99	121	99	
Other relationship	2	1	1	1	1	1	

NOTE: Because of missing data, some numbers might not sum as shown.

In the sample of families who were retained in the study at six and 12 months, these characteristics were similar to those shown in Table F.4 (data not shown). There were also no statistically significant differences between those who were lost to follow-up and those who were retained (data not shown).

We also examined the Philadelphia sample at baseline in terms of caregiver depressive symptoms to describe the level of severity on these indexes among families entering the study (Table F.5). At baseline, the majority of caregivers reported depressive symptoms in the none or minimal (34 percent) or mild (34 percent) range. Only 14 percent reported depressive symptoms in the moderately severe or severe range.

Table F.5. Baseline Assessment Estimates for Philadelphia Families: Caregiver Depression

	Combined		Boys		Girls	
Range	***N***	**Percentage**	***N***	**Percentage**	***N***	**Percentage**
None or minimal	78	34	43	35	35	32
Mild	78	34	41	33	37	34
Moderate	42	18	21	17	21	19
Moderately severe	21	9	13	11	8	7
Severe	12	5	5	4	7	6

NOTE: Because of missing data, some numbers might not sum as shown.

Finally, we examined differences between the intervention and comparison groups at baseline for Philadelphia's primary, secondary, and tertiary outcomes (see Table F.12). At baseline, there were no differences between groups for the primary, secondary, or tertiary child outcomes, indicating that the randomization resulted in balanced groups.

Table F.12 summarizes cross-sectional differences between the intervention and comparison groups at the six- and 12-month follow-up time points for Philadelphia's primary, secondary, and tertiary outcomes. At six months, there was a trend toward significant differences between the two groups on one tertiary outcome, but it did not withstand the correction for multiple testing. We note two significant differences at 12 months, with the intervention group reporting higher symptoms of depression (primary outcome) and more attitudinal barriers to care (secondary outcome) than the comparison group did.

Uptake, Dosage, and Process of Care

The program recorded family-level service data on the follow-up service survey and submitted the data at six and 12 months after the baseline assessment. Table F.6 shows the type and amount of services received for all families who were initially enrolled in the intervention group, regardless of whether they continued to participate in the ongoing research assessment. The data displayed in Table F.6 include services received by summing all time points that the program reported, which was 12 months for most participants in Philadelphia Safe Start. The site did not complete 12-month service surveys after stopping the 12-month assessments at the end of the study.

Table F.6. Services That Philadelphia Safe Start Intervention Families Received

	With Service		Number of Visits		
Service	**Number**	**Percentage**	**Range**	**Mean**	**Median**
Baseline sample (*n* = 111)					
S.E.L.F. sessions	88	79	0–26	8.8	6
Six-month analysis sample (*n* = 99)					
S.E.L.F. sessions	85	86	0–26	9.8	8
12-month analysis sample (*n* = 79)					
S.E.L.F. sessions	68	86	0–26	10.5	10

As shown in the top portion of Table F.6, 79 percent of the 111 families received some S.E.L.F. services within the intervention, with an average of 8.8 S.E.L.F. sessions. Among only the families who received S.E.L.F. services, the average was 11.1 S.E.L.F. sessions. Table F.6 also shows the services received during the six months between baseline and the six-month assessment by the subgroup of intervention group families who participated in the six-month follow-up research assessment and the subgroup that participated in the 12-month assessments. These are the families included in the intervention group in the outcome analysis sample for the Philadelphia program ($N = 99$ at six months and $N = 79$ at 12 months). As shown in Table F.6, 86 percent of the families assessed at six and 12 months received some services, with an average of 9.8 S.E.L.F. sessions in the six-month sample and 10.5 in the 12-month sample, Among only the families who received S.E.L.F. services, the average was 11.4 S.E.L.F. sessions in the six-month analysis sample and 12.2 S.E.L.F. sessions in the 12-month analysis sample. We did not collect service information on usual care in the comparison group.

On the caregiver surveys, we asked caregivers in the intervention group about their satisfaction with Safe Start services (Table F.7). Results show that the caregivers who took part in the enhanced intervention reported high satisfaction on every item (between mostly satisfied and very satisfied or the highest relevant rating on each item).

Table F.7. Satisfaction with Services That Philadelphia Safe Start Intervention Families Received

	Six Months			12 Months		
Satisfaction	***N***	**Mean**	**SD**	***N***	**Mean**	**SD**
Rate quality of service	80	3.76	0.48	50	3.88	0.33
Got the kind of service wanted	78	3.72	0.53	50	3.76	0.52
Program met needs	77	3.18	0.88	50	3.42	0.67
Would recommend to a friend	80	3.91	0.28	50	3.94	0.24
Satisfied with help received	78	3.64	0.56	50	3.72	0.64
Helped deal more effectively with problems	77	3.58	0.59	50	3.76	0.43
Satisfied with service	79	3.78	0.44	50	3.84	0.42
Would come back to program	78	3.90	0.35	50	3.90	0.3

Key Outcome Findings

We begin by analyzing changes in mean scores over time both within the intervention and comparison groups and between the groups. For these analyses, we used an intent-to-treat approach that included all families allocated to the intervention, regardless of the level of service they received. The first set of columns of numbers in Table F.8 describes differences within groups between the baseline and the six-month assessment, with paired *t*-tests comparing each person's score at each follow-up wave with his or her own score at the baseline assessment and adjusting for multiple testing. At six months, we observed statistically significant decreases in only the comparison group for the primary outcome of caregiver depressive symptoms, as well as one of the 13 tertiary outcomes, family conflict.

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

						Group-Level Comparison of Mean Changes					
		Within-Family Mean Change[a]				Unadjusted Model[b]			Adjusted Model[c]		
Outcome		***N***	**Estimate**	**SE**	***p*-Value**	**Estimate**	**SE**	***p*-Value**	**Estimate**	**SE**	***p*-Value**
Primary											
CR of caregiver depression	Intervention	98	–0.65	0.50	0.19	0.77	1.15	0.50	1.03	1.14	0.37
	Comparison	94	–1.42	0.63	0.03[d]*						
Secondary											
CR of attitudinal barriers to care	Intervention	99	–0.22	0.21	0.28	–0.02	0.34	0.94	0.35	0.36	0.33
	Comparison	96	–0.20	0.17	0.24						
CR of child social–emotional competence (ages 0–3)	Intervention	74	–3.97	2.95	0.18	–1.07	7.34	0.88	–1.64	6.34	0.80
	Comparison	75	–2.90	2.66	0.28						
CR of caregiver PTSD	Intervention	99	–0.27	0.16	0.09	–0.09	0.28	0.74	–0.22	0.28	0.43
	Comparison	95	–0.18	0.14	0.21						
Tertiary											
CR of total stressors	Intervention	99	–1.37	1.04	0.19	1.03	2.24	0.65	–0.16	2.24	0.94
	Comparison	96	–2.41	1.08	0.03[d]						
CR of resource problems	Intervention	99	–0.30	0.49	0.53	0.33	1.02	0.75	–0.06	1.07	0.95
	Comparison	96	–0.64	0.57	0.27						
CR of personal problems	Intervention	99	–1.07	0.69	0.13	0.70	1.43	0.62	–0.10	1.42	0.95
	Comparison	96	–1.77	0.66	0.01[d]						
CR of child behavior problems (ages 0–3)	Intervention	52	–0.42	1.02	0.68	–0.28	2.45	0.91	–2.03	2.31	0.38
	Comparison	56	–0.14	0.88	0.87						
CR of child behavior total problems (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
CR of child externalizing problems (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child cooperation (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child assertion (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child self-control (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	0	—	—	—						
CR of family conflict (age 3)	Intervention	98	–0.44	0.27	0.10	0.32	0.52	0.54	0.43	0.53	0.42
	Comparison	96	–0.76	0.22	<0.01[d]*						
CR of child PTSD symptoms (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child internalizing problems (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of total child victimization experiences (ages 0–3)	Intervention	99	–0.06	0.11	0.57	0.02	0.19	0.90	0.11	0.20	0.57
	Comparison	96	–0.08	0.12	0.49						
CR of child maltreatment (ages 0–3)	Intervention	99	–0.05	0.04	0.17	0.02	0.06	0.70	–0.01	0.06	0.87
	Comparison	96	–0.07	0.03	0.03[d]						
CR of child assault (ages 0–3)	Intervention	99	–0.06	0.03	0.03[d]	0.00	0.05	0.97	0.03	0.05	0.61
	Comparison	96	–0.06	0.03	0.06						
CR of child sexual abuse (ages 0–3)	Intervention	99	–0.01	0.01	0.32	0.00	0.01	0.98	0.01	0.01	0.32
	Comparison	96	–0.01	0.01	0.32						
CR of child witnessing violence (ages 0–3)	Intervention	99	0.04	0.09	0.66	0.02	0.15	0.90	0.11	0.16	0.49
	Comparison	96	0.02	0.10	0.83						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
CR of caregiver total number of traumatic experiences	Intervention	99	–0.24	0.11	0.03[d]	–0.09	0.17	0.61	–0.04	0.16	0.78
	Comparison	96	–0.16	0.09	0.07						
CR of caregiver experience of any non-DV trauma	Intervention	99	–0.01	0.05	0.84	0.04	0.08	0.58	0.06	0.08	0.41
	Comparison	96	–0.05	0.05	0.25						
CR of caregiver experience of any DV	Intervention	99	–0.13	0.04	0.00[d]	–0.07	0.07	0.29	–0.02	0.06	0.73
	Comparison	96	–0.06	0.03	0.06						

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

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

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

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

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

Table F.8 also shows intervention effects over time using an intent-to-treat approach in which all families in the intervention are compared with all those in the comparison group, regardless of the actual amount of intervention received in the intervention group. Because any change in outcomes observed can potentially be the result of a time trend observed in all children in the study, we used a difference-in-differences method to assess the unadjusted impact of the program. In the difference-in-differences models presented in the second set of data columns, there was no evidence of intervention effects associated with any of the primary, secondary, or tertiary outcomes at six months. The adjusted models in the third set of columns control for child age, child gender, child race and ethnicity, and parent report of violence exposure; they likewise indicate no difference in differences in the child outcomes. That is, changes observed within the two groups did not differ significantly from one another. The results in Table F.8 can also be discussed in terms of effect sizes for Philadelphia's primary outcome. Within-group change in the caregiver depression scale was small for the S.E.L.F. intervention (–0.13 [–0.41 – 0.15]). Further, the adjusted between-group effect size was very small (0.08 [–0.09 – 0.25]), indicating again that the two groups did not differ much in terms of their improvement.

Table F.9 presents changes in outcomes from baseline to 12 months for each group, as well as group comparisons on the 12-month outcomes. In the first set of columns, we observed improvement over time in the comparison group for caregiver depressive symptoms (primary outcome) and all three of the secondary outcomes (attitudinal barriers to care, child social–emotional competence, and caregiver PTSD symptoms). In terms of the 12 tertiary outcomes with a large enough sample size, we observed changes in the intervention or comparison group or both, including decreases in caregiver stressors (both groups), caregiver resource problems (intervention group only), caregiver personal problems (both groups), family conflict (comparison group only), caregiver traumatic experiences (intervention group only), and caregiver exposure to DV (intervention group only). In the second and third sets of columns, however, we observed no evidence of intervention effect (or difference in differences) on any of these measures.

Table F.9. Changes in Means for Outcome Variables Between Baseline and 12-Month Assessment and Group-Level Comparison of Mean Changes

						Group-Level Comparison of Mean Changes					
		Within-Family Mean Change[a]				Unadjusted Model[b]			Adjusted Model[c]		
Outcome		***N***	**Estimate**	**SE**	***p*-Value**	**Estimate**	**SE**	***p*-Value**	**Estimate**	**SE**	***p*-Value**
Primary											
CR of caregiver depression	Intervention	77	–1.12	0.65	0.09	1.90	1.26	0.13	2.19	1.28	0.09
	Comparison	83	–3.02	0.63	<0.01[d]*						
Secondary											
CR of attitudinal barriers to care	Intervention	79	–0.20	0.22	0.36	0.24	0.38	0.52	0.42	0.42	0.32
	Comparison	85	–0.45	0.20	0.03[d]*						
CR of child social–emotional competence (ages 0–3)	Intervention	42	8.48	5.19	0.11	–2.25	7.69	0.77	–2.65	7.72	0.73
	Comparison	47	10.73	4.32	0.02[d]*						
CR of caregiver PTSD	Intervention	79	–0.32	0.17	0.07	0.33	0.31	0.28	0.30	0.31	0.34
	Comparison	83	–0.65	0.16	<0.01[d]*						
Tertiary											
CR of total stressors	Intervention	79	–4.00	1.00	<0.01[d]*	–0.25	2.42	0.92	–0.92	2.43	0.70
	Comparison	84	–3.75	1.13	<0.01[d]*						
CR of resource problems	Intervention	79	–1.58	0.52	<0.01[d]*	–0.12	1.11	0.92	–0.31	1.14	0.79
	Comparison	84	–1.46	0.55	0.01[d]						
CR of personal problems	Intervention	79	–2.42	0.65	<0.01[d]*	–0.13	1.55	0.93	–0.61	1.56	0.69
	Comparison	84	–2.29	0.71	<0.01[d]*						
CR of child behavior problems (ages 0–3)	Intervention	40	–1.18	1.28	0.36	–0.13	2.50	0.96	–0.58	2.44	0.81
	Comparison	46	–1.04	1.00	0.30						
CR of child behavior total problems (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
CR of child externalizing problems (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child cooperation (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child assertion (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child self-control (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	0	—	—	—						
CR of family conflict (age 3)	Intervention	78	–0.65	0.31	0.04[d]	0.15	0.59	0.80	0.34	0.62	0.58
	Comparison	84	–0.81	0.27	<0.01[d]*						
CR of child PTSD symptoms (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of child internalizing problems (age 3)	Intervention	0	—	—	—	—	—	—	—	—	—
	Comparison	1	—	—	—						
CR of total child victimization experiences (ages 0–3)	Intervention	79	–0.01	0.13	0.92	0.13	0.24	0.59	0.15	0.25	0.56
	Comparison	84	–0.14	0.13	0.28						
CR of child maltreatment (ages 0–3)	Intervention	79	–0.05	0.05	0.32	–0.01	0.08	0.85	–0.05	0.08	0.55
	Comparison	84	–0.04	0.04	0.32						
CR of child assault (ages 0–3)	Intervention	79	–0.05	0.03	0.10	0.03	0.06	0.56	0.04	0.06	0.47
	Comparison	84	–0.08	0.05	0.07						
CR of child sexual abuse (ages 0–3)	Intervention	79	0.01	0.01	0.32	0.01	0.03	0.65	0.02	0.03	0.56
	Comparison	84	0.00	0.02	1.00						
CR of child witnessing violence (ages 0–3)	Intervention	79	0.05	0.11	0.64	0.10	0.19	0.60	0.14	0.20	0.49
	Comparison	84	–0.05	0.11	0.67						

Outcome		Within-Family Mean Change[a]				Group-Level Comparison of Mean Changes					
						Unadjusted Model[b]			Adjusted Model[c]		
		N	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value	Estimate	SE	*p*-Value
CR of caregiver total number of traumatic experiences	Intervention	79	–0.33	0.09	<0.01[d]*	–0.20	0.17	0.26	–0.11	0.16	0.47
	Comparison	84	–0.13	0.12	0.29						
CR of caregiver experience of any non-DV trauma	Intervention	79	–0.11	0.04	0.01[d]	–0.03	0.08	0.69	0.05	0.07	0.46
	Comparison	84	–0.08	0.05	0.11						
CR of caregiver experience of any DV	Intervention	79	–0.13	0.04	<0.01[d]*	–0.10	0.07	0.15	–0.12	0.07	0.07
	Comparison	84	–0.02	0.04	0.60						

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

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

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

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

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

Safe Start Service Dosage and Changes in Primary Outcomes

To examine any intervention effects that Safe Start service dosage has on outcomes, we divided the intervention families into three dosage groups because families tend to receive different amounts of services. Divided into three levels, the variable defines a low dosage as zero to two sessions, a medium dosage as three to 11 sessions, and a high dosage as 12 or more sessions.

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

At six months, the propensity score matching analyses for the primary outcome of depressive symptoms showed that, although most of the changes were in the expected direction, changes were modest even in the high-dosage group at six months (see Table F.10). These findings also show that the high-dosage group had somewhat higher depressive symptoms at baseline and thus a greater need for services. The statistical test comparing the two groups on changes in mean scores between baseline and six months revealed no significant difference between the intervention and comparison groups in any of the dosage groups. At 12 months, a similar pattern of findings emerged, though the changes across time were more pronounced than at six months (see Table F.11).

Table F.10. Changes in Means by Dosage Group for Primary Outcome, Caregiver Report of Caregiver Depression, Between Baseline and Six-Month Assessment

			Baseline		Six Months		
Dosage	**Group**	***N***	**Mean**	**SD**	**Mean**	**SD**	**Difference in Differences**
Low	Intervention	29	7.07	4.98	6.21	4.74	–1.07
	Comparison	29	7.28	5.96	7.48	5.21	
Medium	Intervention	43	7.07	5.30	6.56	4.80	1.36
	Comparison	43	7.78	6.16	5.91	4.47	
High	Intervention	22	8.45	5.03	7.77	4.89	0.94
	Comparison	22	8.34	6.26	6.73	5.52	

NOTE: We do not show data for outcomes for which the cell size is less than five for the group. We did not test comparisons for group sizes less than ten for either group.

Table F.11. Changes in Means, by Dosage Group, for Primary Outcome, Caregiver Report of Caregiver Depression, Between Baseline and 12-Month Assessment

			Baseline		12 Months		
Dosage	**Group**	***N***	**Mean**	**SD**	**Mean**	**SD**	**Difference in Differences**
Low	Intervention	22	7.77	4.91	6.36	6.16	1.59
	Comparison	22	8.82	5.40	5.82	5.63	
Medium	Intervention	22	7.45	6.70	7.27	6.37	3.23
	Comparison	22	6.95	5.05	3.55	4.25	
High	Intervention	27	8.37	4.66	6.41	4.53	0.32
	Comparison	27	7.73	6.48	5.44	4.87	

NOTE: We do not show data for outcomes for which the cell size is less than five for the group. We did not test comparisons for group sizes less than ten for either group.

Conclusions

Philadelphia's Safe Start program, Enhanced Home Visiting Program, augmented standard home visiting and case management in EHS with the S.E.L.F. trauma-focused intervention and was evaluated in an RCT comparing the S.E.L.F. intervention plus EHS with a comparison group that received EHS services only. In the study, the Safe Start program recruited 233 families and retained 196 of them at six months, representing an 84-percent retention rate. The families in the study were largely minorities (68 percent black, 22 percent Hispanic) and impoverished (69 percent had family incomes less than $10,000), with nearly one-third of caregivers (32 percent) reporting depressive symptoms in the moderate to severe range. In terms of uptake of services, 79 percent of the intervention group families received the S.E.L.F. individual therapy sessions, with the vast majority (86 percent) of those retained in the six- and 12-month samples taking part in the S.E.L.F. therapy sessions.

In this trial, we expected a small effect size due to the robust services offered within the usual EHS services that both groups received. Despite very strong retention, the EHS program had limited capacity to recruit the large sample needed to detect a small effect. Given the final enrollment and retention, there was power to detect only a medium effect of about 0.40 at both six and 12 months. Intent-to-treat analyses showed that mean scores in the intervention and comparison groups were in the expected direction across all outcome measures, but, overall, there was no evidence of intervention effects at six or 12 months for any of the outcomes. Examination of families who received low, medium, and high doses of the S.E.L.F. intervention likewise did not reveal any statistically significant differences between families in the intervention group and comparable families in the comparison group. Satisfaction with the intervention services was high.

In sum, the study was not powered to detect the small differences that might be expected with this set of intervention activities, so the study cannot determine whether there is such an effect. More research would be needed to determine whether these activities can produce small improvements in child and family outcomes over and above the usual services that EHS offers.

Table F.12. Comparison of Means for Philadelphia Outcome Variables

		Baseline				Six Months				12 Months			
Outcome		***N***	**Mean**	**SD**	**Difference**	***N***	**Mean**	**SD**	**Difference**	***N***	**Mean**	**SD**	**Difference**
Primary													
CR of caregiver depression	Intervention	110	7.64	5.42	–0.16	99	7.07	5.10	0.41	78	6.90	5.73	1.81[a]*
	Comparison	121	7.80	6.37		95	6.66	5.59		84	5.08	5.13	
Secondary													
CR of attitudinal barriers to care	Intervention	111	1.56	1.96	0.26	99	1.32	1.74	0.20	79	1.48	1.69	0.59[a]*
	Comparison	122	1.3	1.54		96	1.13	1.45		85	0.89	1.3	
CR of child social–emotional competence (ages 0–3)	Intervention	92	35.38	31.99	5.55	75	32.78	33.82	3.87	51	33.38	31.22	1.22
	Comparison	108	29.83	27.75		75	28.9	27.87		55	32.16	28.39	
CR of caregiver PTSD	Intervention	111	1.67	1.47	0.20	99	1.47	1.39	0.18	79	1.39	1.38	0.40
	Comparison	121	1.46	1.40		96	1.29	1.34		84	0.99	1.28	
Tertiary													
CR of total stressors	Intervention	111	40.44	11.04	0.71	99	39.16	11.62	1.71	79	37.35	10.86	1.03
	Comparison	122	39.73	11.40		96	37.45	10.26		84	36.32	10.65	
CR of resource problems	Intervention	111	14.32	5.20	0.19	99	13.99	5.08	0.48	79	13.19	4.92	0.58
	Comparison	122	14.12	4.98		96	13.51	4.95		84	12.61	4.95	
CR of personal problems	Intervention	111	26.13	7.02	0.52	99	25.17	7.47	1.23	79	24.16	6.85	0.45
	Comparison	122	25.61	7.38		96	23.94	6.52		84	23.71	6.79	
CR of child behavior problems (ages 0–3)	Intervention	57	14.84	9.18	1.18	54	14.54	9.20	1.60	77	13.09	7.06	1.79
	Comparison	70	13.66	8.33		60	12.93	8.73		83	11.3	6.66	
CR of child behavior total problems (age 3)	Intervention	0	—	—	—	3	—	—	—	25	8.68	4.99	0.49
	Comparison	1	—	—		1	—	—		26	8.19	4.62	
CR of child externalizing problems (age 3)	Intervention	0	—	—	—	3	—	—	—	25	6.08	3.6	–0.46
	Comparison	1	—	—		1	—	—		26	6.54	3.89	

Outcome		Baseline				Six Months				12 Months			
		N	Mean	SD	Difference	*N*	Mean	SD	Difference	*N*	Mean	SD	Difference
CR of child cooperation (age 3)	Intervention	0	—	—	—	3	—	—	—	23	10.52	5.14	–1.29
	Comparison	1	—	—		1	—	—		26	11.81	3.95	
CR of child assertion (age 3)	Intervention	0	—	—	—	3	—	—	—	23	12.87	5.86	–0.75
	Comparison	1	—	—		1	—	—		26	13.62	4.07	
CR of child self-control (age 3)	Intervention	0	—	—	—	2	—	—	—	19	7.58	5.86	–1.27
	Comparison	0	—	—		0	—	—		20	8.85	4.48	
CR of family conflict (age 3)	Intervention	110	3.93	2.65	0.36	98	3.61	2.54	0.72[a]	79	3.56	2.72	0.72
	Comparison	122	3.56	2.61		96	2.88	2.55		84	2.84	2.61	
CR of child PTSD symptoms (age 3)	Intervention	0	—	—	—	3	—	—	—	25	36.64	9.16	0.10
	Comparison	1	—	—		1	—	—		26	36.54	7.71	
CR of child internalizing problems (age 3)	Intervention	0	—	—	—	3	—	—	—	25	2.60	1.89	0.95
	Comparison	1	—	—		1	—	—		26	1.65	1.29	
CR of total child victimization experiences (ages 0–3)	Intervention	111	0.57	0.87	–0.18	99	0.55	0.80	–0.06	79	0.62	1.08	–0.05
	Comparison	121	0.75	1.19		96	0.60	0.93		84	0.67	1.15	
CR of child maltreatment (ages 0–3)	Intervention	111	0.10	0.33	–0.03	99	0.06	0.24	0.02	79	0.09	0.33	–0.01
	Comparison	122	0.13	0.38		96	0.04	0.20		84	0.10	0.37	
CR of child assault (ages 0–3)	Intervention	111	0.07	0.26	–0.02	99	0.01	0.10	–0.03	79	0.03	0.16	–0.01
	Comparison	122	0.09	0.32		96	0.04	0.20		84	0.04	0.19	
CR of child sexual abuse (ages 0–3)	Intervention	111	0.01	0.09	0.00	99	0.00	0.00	0.00	79	0.03	0.16	0.01
	Comparison	121	0.01	0.09		96	0.00	0.00		84	0.01	0.11	
CR of child witnessing violence (ages 0–3)	Intervention	111	0.38	0.66	–0.14	99	0.44	0.67	–0.02	79	0.44	0.86	–0.05
	Comparison	122	0.52	0.87		96	0.47	0.77		84	0.49	0.95	
CR of caregiver total number of traumatic experiences	Intervention	111	0.52	0.93	0.10	99	0.31	0.74	0.05	79	0.15	0.48	–0.15
	Comparison	122	0.42	0.86		96	0.26	0.65		84	0.30	0.80	
CR of caregiver experience of any non-DV trauma	Intervention	111	0.19	0.39	0.03	99	0.18	0.39	0.07	79	0.08	0.27	–0.02
	Comparison	122	0.16	0.37		96	0.11	0.32		84	0.10	0.30	

Outcome		Baseline				Six Months				12 Months			
		N	Mean	SD	Difference	*N*	Mean	SD	Difference	*N*	Mean	SD	Difference
CR of caregiver experience of any DV	Intervention	111	0.18	0.39	0.02	99	0.06	0.24	–0.02	79	0.04	0.19	–0.08
	Comparison	122	0.16	0.36		96	0.08	0.28		84	0.12	0.33	

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