Portland State University

PDXScholar

Criminology and Criminal Justice Faculty Publications and Presentations

Criminology and Criminal Justice

1-2013

Evaluating What Works for Helping Children Exposed to Violence: Results from Nine Randomized Controlled Trials

Laura J. Hickman
Portland State University

Lisa Jaycox
RAND Corporation

Claude Setodji
RAND Corporation

Aaron Kofner
RAND Corporation

Dana Schultz
RAND Corporation

Follow this and additional works at: https://pdxscholar.library.pdx.edu/ccj_fac

 Part of the Criminology and Criminal Justice Commons

See next page for additional authors

Let us know how access to this document benefits you.

Citation Details

Hickman, Laura J.; Jaycox, Lisa; Setodji, Claude; Kofner, Aaron; Schultz, Dana; Barnes-Proby, Dionne; and Harris, Racine, "Evaluating What Works for Helping Children Exposed to Violence: Results from Nine Randomized Controlled Trials" (2013). *Criminology and Criminal Justice Faculty Publications and Presentations*. 1.
https://pdxscholar.library.pdx.edu/ccj_fac/1

This Pre-Print is brought to you for free and open access. It has been accepted for inclusion in Criminology and Criminal Justice Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Authors

Laura J. Hickman, Lisa Jaycox, Claude Setodji, Aaron Kofner, Dana Schultz, Dionne Barnes-Proby, and Racine Harris

This pre-print is available at PDXScholar: https://pdxscholar.library.pdx.edu/ccj_fac/1

Evaluating What Works for Helping Children Expose to Violence: Results from Nine Randomized Controlled Trials

Laura J. Hickman,[1] Lisa Jaycox,[2] Claude Setodji,[2] Aaron Kofner,[2] Dana Schultz,[2] Dionne Barnes-Proby, [2] & Racine Harris[2]

[1] Portland State University
[2] RAND Corporation

Corresponding Author:
Laura Hickman
Associate Professor
Criminology & Criminal Justice Division
Hatfield School of Government
College of Urban and Public Affairs
Portland State University
P.O. Box 751
Portland, Oregon 97207-0751
Lhickman@pdx.edu
Voice: (503) 725-5201
Fax: (503) 725-4014

Submitted for Final Review: Journal of Experimental Criminology January 2013

Abstract

Objectives

The study tests whether participation in interventions offered by a subset of sites from the National Safe Start Promising Approaches for Children Exposed to Violence initiative improved outcomes for children relative to controls.

Methods

The study pools data from the nine Safe Start sites that randomized families to intervention and control groups, using a within-site block randomization strategy based on child age at baseline. Caregiver-reported outcomes, assessed at baseline, six and 12 months, included caregiver personal problems, caregiver resource problems, parenting stress, child and caregiver victimization, child trauma symptoms, child behavior problems, and social-emotional competence.

Results

Results revealed no measurable intervention impact in intent-to-treat analyses at either six- or twelve-month post-baseline. In six-month as-treated analyses, a medium to high intervention dose was associated with improvement on two measures of child social-emotional competence: cooperation and assertion. Overall, there is no reliable evidence of significant site-to-site effect variability, even in the two cases of significant intervention effect.

Conclusions

Since families in both the intervention and control groups received some degree of case management and both groups improved over time, it may be advantageous to explore the potential impacts of crisis and case management separately from mental health interventions. It

may be that, on average, children in families whose basic needs are being attended to improve substantially on their own.

Keywords: Children Exposed to Violence, Case Management, Multi-Site Evaluations, National Safe Start Evaluation, Violence, Victimization

Introduction

There has been a growing awareness of the issue of children's exposure to violence and the multiple harms that may result from this exposure. Definitions of children's exposure to violence (CEV) vary somewhat but typically the term refers to children and adolescents witnessing violent or threatening acts at home (e.g. domestic violence) or other settings such as school and communities. Definitions also include child violent victimization and maltreatment including neglect and other abuse (Kracke & Hahn, 2008). The National Survey of Children's Exposure to Violence recently reported that 61 percent of children had experienced or witnessed violence in the last year, with many exposed to multiple forms of violence (Finkelhor, et al., 2009). Child protective services agencies received 3.3 million referrals for neglect and abuse in 2009 (Department of Health and Human Services, 2011) and estimates based on law enforcement data indicate that nearly half of domestic violence incidents include child witnesses (Fantuzzo, et al., 2007; Fusco & Fantuzzo, 2009).

Aside from the obvious potential for physical injury, prior studies suggest that CEV is associated (directly or indirectly) with a variety of harms. For example, CEV has been linked with posttraumatic stress disorder (PTSD; Berman, et al. 1996; Breslau, et al., 1997; Wolfe, et al., 2003), depression (Kliewer, et al., 1998), anxiety (Singer, et al., 1995), and behavioral or developmental problems (Bell & Jenkins, 1993; Bourassa, 2007; Farrell & Bruce, 1997; Garbarino, et al., 1992; Herrenkohl, et al., 2008; Martinez & Richters, 1993; Sternberg, et al., 2006). CEV- related symptoms can also impact children in school. Poorer school functioning and academic performance has been found among children exposed to community violence (Bowen & Bowen, 1999; Delaney-Black et al., 2002; Hurt, et al., 2001; Schwartz & Gorman, 2003) and school violence (Grogger, 1997). Moreover, a growing literature finds that harms may be more

likely or more pronounced among those children exposed to violence of more than one type (Hickman et al., in press, Turner, Finkelhor & Ormrod, 2010, Finkelhor, Ormrod, Turner, 2009).

As evidence mounts about the widespread and negative impacts of CEV, interest has turned toward developing and testing potential interventions to help ameliorate the harms resulting from it. Among those efforts is the National Safe Start Initiative for Children Exposed to Violence. The U.S. Department of Justice's Office of Juvenile Justice and Delinquency Prevention (OJJDP) launched the Safe Start Initiative in 2000. It is a community-based effort focused on developing, fielding, and evaluating interventions with the goal of preventing and reducing the impact of CEV. Ending in 2006, Phase 1 of the Safe Start Initiative consisted of implementing demonstration projects of various innovative program approaches to addressing CEV (Hyde et al., 2008).

Phase 2, entitled Safe Start Promising Approaches (SSPA), was launched in 2005 and data collection proceeded through 2010. It was intended to implement and evaluate evidence-based and promising approaches to helping children and their families who had been exposed to violence. OJJDP utilized a peer review panel of researchers and subject matter experts to identify the program proposals that demonstrated the strongest evidence base and appeared to be the most feasible to implement (both in terms of program and evaluation implementation).

The SSPA programs employed a range of intervention components. All included a therapeutic component and in some cases the modality offered by the site varied by child age, with dyadic or family therapy for caregivers of younger children and group therapy for older children. Most programs also offered case management and some established or enhanced interagency service coordination for families. Some of the programs had other intervention components, such as family or child-level advocacy, parent/caregiver groups, or other services

(e.g., assessment of family needs or in-home safety assessments). The intervention setting also varied with services delivered in families' homes, clinics, shelters, social service agency offices, or Head Start classrooms. The intervention length ranged from three months to more than one year, though most were six months. Eligible child age varied but most programs enrolled only young children and their caregivers/families. See Schultz et al. (2010) for a detailed description of each program and its implementation.

Longitudinal data were collected from caregivers and children on standardized, age-appropriate measures at 6-, 12-, 18-, and 24-months post-enrollment. The data included demographic, violence exposure, negative symptoms (post-traumatic stress, depression, behavior problems, parenting stress), and social-emotional competence measures. Caregivers reported on their own and family outcomes as well as their perceptions of outcomes for younger children, and older children completed self-report measures. Randomized controlled trials were conducted in 10 sites, with the remainder implementing quasi-experimental designs.

In the SSPA national outcome evaluation (Jaycox et al., 2011), recruitment and retention in many sites met or exceeded typical service retention rates for mental health programs (Wierzbicki & Pekarik, 1993). Nonetheless, sample sizes were inadequate to afford sufficient statistical power for testing intervention effects of the individual programs (see Jaycox et al., 2011). Thus, no clear conclusions could be drawn about whether any individual Safe Start program represented a potentially effective model to address children's exposure to violence.

Given the rigorous design of SSPA evaluation at each individual program, the common data collected, and the general similarity of interventions, the limitations of the national evaluation's within-site analyses can be overcome through analyses that pool data across programs. The present study undertakes these analyses using data from nine SSPA randomized

control trials. The study tests whether (on average) the SSPA interventions make a difference for participants on key outcomes that all were designed to achieve. While not informative about the impact of specific programs, the pooled analysis addresses a question similar to that addressed by a meta-analysis, but using a superior approach to the combination of individual studies.

A meta-analysis "seeks to combine independent studies to identify consistent effects across criminal justice settings or contexts" (Weisman & Taxman, 2000: 316). Weisman and Taxman describe meta-analyses as a commonly used strategy for combining research results but identify a number of common problems with this approach. Key among these problems is the blending of different methodologies and measures, clouding the interpretation of the results particularly when paired with broad variability in intervention content and sample characteristics. A vast improvement on combining multi-site studies, Weisman and Taxman (2000) argue, is the multicenter clinical trial approach. In the method, a single intervention is implemented in multiple sites and evaluated experimentally using within-site randomization strategies and standardized measures and data collection procedures. The resulting data may then be pooled and analyzed to increase statistical power and support clearer conclusions about the effectiveness of the overall intervention model.

While the SSPA evaluation shares some similarities with a multicenter clinical trial, the SSPA interventions cannot be defined as such due to the variation in individual interventions. By the same token, the SSPA interventions share far more in common than studies combined in most well-known and frequently cited meta-analysis studies in criminology (e.g. Andrews et al., 1990; Lipsey 1992). Thus, in terms of clarity of interpretation about a specific intervention approach, the present study falls somewhere between a multicenter clinical trial and a meta-analysis. The shared research design, common goals and measures, and extensive coordination

and oversight by a single independent evaluation team make the interpretation of the findings much more clear than most published meta-analysis studies. The interventions themselves arguably shared more in common than the very well-regarded quasi-experimental evaluation of a "systems of care continuum" implemented in Fort Bragg, North Carolina for children and youth referred to mental health services (Bickman, Sumerfelt, & Noser, 1997). In short, the study overcomes the major limitations of a meta-analysis via the advantages of a multicenter clinical trial in assessing whether, on average, a set of interventions that combine case management and evidence-based mental health services designed to improve outcomes for children exposed to violence.

Randomizing SSPA Programs

To facilitate clearer interpretation of results, we have restricted the analyses to include only those SSPA programs that were evaluated using the most rigorous of research designs, i.e., a randomized controlled trial. Of the ten SSPA programs implemented utilizing such designs, one program was excluded here because it was substantially different from the rest, both in intervention delivery method and randomization strategy (i.e. entire Head Start classrooms were randomized to receive a curriculum-based Safe Start intervention or to a standard curriculum). The remaining nine programs randomized individual families to an intervention or control group. One of these nine programs utilized a wait-list design which randomized families to immediately begin its intervention program or to a wait-list group which would be eligible to begin the intervention after six months. These families, therefore, are not included in the 12-month analysis reported here.

The nine SSPA programs were situated locally within different kinds of lead agencies or organizations, including a health clinic, several human services agencies, a university

department, domestic violence or child maltreatment services agencies, and a county-level government office. The programs varied in their source of referrals, including health care systems, child welfare systems, domestic violence shelters, and human services agencies. The intervention length ranged from three months to one year, with length in most programs determined by family need rather than by a pre-determined number of sessions or services.

All interventions' therapy component (dyadic or family) focused to a greater or lesser extent on caregiver-child interaction and caregiver management of the emotional and/or behavioral impacts of violence exposure on children. All interventions also contained a case management component that assisted caregivers with accessing for basic needs (such as food, housing, clothing, employment, subsidized childcare, and healthcare). This may have been integrated into therapy sessions or coordinated with a case management service provider. Some of the sites had additional intervention components, such as caregiver or child-level advocacy, parent/caregiver support groups, and provider service coordination meetings. All interventions included an emphasis on reducing the risk of child repeat exposure to violence. This was largely through caregiver education about the impact of violence exposure and support to end domestic violence or other living circumstances that increased exposure risk.

For all programs, control group conditions included either services-as-usual or some enhancement. These were most often limited case management services or monthly check-in contacts and provision of community service referrals. A brief summary of the nine individual programs is provided here and in Table 1 (see Schultz et al. 2010 for a detailed description of each program model and its implementation).

Child-Parent Psychotherapy. Four of the nine programs implemented Child-Parent Psychotherapy (CPP) as the therapy component. The approach is designed to restore the parent-

child relationship, the child's mental health, and the child's development following exposure to violence, as well reduce the risk of re-exposure. The overarching goals of CPP include increasing the caregiver's and child's (age-appropriate) capacity to be emotionally attuned to each other's needs and changing negative patterns of interaction into positive and nurturing ones. Case management services are built in to the CPP model (NCTSN, 2008, Lieberman and Van Horn, 1995, 2008). In one SSPA program, the therapist provided both CPP therapy and case management, while three had a separate provider for case management. In addition to CPP, one SSPA program also included an extensive developmental assessment and plan for each child (and family). In one program, 12 CPP sessions were provided over three months with the remaining three programs providing CPP and other services for up to one year. Two programs provided CPP sessions in a clinic setting, while two worked in-home. One program served children ages five and younger who had been exposed to domestic violence. The remaining three programs recruited families with children exposed to all forms of violence of varying ages.

Family-Centered Treatment. One program implemented the Family-Center Treatment model, delivered to families where a child (age birth to six) had been exposed to violence, primarily focusing on domestic violence). The model was designed to stabilize the family, build on family strengths, and support families in taking up other needed services in the community (IFCS, 2004). Therapy sessions involved the whole family and individual members were directed toward accomplishing goals such as improving parenting, safety planning to reduce the risk of repeat violence exposure, coping with loss and separation, reducing specific behavioral or emotional problems in the parent or child, and overcoming effects of trauma. All sessions were delivered in-home, by the same therapist, typically over the course of six months. Integrated case management services were also provided by therapists, who helped families identify and access

needed services and facilitated monthly team meetings between each family and other social service providers.

Project SUPPORT. One program implemented Project SUPPORT, targeting children ages 3-9 who exhibit clinical levels of conduct problems upon exit from domestic violence shelters with their mothers. The intervention addressed children's mental health problems related to domestic violence exposure, particularly conduct problems and symptoms of depression and trauma (McDonald, Jouriles, and Skopp, 2006). The intervention sessions employed a behavior training model that involves assessing mothers' current parenting knowledge and skills, and providing education and training to enhance a specific skill set. Therapists devoted a portion of each session to case management activities to assist families with obtaining such things as food, clothing, rental assistance, child care, employment assistance, and health care. The sessions were provided by a single therapist within the context of weekly home-based sessions over a six month period.

Caregiver-Child/Infant Psychotherapy. One program implemented Caregiver-Child/Infant Psychotherapy, a dyadic approach for caregivers and children ages five and younger exposed to all forms of violence. It was targeted toward improving child social functioning, building positive caregiver-child interaction and attachment, and identifying root causes of maladaptive child behavior (such as anxiety, depression, or impulse control). The therapy involved feedback based on observation of caregiver-child interaction and education of caregivers about the negative behavioral and emotional impacts of exposure to violence on children. Therapists also assessed each family's basic needs (e.g. food, housing, employment, etc.) and delivered intensive case management services in combination with therapy to address these needs as well as reduce the risk of repeat violence exposure. Weekly sessions were

delivered in-home over a six-month period and sometimes involved additional family members as deemed necessary by the therapist.

Trauma-Focused Cognitive-Behavioral Therapy. One program employed Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) with children ages 3-9 exposed to domestic violence and identified within a child welfare population. TF-CBT targets symptoms of PTSD that often co-occur with depression, anxiety, and behavior problems (Cohen, Mannarino, and Deblinger, 2003). It seeks to teach children skills to cope with the difficulties related to PTSD and with confronting the underlying traumatic experience (i.e. violence exposure). Child sessions were accompanied by individual caregiver meetings, where therapists provided education on trauma, behavior problems, and parenting skills and strategies. The therapy involved one or more assessment sessions and approximately 20 therapy sessions, delivered weekly in a clinic setting over a six-month period. The SSPA program also held regular service coordination meetings between the child welfare case manager, therapist, and a Safe Start advocate. Advocacy services focused on domestic violence–related services to address family needs and reduce risk of repeat exposure, including instrumental support, housing assistance, and legal support and assistance. Advocacy services varied in length depending on family need.

Mixed Modality Therapy. One program's therapy component (for children birth to 12 exposed to all forms of violence) involved both home-based therapy for individual families and caregiver education groups. The "individualized" family component involves an initial assessment of the target child's developmental history, family situation, and home environment. This was used to develop an integrated treatment plan. Caregiver education groups were offered as a compliment to the individual family home-based sessions. The 12 weekly group sessions focused on expanding parenting knowledge, improve caregiver-child bonding, and building child

management and child protection skills. A designated case manager also assisted families with obtaining community services and support for as long as they participated in the research study.

-- Insert Table 1 About Here -

Methods

Hypotheses

The primary goals of the SSPA programs were to reduce repeat violence exposure and intervene with both children and caregivers in ways expected to reduce the negative impact of violence exposure on children. Among children, the interventions sought to reduce child conduct and behavior problems (both internalizing and externalizing), reduce post-traumatic stress disorder symptoms, and improve social-emotional competence. Among caregivers, the programs were expected to reduce parenting stress, reduce perceptions of child difficulty, improve caregiver-child interaction, and improve the overall caregiver-child relationship. The case management component of the programs sought to alleviate the everyday stressors of families, such as resource problems and other stressors that often co-occur in families of children exposed to violence. Thus, the hypotheses tested in the present study are as follows:

Child Outcomes

H1 Children participating in the SSPA interventions will be less likely to have violence re-exposure compared to non-participants.

H2 Children participating in SSPA will exhibit fewer behavior problems and symptoms of post-traumatic stress relative to non-participants.

H3 Children participating in SSPA will show greater social-emotional competence than non-participants.

Caregiver Outcomes

H4 Caregivers participating in the SSPA interventions will be less likely to experience domestic violence victimization than non-participants.

H5 Caregivers participating in the SSPA interventions will report reduced overall parenting stress relative to non-participants.

H6 Caregivers participating in the SSPA interventions will report their children are less difficult relative to the children of non-participating caregivers.

H7 Caregivers participating in the SSPA interventions will report reduced parental distress relative to non-participating caregivers.

H8 The caregiver-child relationship of SSPA participants will exhibit less dysfunction relative to the caregiver-child relationship of non-participants.

H9 Caregivers participating in SSPA will report fewer resource and personal problems relative to non-participants.

Screening and Randomization Procedures

Eligibility screening procedures varied by site, as did the approved Institutional Review Board procedures for obtaining informed consent (from primary caregivers and legal guardians) and child assent (for children age 7 and older). At a minimum, enrollment required a willing English- or Spanish-speaking primary caregiver who had lived with the child for at least 30 days and a child exposed to violence (in some form) falling within each program's specific age range. If more than one child fell within the study's eligibility criteria, a "target" child was selected to serve as the focus of the longitudinal research assessments (usually according to the most recent birthday, though some sites asked caregivers to select the child based on need). Additional family members could participate in the intervention services, dependent upon the program model.

Data on the characteristics of potentially-eligible families who declined participation were not systematically collected across sites. For seven of the nine sites, enrollment ranged from 41 to 83 percent of those deemed eligible for recruitment. Data from two sites were inadequate to derive an overall percentage of eligible families actually recruited.

Thus, the sample for the present study is defined as consisting of families identified by local social service agencies that were willing to participate both in services addressing violence exposure and in an on-going research study. Clearly, this limits the potential generalizability of these data to families of children exposed to violence in community settings overall but a limitation of this sort is inherent in community-based experimental research.

Block randomization was used to equalize the intervention and control groups on child age within the program's own specific age range, stratified into four possible groups: birth-2, 3-6, 7-12, and 13-17. Sites were provided with a set of color-coded envelopes that corresponded to each age strata applicable to their program. (The maximum age in the present study's sample is 13 years). After the baseline research assessment was completed, site research staff would select and open the next sealed envelope in the appropriate age strata sequence. The child was assigned the unique study identification number on the outside of the envelope and its contents revealed the family's assignment to either the intervention or control group. The randomization process was monitored closely by the independent evaluation team to insure compliance with the designed procedures. While there were a few minor problems identified with randomization procedures early on, compliance with enrollment and randomization procedures was very high across all nine sites. Specifically, 16 total mishandled cases were identified, coming from just four of the nine sites. These cases were dropped from the study. Since they represented such a small number of cases overall (0.017%) there is little risk of impact on the overall study results.

The most common randomization error was failure to select the appropriate sequential envelope in each age strata. This was followed by a few incidents where a family was randomized to a group prior to completion of the baseline assessment, in violation of the study protocol that interviewers be blind to treatment condition at baseline.

Study Sample

Across the nine programs, 920 families were appropriately enrolled in the study. Of those, 463 (50.3%) were randomized to the intervention group and 457 (49.7%) to the control group. Overall, biological or adoptive mothers represented the vast majority of primary caregivers (85%). The mean child age was 5 (S.D. = 2.7), with a range of infants to age 13. The majority of households (71%) had income at or below the federal poverty line. As Table 2 shows, the randomization procedure successfully equalized the groups on key demographic characteristics (compared using chi-square and t-tests), with the exception of a significant difference on caregiver education. In this case, caregivers in the intervention group reported lower levels of education than those in the control group. The success of the randomization at the aggregate level mirrored the overall success of the randomization within individual sites. In seven of the nine sites, there were no significant differences on any measured demographic characteristics or violence exposure characteristics.

-- Insert Table 2 About Here--

At six months, a second assessment battery was completed with the families. Unlike the baseline assessment, interviewers were not blind to a family's assigned treatment condition at six months. Given that interviewers were not rating or judging respondents, but rather assisting them with self-report measures, and extensive training protocols and supervision in place to minimize

potential bias, we believe any bias introduced by non-blind interviewers at follow up would be small.

Of the 920 families enrolled at baseline, 529 were retained in the study at the six month assessment point, 286 (54%) in the intervention and 243 (46%) in the control. This represents 58 percent retention, with somewhat greater retention in the intervention group. Due to the focus of this paper, the sample was further restricted to only those families with the same primary caregiver participating in the baseline and six month research assessment. This subgroup represents 96 percent of all retained six-month families. The final sample totals 508, with 276 in the intervention group and 232 in the control group. Of these, 169 families (93 in the intervention group and 76 in the control) were retained in the study at the 12 month follow-up assessment point. The reduction in the 12 month sample size is not entirely due to attrition. One of the nine sites utilized a 6-month waitlist design and therefore the data from this site were not included in the 12 months outcome analyses. Adjusting for the 94 cases this site contributed at 6 months, overall 12 month retention was 41 percent. (See Cross et al., in press, for a more detailed analysis of study retention).

Data Collection and Measures

Data was collected in-person by trained interviewers who administered an assessment battery to primary caregivers. Measures of caregiver-reported demographic characteristics, child exposure to violence, caregiver victimization, caregiver everyday stressors, child post-traumatic stress disorder (PTSD) symptoms, child behavior problems, child social-emotional competence, and caregiver-child relationship. Caregiver and child demographic information was collected using a modification of the instrument employed in the Longitudinal Studies of Child Abuse and Neglect (LONGSCAN study; LONGSCAN, 2010), a consortium of longitudinal research studies

assessing the etiology and impact of child maltreatment. The questions covered a host of demographic characteristics (sex, race/ethnicity, marital status, caregiver and household income, etc.) and also caregiver and child physical health, emotional problems, and support or assistance received. Caregivers were asked race and Hispanic ethnicity questions separately. Because Hispanic caregivers very frequently declined to also select a race category, we collapsed race/ethnicity into a variable that indicated whether caregivers selected white race only, black race only, Hispanic ethnicity only, or other (indicating multiple race/ethnicity endorsements).

Child exposure to violence was measured via the caregiver report version of the Juvenile Victimization Questionnaire (JVQ; Hamby et al., 2004a, 2004b). The questionnaire includes several domains: conventional crime, child maltreatment, peer and sibling victimization, witnessing and indirect violence, and sexual assault. The National SSPA evaluation shortened the original JVQ instrument from 34 items to 17 items, which asked about 17 forms of violence exposure within the categories of child physical assault, child maltreatment, witnessing and indirect violence, and sexual abuse. For every form of violence a caregiver endorsed, they were asked how many times this had happened to the child. At baseline, caregivers were asked to report over the child's entire lifetime, and about the prior six months for each follow-up assessment.

While the JVQ asks about caregiver victimization witnessed by the child, caregivers were also asked separately about their own domestic violence victimization based on an item modified from the National Victimization Crime Survey (NCVS) and two general questions about whether in the past year the caregiver had been threatened or attacked by a present or former intimate partner. The baseline time period covered one year and each follow-up assessment asked the caregiver about the prior 6 months.

The caregiver-child relationship was measured via caregiver reports of parenting stress, using the 36-item Parenting Stress Index—Short Form (PSI-SF; Reitman, Currier, and Stickle, 2002). The PSI-SF measures total parenting stress and contains three scales, each with 12 items: parental distress, dysfunctional parent-child interaction, and difficult child characteristics. Caregivers indicated their level of agreement (on a five point scale) with statements about themselves or feelings about/interactions with their child (e.g., I often have the feeling that I cannot handle things very well, my child rarely does things for me that make me feel good). Like prior research, the SSPA national evaluation found the scale to show good internal consistency, with Cronbach's alphas of 0.87 for the parental distress scale, 0.88 for the parent-child dysfunction scale, 0.89 for the difficult child scale, and 0.94 for the total stress scale (Jaycox et al., 2011). For both the computed total score and each subscale, higher scores indicate more stress.

Caregiver everyday stressors, or problems faced in everyday life, were measured via the 20-item Everyday Stressors Index (ESI) used in the LONGSCAN study. In prior research, the ESI construct validity was supported by discrimination of everyday stressors from measures of maternal depression and psychosomatic symptoms using factor analytic procedures (Hall, 1983). Yet, ESI scores also were found to be positively and significantly associated with these symptoms (Hall and Farel, 1988). The National SSPA evaluation used factor analysis of ESI baseline data from 1,517 baseline respondents and identified a two-factor solution that was easily interpretable and had high internal consistency. One represented a 7-item resource problems score (Cronbach's alpha = 0.81) which included items tapping issues related to poverty, such as owing money or getting credit, not having enough money for basic necessities, problems with housing, and employment concerns. The second represented a 13-item personal/family problems

score (Cronbach's alpha = 0.80) including items having to do with concerns about health and about children, interpersonal conflicts, and having too many responsibilities. On both measures, higher scores indicate more problems.

Caregiver reports of child PTSD symptoms were captured using the Trauma Symptom Checklist for Young Children PTSD subscale (TSCYC; Briere et al., 2001). It was administered for children ages 3 to 10 and consisted of 27 items asking caregivers to rate the frequency in the last month of things the child does, feels, or experiences (e.g., bad dreams or nightmares, being bothered by memories of something that happened to him or her) on a 4-point scale with higher scale scores indicating more PTSD symptoms. Discriminant, predictive, and construct validity have been demonstrated for the TSCYC in multiple samples and studies (Briere et al., 2001; Pollio, Glover-Orr, and Wherry, 2008). In the SSPA national evaluation, the Cronbach's alpha for this scale was 0.93 (Jaycox et al., 2011).

Child problem behaviors were captured in several ways. For children age three and older, measures included the internalizing/depression and externalizing behavior problems scales from the Behavior Problems Index (BPI; Peterson & Zill, 1986). Caregivers were asked about their agreement with a series of statements about the child's behavior in the past month (e.g. has been too fearful or anxious, has argued too much). A combined "total behavior problems" measure was constructed to develop a calibrated measure that could be used for all children one age and older. This was done by combining total BPI scores with a second problem behavior measure for younger children using an item response theory (IRT) factor analysis procedure (see Jaycox et al., 2011). For children age one to three, a second measure, the Brief Infant-Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan and Carter, 2002), was also employed. It contains 31 items that ask caregivers to rate behavioral problems (e.g., seems nervous, tense, or

fearful; is restless and can't sit still) on a three-point scale (1 = not true or rarely, 2 = somewhat true or sometimes, and 3 = very true or often). In previous research, the BITSEA scores have correlated highly with concurrent evaluator problem ratings and predicted problem scores one year later (Briggs-Gowan et al., 2004).

Three measures were available to tap social-emotional competence in the sample's age range. The Social Skills Rating System (SSRS; Gresham and Elliott, 1990) was used to assess cooperation (ten items), assertion (ten items), and self-control (ten items). Caregivers of children ages 3 to 12 were asked to rate the frequency (2 = very often, 1 = sometimes, and 0 = never) of a series of behaviors (e.g., How often does your child use free time at home in an acceptable way? How often does your child help you with household tasks without being asked? How often does your child avoid situations that are likely to result in trouble?). Some items varied depending on child age. The SSRS was found to be reliable in prior studies, with Cronbach's alphas of 0.90 for its social skills scale and 0.84 for problem behaviors (Gresham and Elliot, 1990). Other studies have examined convergent validity and found moderate to high correlations between the SSRS and other social competence measures (Merrell and Popinga, 1994; Albertus et al., 1996; Flanagan et al., 1996). For the present study, we utilized the SSRS cooperation measure for a child age 3 and older. For assertion and self-control, we utilized the SSPA national evaluation's age calibrated measures that combined these two SSRS scales with the BITSEA social-emotional scale items. Derived using IRT factor analysis procedures, the resulting scores for children ages 1 and older have a mean value of 0 and a standard deviation of 1, with a higher score indicating more assertion or more self-control. Table 3 presents a list of the measures by child age and programs with children in those age ranges.

-- Insert Table 3 About Here--

In addition to the caregiver report data, SSPA program staff separately documented the type and amount of SSPA services provided to intervention group families and submitted those data at six intervals until services and/or study participation ended. Thus, we are able to take service dose into account in the analyses. The interventions did not define completion according to specific number of sessions and most did not systemically record the reason that services ended. Therefore, a "completion" measure could not be constructed from the data. Alternatively, we measured service dose by summing the number of sessions/services of all types, yielding a family total service count. Within each site, we divided the total service count into quartiles and created three variables which indicated: (1) intervention group families receiving a "weak" dose of SSPA services (as defined by a maximum of 25th percentile or fewer within site), (2) low service dose families receiving services falling between the 25th and 50th percentile, and (3) medium to high dose families in the 50th percentile or above for services received.

These categorical variables were used in conducting "as treated" analyses that capture service dose while standardizing the intensity of SSPA dose delivered across sites.

There were 14 intervention group families (0.05%) who received no reported services. In "intent-to-treat" analyses, these families were analyzed as assigned. In as-treated analyses, however, these families were pooled with control group families.

Since the quartiles were created according to the number of services received, the percent of intervention group families that fell into each group followed a similar distribution: 24 percent received a "weak" dose, 25 percent received a low dose, and 51 percent received a medium to high service dose of the intervention service.

According to the SSPA process evaluation results, nearly all interventions sought to tailor the amount of mental health and case management services to the needs of each family (Schultz et al., 2010). While we are unable to estimate the approximate share of families who received what was intended, it seems reasonable to expect that those in the medium/high service group included families likely to have received the most of what was intended. By the same token, we would expect this group to contain those with the highest level of need and thus representing the group for whom the interventions should have the greatest impact. Conversely, process evaluation observations suggest that those in the first quartile (receiving just one or only a few sessions) received much less of the intervention dose than any of the program models would recommend, rather than representing a dose adjusted for family need (Schultz et al., 2010). Thus, we would expect this weak dose group to have outcomes similar to the control group. Families in the low dose category likely fell short of the service dose that would have been recommend by the respective program model and/or have less pronounced need for continued services. This low dose group allows us to observe whether even limited services may make some difference relative to control families.

Statistical Analyses

In estimating the impact of SSPA participation, the primary comparison is between the pooled intervention group and control group at six months and twelve months post-intervention. Seven caregiver-reported outcomes and nine child-reported outcomes were examined, with each tested separately in multilevel regression models (Hox, 2002; Snijders & Bosker, 1999). These take into account the nested structure of participants clustered in sites where variation in interventions' impact can be estimated.

Since the sites delivered varied interventions, there is a need to estimate the average SSPA intervention effect as well as the site to site variability around the average effect. Our analytic method assesses such variation through a hierarchical regression model with site random intercept and a random intervention impact. For testing scaled outcomes, we employed multivariate hierarchical linear regression to model intervention effects and dichotomous outcomes were tested via hierarchical logistic regression, with both modeling methods incorporating clustering of participants within sites.

The models were estimated using WinBUGS 1.4.1 (Windows Bayesian Inference Using Gibbs Sampling; Lunn, Thomas, Best, & Spiegelhalter, 2000). We used Bayesian Markov Chain Monte Carlo (MCMC) modeling (e.g., Gelman & Hill, 2007) because of convergence limitations that arise in currently available maximum likelihood (ML) approach software packages when estimating hierarchical regression models where small random variations are being estimated. Flexibilities in Bayesian methods can avoid such limitations. The MCMC method generates a large number of simulated random draws from conditional distributions of all the model parameters. Parameter estimations are continuously updated by drawing values from the respective distributions assuming that the current estimated values for the other parameters are true. The basic principle in Bayesian estimation is that once these repeated updates have run long enough, they will approach the desired posterior distribution (Gill, 2002). It is then possible to calculate the posterior mean of this distribution as the best point estimate for each parameter. Bayesian credible intervals (BCI) are reported with the point estimate. They represent the posterior probability interval in which an estimated parameter lies with a specified probability. In analogy to confidence intervals in classical ML statistics, the BCI is based on the 2.5th and 97.5th percentile points of the posterior distribution. That is, the true value of the estimated

parameter lies within this interval with a probability of 0.95 (Gelman & Hill, 2007; Spiegelhalter, Best, Carlin, & van der Linde, 2002). BCIs not including zero as possible values can be interpreted as estimated parameter values that are reliably different from zero.

As previously discussed, the intent-to-treat intervention and control groups were very similar at baseline due to the randomization strategy, with the single exception of caregiver education. Therefore, all models controlled for the child's age categories used for the block randomized design and the caregiver education level using three variables indicating less than high school education (=1), high school diploma or equivalent (=1), and some college or more (=1), respectively. Less than high school education served as the excluded category. Since the sample was balanced on all other measured baseline characteristics, no other control variables were included. In addition, we controlled for the baseline score of the outcome measure in each model in order to account for unmeasured time-invariant characteristics affecting that outcome. Baseline outcomes are expected to explain much of the variability in each of the six and 12 months outcome models, but are included because they provide a more accurate estimate of change than a simple change score model can produce (Markus, 1980).

For the six month time point only, intervention effect was assessed two ways. First, we conducted an intent-to-treat analysis, assessing intervention impact through the use of a single variable indicating whether the family was randomized to the SSPA intervention (=1) or control (=0) group. This independent variable allowed for the examination of overall intent-to-treat effect. Second, we conducted an additional "as-treated" test which replaced the dichotomous intervention variable with the set of four variables capturing the "dose" of SSPA services that each intervention family received, with no services serving as the excluded category. For the

twelve month assessment point, an inadequate number of families was retained to allow for meaningful tests of the impact of service dose.

While the randomization strategy proved highly successful in equalizing the groups on measured characteristics at baseline (see Table 2), the self-selection of the level of treatment dosage received can introduce selection bias. To control for the potential effects of such self-selection in the "as-treated" analyses, we added covariates in the models including gender, race, citizenship, health status, number of children in the household, caregiver employment and caregiver marital status.

Results

The goal of these analyses is to test whether intervention participation, on average, is associated with improved outcomes for children and caregivers. Table 4 shows the simple bivariate relationships between the outcome measures and intervention condition at baseline, then again at the six month and twelve month follow-up assessment points. The group means and distribution of cases across the indicator variable were very similar in practical terms and chi-square and t-tests (as appropriate) identified no significant differences at six or twelve months across the outcomes measures.

-- Insert Table 4 about here ---

Caregiver Outcomes

First, we discuss multivariate analyses of caregiver outcomes at both the six and twelve-month time points. Table 5 presents the results of the six month intent-to-treat and service dose models and the twelve month intent-to-treat models. Since only the intervention-related effects and site to site variation are of interest, only the coefficients related to these variables and variance components are presented in this and subsequent tables. (In the Appendix, we provide

an additional table which adds model coefficient standard errors and effect sizes to the results displayed in Table 5). The site to site variability in intervention (or in dosage group) effect is reported as the proportion of variance associated with the intervention (or dosage) when compared to the total variance. For dichotomous outcomes, the person-level variance was assumed to be $\pi^2/3$ as suggested by Snijders and Bosker (1999) and Larsen and Merlo (2005) in the estimation of such proportions. Each row represents the results of a four separate outcome models, summarizing the results for each outcome measure at the two time points and across service dose levels.

As shown in Table 5, the results for the caregiver intent-to-treat models reveal no significant (i.e. reliable) differences between the intervention and control groups in any of the seven caregiver outcomes. This was true of both the six- and twelve-month models. Site to site variability in intervention effect was also small for all continuous outcomes. It ranged from 0.38% to 0.67% of total variance for six months models and 1.59% to 6.28% of total variance for twelve months models. For the dichotomous outcome of the likelihood of at least one repeat domestic violence incident, the variability is 3.73% and 35.53% for six and twelve month models respectively. While this upper range appears large, this is probably driven by the person-level variance assumption discussed above.

To assess potential sample bias due to attrition by 12 months, we repeated the six-month caregiver outcome models with the subset of 169 families who were retained in the study by twelve months. No differences were observed on six month caregiver outcomes for families who were retained by the 12-month follow-up period. In other words, there was no evidence of any differential intervention effects on shorter term caregiver outcomes among those who were retained over a longer period.

Table 5 also shows the results of the six-month service dose analyses. The comparison in these models is whether an impact can be detected for any of the three SSPA service dosage levels relative to those who received no services. There were no significant differences on any of the outcomes for any service level. In this case, site to site variability in dosage effect was also small for all continuous outcomes, ranging from 0.45% to 2.97% of the total variance. For the dichotomous repeat domestic violence outcome, the variability ranges from 1.59% to 14.87% across dose groups. In all models, all the proportion of variance estimates had BCIs with a lower bound of almost 0, suggesting that there is no reliable evidence of significant site to site effect variability.

-- Insert Table 5 about here –

Child Outcomes

Like the caregiver outcomes, we estimated separate hierarchical random effects models for each of the nine child outcomes, controlling for baseline outcome levels, child age category blocks, and caregiver education. Additional child covariates were added in the six-month as-treated analysis. The results of all four models for each child outcome are displayed in Table 6 (see Appendix Table A2 for model coefficient standard errors and effect size results). Consistent with the caregiver results, in the intent-to-treat analyses, no significant differences were detected between the two groups at either the six or twelve month time points. In other words, there was no evidence of an overall intervention effect either in the short or longer term. Even the site to site variability in intervention effects was small, with BCIs suggesting no reliable evidence of significance. In the six-month service dose models, however, two significant differences in average intervention effect emerged. These measures represent two of the three measures of child social-emotional competence. Specifically, caregivers in families receiving a medium/high

dose of the SSPA services reported improved child cooperation (for those ages 3 and older) relative to those who received no services. (Recall child cooperation is a scaled measure containing items largely relating to caregiver reports of child cooperativeness on household tasks.) This improvement is in the direction expected by the SSPA interventions.

The other significant difference in the six-month service dose analyses is an increase in assertion for children in families receiving a medium/high service dose relative to those receiving no services. The improvement in this measure of social-emotional competence is also in the direction expected by the SSPA interventions.

In all these models, all the proportion of variance estimates again have BCIs with a lower bound almost 0, suggesting that there is no reliable evidence of significant site to site effect variability, even in the two cases of significant intervention effect.

-- Insert Table 6 About Here ---

Discussion

The SSPA initiative, and its national evaluation, was an ambitious attempt to evaluate promising and evidence-based programs under real-world conditions. The effort described here implemented randomized control trials and experienced the many challenges that come with conducting such studies in community settings. These include difficulty enrolling and retaining families, challenges around creating services for the control group families that were both ethical and feasible, and allowing flexible and sometimes shortened administration of the planned interventions. These issues resulted in underpowered studies at the site level (Jaycox et al., 2011) but still allow for examination of pooled data. Thus, akin to a meta-analysis, the goal of the present study was to test the average impact of the SSPA intervention approach on outcomes for children exposed to violence and their families. The working hypotheses in many funding

agencies and community-based organizations is that offering case management and some type of promising mental health intervention will make a difference for children exposed to violence.

The present study tested whether an average intervention effect could be observed on repeat caregiver domestic violence victimization, repeat child exposure to violence, caregiver resource or personal problems, multiple measures of parenting stress, child behavior problems, child PTSD symptoms, or child social-emotional competence. Using hierarchical regression models with site random effects within a Bayesian framework, we found no overall intervention effect on nine child outcomes or on seven caregiver outcomes in the intent-to-treat analyses at both six and 12 months.

In the six-month as-treated analyses, there were some promising findings on two of the three child social-emotional measures, but limited to only to those intervention group families receiving a medium to high dose of intervention services when compared to the control group that did not receive the intervention. Specifically, we found that caregivers in this group reported the participating child to be more cooperative around the house and had higher levels of assertion after six months of SSPA participation, relative to those that received no services. In other words, the results indicate there may be some modest impact on child socio-emotional competence but it may also be limited and observable only among those who receive the most services.

One question raised by the analyses is why no overall intervention effect could be observed on any of the multiple outcomes tested using pooled data from nine carefully executed randomized controlled trials. Indeed, for some outcome measures, it seems reasonable to expect at least minor improvement by six months. For example, given that all programs delivered case management (along with family advocacy for some) it seems reasonable to expect that six

months would be enough for caregivers in the intervention group to begin reporting fewer resource problems (including financially-related worries). Despite the considerable poverty of families within the sample, we did not observe this, even for the families who received the higher levels of intervention services.

One possible explanation is that many of the SSPA programs provided to control group families some form of enhanced service over what might be otherwise available outside the study. In particular, referrals and case management support of varying intensity was provided to many control group families. Alternative therapy services were typically not available, with the exception of the program operating within a child welfare setting. Thus, it is possible that the intervention and control groups differed primarily in the receipt of the specific therapy portion of the overall SSPA interventions, and to a lesser extent in their receipt of other elements of the planned SSPA services such as case management and general crisis support.

Indeed, there is some evidence in the available data that these control group services may have "washed out" measureable impacts among the intervention group. Control group caregivers who were retained at six months reported higher mean resource problems than intervention group caregivers at baseline (15.06 and 14.67, respectively). By the six month follow-up, reported resource problems had declined for both groups ($m = 13.83$ for the control group versus $m=13.54$ for the intervention). The six month decline, however, was only significant (using a paired-samples t-test) for control group caregivers ($t = 3.89$, $p<.01$). Overall group means show that both groups improved on every outcome measure by the six month mark (Table 4).

The possible impact of control group services provides an illustration of the inherent challenges of identifying evidenced-based programming for vulnerable and disadvantaged populations in general and children exposed to violence in particular. Once their troubled life

circumstances have been documented by an enrollment process or research assessment, withholding available assistance to struggling families with young children is not possible or ethical. Ethical and Institutional Review Board concerns often compel service providers and researchers alike to assist control group members at least to some degree. Yet these efforts, largely untested themselves, may contribute to improvements in the control group. Studies powered to detect small overall effects may still be underpowered to detect the difference between good "usual" care and better care offered by the intervention under study.

In addition potential impacts of control group services, the present study contains other limitation which should be considered in the interpretation of its findings. Among these is a reliance on primary caregiver report for all measures except service dose. While this data source has considerable advantages in studies of families and young children, it also can be biased in a variety of ways including problems with recall, denial, and incomplete knowledge of child experiences (Acosta et al., 2012). While caregiver report bias would impact both groups, it is possible that intervention participation may have produced a systematic bias through caregiver sensitization toward the issues under study. It is also possible that the programs examined here may have improved the lives of children and families in ways that were not measured or not measured adequately. Resource constraints also limited the data collection so it was not possible to systematically document or rigorously monitor intervention integrity overall or as-delivered for individual families. While a process evaluation of the SSPA interventions was conducted (Schultz et al., 2010), the multi-site nature of the initiative meant that implementation was documented in a broad fashion and only at limited intervals. It could be that models were delivered in an inconsistent fashion over time, lessening their potential impact.

Another possibility is that data collection may not have been long enough to observe full intervention effects. Many SSPA programs continued to provide services to some families beyond the six month research assessment. Some intervention providers expected that families participating in services may "get worse before they get better," as part of the normal course of progress in mental health service components (Schultz et al., 2010). Thus, failure to detect a short term average intervention effect does not negate the potential for longer term outcomes. Moreover, though we did not detect substantial site-to-site variation in the analyses, these results do not provide insights about the effectiveness of individual SSPA interventions over the short or longer term.

The ideal future study of this type would include many of the features here, including the consistency in measures and data collection, rigorous randomization procedures that resulted in nearly equivalent groups, and ethical and responsible consideration of family needs in study protocols. Improvements could be made with larger sample sizes, more extensive strategies for enhancing recruitment and retention for longer term data collection, and systematic documentation of services to control group families. With the typical funding levels available to support rigorous intervention research, however, what is ideal may not be practically feasible.

Weisburd and Taxman (2000) recommend multicenter randomized trials as a partial solution by targeting funding resources into the multisite implementation of a single program model. This may lead to more economies of scale in process evaluation work and some aspects of outcome evaluation activities, relative to the centralized evaluation of a multi-site, multi-program initiative of this sort. Large federal funding agencies however may find it very difficult to invest significant funds over many years to test a single program model across multiple sites, unless and until there is a particularly appealing intervention that may be worth taking such a

substantial financial and political risk. While there are certainly some examples of such endeavors (e.g. the HIDTA model of treatment for drug-involved offenders described by Weisburd and Taxman), they are few and far between and none exist in areas as new to rigorous evaluation as interventions for children exposed to violence.

In the meantime, studies that combine data in other ways, like the present, can at least contribute some much needed knowledge upon which to build. While no overall intervention effect could be detected for these evidence based therapies offered in conjunction with case management for families with children exposed to violence. There was, however, a modest impact on short-term child social-emotional outcomes for families who receive at least a moderately intense service dose. At the same time, families in the control groups improved as well, perhaps due in part to their receipt of crisis and case management type services. Therefore, it may be advantageous to explore the potential impacts of crisis and case management separately from mental health interventions for children exposed to violence. It may be that, on average, children in families whose basic needs are being attended to improve substantially on their own.

Acknowledgements

This work was made possible through a grant (2005-JW-BX-0001) from the U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention. The opinions expressed here are those of the authors alone. We also appreciate the helpful comments provided by David Weisburd and the anonymous journal reviewers.

References

Acosta, J., Barnes, D., Harris, R., Francois, T., Hickman, L. J., Jaycox, L., & Schultz, D. (2012). An Examination of Measures Related to Children's Exposure to Violence for Use by Both Practitioners and Researchers. *Trauma, Violence, & Abuse, October 2012*, 13 (4): 187-197.

Albertus, C., Birkinbine, J., Lyon, M. A., & Naibi, J. (1996). A validity study of the sSocial Skills Rating System-Teacher Version with Disabled and Nondisabled Preschool Children. *Perceptual and Motor Skills, 83*, 307-316.

Andrews, D.A., Zingler, I., Hoge, R.D., Bonta, J., Gendreau, P., & Cullen, F.T. (1990). Does correctional treatment work? A clinically-relevant and psychologically-informed meta-analysis. *Criminology, 28:* 369-404.

Bickman, L., Sumerfelt, W. T., & Noser, K. (1997). Comparative outcomes of emotionally disturbed children and adolescents in a system of services and usual care. *Psychiatric Services, 48,* 1543–1548.

Bourassa, C. 2007. Co-occurrence of interparental violence and child physical abuse and its effect on the adolescents' behavior. *Journal of Family Violence*, 22(8): 691-701

Briere, J., Johnson, K., Bissada, A., Damon, L., Crouch, J., Gil, E., et al. (2001). The Trauma Symptom Checklist for Young Children (TSCYC): Reliability and Association with Abuse Exposure in a Multi-Site Study. *Child Abuse and Neglect, 25*, 1001-1014.

Briggs-Gowan, M. J., & Carter, A. S. (2002). *Brief Infant-Toddler Social and Emotional Assessment (BITSEA) manual, version 2.0*. New Haven, CT: Yale University.

Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: Screening for Social-Emotional Problems and Delays in Competence *Journal of Pediatric Psychology, 29*, 143-155.

Chaplin, D. (2003). *Hierarchical linear models: Strengths and weaknesses*. Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, Washington, DC.

Cohen, J. (1998). *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

Cross, A., Jaycox, L.H., Hickman, L.J., Schultz, D., Barnes-Proby, D., Kofner, A., & Setodji, C. M. (in press). Predictors of study retention from a multi-site study of interventions for children and families exposed to violence. *Journal of Community Psychology, in press.*

Finkelhor, D., Ormrod, R.K., & Turner, H.A. (2009). Lifetime Assessment of Polyvictimization in a National Sample of Children and Youth. *Child Abuse & Neglect, 33*, 403-411.

Flanagan, D.P., Alfonso, V.C., Primavera, L., Povall, L., & Higgins, D. (1996). Convergent validity of the BASC and SSRS: Implications for social skills assessment. *Psychology in the Schools, 33,* 13-23.

Gresham, F. M., & Elliott, S. (1990). *Social Skills Rating System Manual*: American Guidance Service, Inc.

Gelman, A., & Hill, J. (2007). Data analysis using regression and multilevel/ hierarchical models. New York, NY: Cambridge University Press.

Gill, J. (2002). Bayesian methods: A social and behavioral sciences approach. Boca Raton, FL: Chapman & Hall/CRC

Gopalan, G., Goldstein, L., Klingenstein, K., Sicher, C., Blake, C., & McKay, M. M. (2010). Engaging families into child mental health treatment: Updates and special considerations. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 19*(3), 182-196.

Hall, L. (1983). Social Supports, Everyday Stressors, and Maternal Mental Health. Unpublished doctoral dissertation. University of North Carolina at Chapel Hill.

Hall, L. & Farel, A. M. (1988). Maternal Stresses and Depressive Symptoms: Correlates of Behavior Problems in Young Children. *Nursing Research, 37*, 156-161

Hamby, S. L., Finkelhor, D., Ormrod, R. K., & Turner, H. A. (2004a). *The Juvenile Victimization Questionnaire (JVQ): Caregiver Version*. Durham, NH: Crimes against Children Research Center.

Hamby, S. L., Finkelhor, D., Ormrod, R. K., & Turner, H. A. (2004b). *The Juvenile Victimization Questionnaire (JVQ): Child Self-Report Version*. Durham, NH: Crimes against Children Research Center.

Herrenkohl, T. I., Sousa, C., Tajima, E. A., Herrenkohl, R. C., & Moylan, C. A. (2008). Intersection of child abuse and children's exposure to domestic violence. *Trauma, Violence & Abuse*, 9(2), 84–99.

Hickman, L. J., Jaycox, L. H., Setodji, C. M., Kofner, A., Schultz, D., Barnes-Proby, D. & Harris, R. (in press). How much does "how much" matter? Assessing the relationship between children's lifetime exposure to violence and trauma symptoms, behavior problems, and parenting stress. *Journal of Interpersonal Violence, in press.*

Hox, J. (2002). Multilevel analysis: Techniques and applications. Mahwah, NJ: Erlbaum.

Hyde, M.M., Kracke, K., Jaycox, L.H., & Schultz, D. (2008). The Safe Start Initiative: Advancing system and practice responses to children exposed to violence. *Protecting Children*, 22 (3/4), 80-95.

Institute for Family Centered Services (IFCS), Inc. (2004). *Funding Proposal to the Office of Juvenile Justice and Delinquency Prevention*. CFDA Title: Safe Start Promising Approaches for Children Exposed to Violence. Denver, North Carolina.

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

Larsen K. and Merlo J. (2005). Appropriate assessment of neighborhood effects on individual health—integrating random and fixed effects in multilevel logistic regression. *American Journal of Epidemiology, 16,* 181–188.

Lieberman, A.F. & Van Horn, P. (2005). *Don't Hit My Mommy: A Manual for Child-Parent Psychotherapy with Young Witness of Family Violence*. Washington, DC: Zero to Three.

Lieberman, A.F. & Van Horn, P. (2008). *Psychotherapy with Infants and Young Children: Repairing the Effects of Stress and Trauma on Early Attachment.* New York: Guilford Press.

Lipsey, M. (1992). Juvenile delinquency treatment: a meta-analytic inquiry into the variability of effects. In T. Cook, et al. (Eds.). *Meta-analysis For Explanation: A Casebook.* New York, NY: Russell Sacge Foundation.

LONGSCAN Study. (2010). Available at: http://www.iprc.unc.edu/longscan/pages/measelect/Measure%20Table%20(up%20through%20Age%2018%20Interviews).pdf.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS: A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing, 10, 325–337.

Markus, G. (1980). *Models for the Analysis of Panel Data.* Beverly Hills: Sage.

McKay, M. M. & Bannon, W. M. (2004). Engaging families in child mental health services. *Child and Adolescent Psychiatric Clinics of North America, 13*, 905-921.

National Child Traumatic Stress Network, "CPP: Child Parent Psychotherapy," Raleigh, N.C., 2008. As of July 17, 2011: http://www.nctsnet.org/nctsn_assets/pdfs/promising_practices/cpp_general.pdf

Peterson, J. L., & Zill, N. (1986). Marital disruption, parent-child relationships, and behavioral problems in children. *Journal of Marriage and the Family, 48*(295-307).

Pollio, E. S., Glover-Orr, L. E., & Wherry, J. N. (2008). Assessing Posttraumatic Stress Disorder Using the Trauma Symptom Checklist for Young Children. *Journal of Child Sexual Abuse, 17*(1), 89-100.

Reitman, D., Currier, R. O., & Stickle, T. R. (2002). A critical evaluation of the Parenting Stress Index-Short Form (PSI-SF) in a Head Start population. *Journal of Clinical Child and Adolescent Psychology, 31*(3), 384-392.

Schultz, D., Jaycox, L.H., Hickman, L.J., Chandra, A., Barnes-Proby, D., Acosta, J., Beckman, A., Francois, T., & Honess-Morealle, L. (2010). *National Evaluation of Safe Start Promising Approaches: Assessing Program Implementation*, Santa Monica, Calif.: RAND Corporation, TR-750-DOJ.

Snijders, T. A. B., & Bosker, R. J. (1999). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Thousand Oaks, CA: Sage.

Spiegelhalter, D. J., Best, N. G., Carlin, B. R., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, B, 64*, 583–639.

Sternberg, K. J., Baradaran, L. P., Abbot, C. B., Lamb, M. E., & Guterman, E. (2006). Type of violence, age, and gender differences in the effects of family violence on children's behavior problems: a mega-analysis. *Developmental Review*, 26(1), 89–112.

Turner, H.A., Finkelhor, D., Ormrod, R. (2010) Poly-victimization in a national sample of children and youth. *American Journal of Preventive Medicine, 38,* 323-330.

Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice, 24*(2), 190–195.

Williams, R. L. (2000). A note on robust variance estimation for cluster-correlated data. *Biometrics*, 56(2), 645–646.

Weisburd, D. & Taxman, F., S. (2000). Developing a multicenter randomized trial in criminology: The case of HIDTA. *Journal of Quantitative Criminology, 16 (3)*, 315-339.

Wolfe, D. A., Crooks, C. V., Lee, V., McIntyre-Smith, A., & Jaffe, P. (2003). The effects of children's exposure to domestic violence: a meta-analysis and critique. *Clinical Child and Family Psychology Review*, 6(3): 171–187.

Table 1. Summary of SSPA Program Components, Baseline Sample Size, and Target Population

Site	Intervention Components	Target Population
Program 1 N = 126	Clinic-based, Child-Parent Psychotherapy and additional case management	Children ages 0–6 within a medical home setting exposed to all types of violence or abused
Program 2 N = 201	Clinic-based, Child-Parent Psychotherapy and additional case management	Children ages 0–13 exposed to all types of violence referred from a social service agency
Program 3 N = 68	Home-based, Child-Parent Psychotherapy and additional case management	Children ages 0–5 exposed to domestic violence referred from DV shelters
Program 4 N = 55	Home-based, Child-Parent Psychotherapy (caregiver-child therapy integrated with case management)	Children ages 0–7 exposed to all types of violence in kinship family settings referred from a social service agency
Program 5 N = 166	Home-based, Family-Centered Treatment (caregiver-child therapy integrated with case management)	Children ages 0–8 who have been exposed to all types of violence from multiple social service agencies
Program 6 N = 85	Home-based, Project Support (caregiver-child therapy integrated with case management)	Children ages 3–9 exposed to domestic violence and exiting domestic violence shelters with their mothers
Program 7	Home- or clinic-based, Dyadic caregiver/child	Children ages 0–5 exposed to all types of violence or

N = 58	psychotherapy integrated with case management	abused referred from multiple social service agencies
Program 8 N = 104	Clinic-based, Trauma-Focused Cognitive-Behavioral Therapy, domestic violence advocacy, and case coordination	Children ages 3–12 exposed to domestic violence identified within a county child welfare setting
Program 9 N = 57	Home- or clinic-based, Individualized therapy, case coordination, and parent education groups	Children ages 0–12 exposed to all types of violence or abused referred from a social service agency

Table 2 Baseline Sample Characteristics By Group

	Intervention N = 463		Control N = 457	
	***N*[1]**	***Mean (S.D.)***	***N***	***Mean/S.D.***
Caregiver Age	463	32.3 (9.73)	455	32.6 (9.67)
Target Child Age	463	4.98 (2.76)	457	4.93 (2.69)
Children <18 in home	462	2.5 (1.34)	455	2.5 (1.45)
	N	***%***	***N***	***%***
Caregiver Relationship to Child				
Mother	391	84.6	383	84.4
Other Relationship	71	15.4	71	15.6
Target Child Sex				
Female	249	53.8	238	52.2
Male	214	46.2	218	47.8
Child Race/Ethnicity				
Hispanic	130	28.2	141	30.9
White	117	25.4	131	28.7
Black	109	23.6	106	23.2
Other	105	22.8	79	17.3
Caregiver Born in United States				
Yes	291	63.1	299	65.9
No	170	36.9	155	34.1
Caregiver Education*				
Less than high school	171	37.0	130	28.1

High school or GED	129	27.9	138	30.3
Some college or above	162	35.0	188	41.2
Caregiver Employment Status				
Full-Time	134	29.0	123	27.0
Part-Time	47	10.2	56	12.3
Unemployed & Looking	107	23.2	106	23.2
Unemployed, Not looking for work (retired, other)	174	37.7	171	37.5
Household At or below Federal Poverty Line				
Yes	290	72.9	280	69.8
No	108	27.1	121	30.2
Household Receiving Financial Assistance				
Yes	395	85.3	377	82.5
No	68	14.7	80	17.5
Caregiver Health Status				
Poor to Fair	153	33.0	153	33.8
Good to Excellent	310	67.0	300	66.2
Caregiver Living with Spouse/Partner				
Yes	133	29.0	132	29.1
No	326	71.0	322	70.9

[1] This column represents the number of cases for which there are valid data.

* p<.05

Table 3. Caregiver Report Measures by Domain, Child Age and Program

Domain	Measure	Age Range
Demographics and Caregiver Problems	Child and Caregiver Demographics and Service Use	All
	Everyday Stressors Index	All
PTSD Symptoms	Trauma Symptom Checklist for Young Children (PTSD Scale)	3-10
Child Behavior Problems	Total Problems: IRT combination of (1) Brief Infant-Toddler Social and Emotional Assessment (Problem Scale) and (2) Behavior Problem Index	1+
	Externalizing: Behavior Problem Index (Externalizing Scale)	3+
	Depression: Behavior Problem Index (Internalizing Scale)	3+
Social-Emotional Competence	Assertion & Self-Control: IRT combination of Brief Infant-Toddler Social and Emotional Assessment (Social-Emotional Competence Scale) and Skills Rating System Assertion and Self-Control Scales	1+
	Cooperation: Social Skills Rating System Cooperation scale	3+
Caregiver-Child Relationship	Parenting Stress Index –Short Form	All
Violence Exposure & Victimization	Juvenile Victimization Questionnaire	All
	Caregiver Victimization	All

Table 4. Bivariate Relationships Between Group and Outcome Measures at Baseline, Six and Twelve Months

	Baseline		**Six Months**			**Twelve Months**		
	Intervention N = 463	**Control N = 457**	**Intervention N = 276**	**Control N = 232**		**Intervention N = 93**	**Control N = 76**	
Caregiver Outcomes	**Mean (s.d.)**	**Mean (s.d)**	**Mean (s.d.)**	**Mean (s.d)**	**Effect size**	**Mean (s.d.)**	**Mean (s.d)**	**Effect size**
Resource Problems	15.07 (5.67)	15.12 (5.66)	13.54 (5.14)	13.83 (5.29)	-0.0559	13.52 (5.54)	13.25 (6.02)	0.0466
Personal Problems	25.26 (7.17)	25.36 (7.08)	23.16 (6.76)	23.47 (7.07)	-0.0443	23.12 (6.69)	21.85 (5.95)	0.1994
Total Parental Stress	88.88 (23.39)	87.69 (24.40)	82.18 (22.73)	81.93 (23.77)	0.0107	83.59 (23.25)	82.30 (27.23)	0.0512
Child Difficulty	32.37 (9.66)	32.02 (9.84)	30.01 (9.29)	30.38 (9.62)	-0.0390	30.43 (8.90)	29.89 (10.75)	0.0552
Parent-Child Dysfunction	25.00 (8.48)	24.85 (8.75)	23.65 (7.69)	23.30 (8.01)	0.0441	24.37 (8.46)	24.18 (9.43)	0.0216
Parental Distress	31.53 (8.92)	30.84 (9.42)	28.49 (9.19)	28.25 (9.19)	0.0261	28.85 (9.21)	28.23 (10.16)	0.0643
Repeat Violence/Exposure	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**
Caregiver experience DV Yes	58%	62%	15%	18%	-0.0663	12.22%	14.67%	-0.0717

No	42%	39%	85%	82%		87.78%	85.33%	
Child Exposed to Violence								
Yes	100%	100%	49%	47%	0.0400	52.22	45.33	0.1379
No	0	0	51%	53%		47.78	54.67	
Child Outcomes	**Mean (s.d.)**	**Mean (s.d.)**	**Mean (s.d.)**	**Mean (s.d.)**	**Effect Size**	**Mean (s.d.)**	**Mean (s.d.)**	**Effect Size**
Violence Exposure Frequency	14.71 (20.10)	14.77 (20.17)	3.42 (7.16)	3.34 (8.68)	0.0104	5.37 (10.80)	2.67 (5.12)	0.3112
PTSD Symptoms	45.74 (13.04)	45.20 (13.24)	42.99 (12.92)	42.24 (12.11)	0.0590	40.57 (12.15)	37.39 (9.28)	0.2935
Externalizing Behavior	29.46 (8.19)	28.92 (8.16)	27.07 (7.59)	27.21 (7.80)	-0.0187	27.12 (7.50)	26.63 (6.40)	0.0702
Internalizing Behavior	16.59 (4.46)	16.52 (4.22)	15.26 (3.98)	15.25 (4.05)	0.0041	14.53 (3.73)	14.28 (3.19)	0.0725
Behavior Problems	0.25 (0.99)	0.25 (0.96)	-0.03 (0.93)	0.01 (0.99)	-0.0362	-0.03 (0.96)	-0.09 (0.81)	0.0615
Assertion	-0.18 (0.80)	-0.13 (0.88)	-0.06 (0.79)	-0.09 (0.86)	0.0378	0.10 (0.73)	0.01 (0.74)	0.1251
Self-Control	-0.15 (0.89)	-0.11 (0.92)	0.02 (0.88)	-0.01 (0.97)	0.0324	0.22 (0.83)	0.07 (0.87)	0.1872
Cooperation	10.46 (4.32)	10.88 (4.51)	11.57 (3.80)	11.11 (4.23)	0.1131	11.74 (3.83)	11.42 (4.19)	0.0787

* p<.05 Note: Effect size is calculated as the difference between the intervention and control group divided by the standard deviation at each time point.

Table 5. Multivariate Mixed Effects Six- and Twelve-Month Caregiver Outcome Models Controlling for Caregiver Education, and Baseline Outcome

Outcomes Models		Six Month: Intervention Service Dose[1] N = 508			Six Month: Intent-to-Treat N = 508	Twelve Month: Intent-to-Treat N = 169
Caregiver Outcomes		Weak	Low	Medium/High	Intervention Group	Intervention Group
Resource Problems	Point Est	0.22 (-1.17, 1.65)	-0.34 (-1.51, 1.03)	0.04 (-0.82, 0.92)	-0.12 (-0.88, 0.70)	-0.58 (-2.13, 0.87)
	% Variance	2.97 (0.01, 21.13)	2.35 (0.00, 17.87)	0.77 (0.00, 5.34)	0.59 (0.00, 3.93)	1.59 (0.00, 12.64)
Personal Problems	Point Est	0.18 (-1.53, 2.11)	-0.91 (-2.53, 0.91)	0.10 (-1.04, 1.46)	-0.14 (-1.23, 1.04)	0.80 (-1.01, 2.73)
	% Variance	1.45 (0.00, 11.58)	1.23 (0.00, 9.61)	1.37 (0.00, 9.70)	0.65 (0.00, 4.58)	5.37 (0.00, 31.06)
Parental Distress	Point Est	1.39 (-0.81, 3.72)	-1.29 (-3.24, 1.00)	-0.32 (-1.64, 1.10)	-0.24 (-1.62, 1.02)	-1.56 (-3.76, 0.48)
	% Variance	2.63 (0.00, 22.66)	1.62 (0.00, 13.67)	0.61 (0.00, 4.87)	0.66 (0.00, 5.13)	1.89 (0.00, 15.56)
Parent-Child Dysfunction	Point Est	0.65 (-1.22, 2.65)	-0.31 (-1.87, 1.60)	-0.07 (-1.20, 1.02)	-0.05 (-1.13, 1.06)	-0.49 (-2.65, 1.86)
	% Variance	2.64 (0.00, 19.08)	2.08 (0.00, 17.68)	0.63 (0.00, 4.69)	0.67 (0.00, 4.73)	2.10 (0.00, 16.35)
Child Difficulty	Point Est	0.74 (-1.23, 3.05)	-0.32 (-2.24, 1.79)	-0.86 (-2.14, 0.46)	-0.67 (-1.91, 0.59)	-0.04 (-2.54, 2.64)
	% Variance	1.26 (0.00, 10.33)	2.90 (0.00, 21.40)	0.85 (0.00, 6.28)	0.50 (0.00, 3.51)	6.28 (0.00, 33.57)
Total Parenting Stress	Point Est	2.96 (-1.73, 8.26)	-1.75 (-5.79, 2.98)	-1.38 (-4.49, 1.63)	-0.96 (-3.80, 2.00)	-2.28 (-8.36, 3.84)
	% Variance	1.22 (0.00, 10.51)	1.16 (0.00, 10.61)	0.45 (0.00, 3.86)	0.38 (0.00, 3.17)	3.51 (0.00, 24.83)
At least one DV incident	Point Est	-0.42 (-1.97, 0.76)	-0.37 (-1.83, 0.84)	-0.02 (-0.70, 0.60)	-0.05 (-0.67, 0.51)	-0.60 (-3.20, 1.13)
	Odds	0.66 (0.14, 2.13)	0.69 (0.16, 2.32)	0.98 (0.50, 1.82)	0.95 (0.51, 1.66)	0.55 (0.04, 3.10)
	% Variance	11.92 (0.02, 69.97)	14.87 (0.02, 74.16)	1.59 (0.01, 10.88)	3.73 (0.03, 21.44)	35.53 (0.06, 88.21)

Point Est. = Point estimate of treatment effect computed as the posterior mean and its Bayesian credible intervals (BCI).

% Variance = Proportion of variance (and BCI) associated with site to site variability in treatment (or dosage group) effect.

[1] In addition to caregiver education, all service dose models contain the following covariates: child gender, race/ethnicity, caregiver immigrant status, health status, employment, caregiver cohabitation with spouse/partner, and number of children in the household.

Table 6. Multivariate Mixed Effects Six- and Twelve-Month Child Outcome Models Controlling for Caregiver Education, and Baseline Outcome

Outcomes Models		Six Month: Intervention Service Dose[1] N = 508			Six Month: Intent-to-Treat N = 508	Twelve Month: Intent-to-Treat N = 169
Child Outcomes		Weak	Low	Medium/High	Intervention Group	Intervention Group
Any Violence Exposure	Point Est	0.14 (-0.67, 0.96)	0.24 (-0.39, 0.92)	0.26 (-0.29, 0.89)	0.17 (-0.27, 0.63)	0.36 (-0.79, 1.49)
	Odds	1.15 (0.51, 2.62)	1.27 (0.67, 2.50)	1.30 (0.75, 2.44)	1.19 (0.76, 1.88)	1.43 (0.45, 4.42)
	% Variance	2.70 (0.02, 18.97)	3.38 (0.02, 23.23)	4.25 (0.02, 24.92)	1.22 (0.02, 7.64)	18.95 (0.05, 69.67)
Violence Exposure Frequency	Point Est	-1.00 (-3.39, 1.64)	2.29 (-0.52, 5.08)	0.08 (-1.45, 1.57)	0.34 (-1.12, 1.80)	3.01 (-0.40, 6.05)
	% Variance	1.36 (0.00, 10.95)	7.70 (0.00, 37.67)	1.00 (0.00, 7.80)	0.50 (0.00, 3.82)	6.20 (0.00, 31.59)
PTSD Symptoms	Point Est	-2.29 (-6.43, 1.72)	-1.29 (-4.60, 1.66)	-0.56 (-3.05, 1.74)	-0.32 (-2.44, 1.57)	2.75 (-1.03, 6.08)
	% Variance	2.43 (0.00, 21.58)	1.43 (0.00, 11.81)	1.90 (0.00, 14.29)	0.74 (0.00, 5.75)	2.03 (0.00, 18.08)
Externalizing Behavior	Point Est	-1.14 (-3.30, 1.29)	0.02 (-2.80, 3.15)	-0.52 (-2.01, 1.13)	-0.55 (-1.77, 0.61)	0.30 (-1.98, 2.55)
	% Variance	2.81 (0.00, 24.14)	18.00 (0.02, 59.03)	3.84 (0.00, 21.33)	0.88 (0.00, 6.62)	4.22 (0.00, 30.80)
Depression	Point Est	-0.27 (-1.66, 0.98)	-0.34 (-1.43, 0.70)	-0.54 (-1.25, 0.16)	-0.19 (-0.90, 0.50)	-0.10 (-1.27, 1.11)
	% Variance	2.46 (0.01, 19.85)	3.27 (0.01, 22.32)	1.29 (0.01, 9.30)	1.11 (0.01, 7.56)	4.79 (0.01, 28.99)
Total Behavior Problems	Point Est	-0.05 (-0.31, 0.21)	0.02 (-0.28, 0.35)	-0.08 (-0.26, 0.11)	-0.03 (-0.17, 0.12)	0.08 (-0.16, 0.33)
	% Variance	4.54 (0.10, 24.38)	13.05 (0.21, 46.27)	3.89 (0.11, 17.85)	1.45 (0.10, 7.06)	4.14 (0.12, 21.65)
Assertion	Point Est	0.03 (-0.26, 0.33)	0.11 (-0.12, 0.34)	0.22 (0.05, 0.41)*	0.10 (-0.05, 0.24)	0.07 (-0.18, 0.35)
	% Variance	12.62 (0.23, 45.37)	3.17 (0.11, 17.20)	4.27 (0.14, 18.90)	2.56 (0.14, 12.13)	10.31 (0.20, 41.82)
Self-Control	Point Est	0.14 (-0.13, 0.41)	0.16 (-0.10, 0.41)	0.10 (-0.10, 0.31)	0.07 (-0.08, 0.22)	0.10 (-0.15, 0.36)
	% Variance	4.99 (0.12, 26.50)	4.01 (0.11, 21.22)	6.16 (0.18, 23.39)	2.28 (0.12, 10.28)	3.15 (0.11, 16.70)
Cooperation	Point Est	0.51 (-1.09, 2.35)	-0.03 (-1.28, 1.19)	1.02 (0.14, 2.06)*	0.65 (-0.16, 1.54)	-0.01 (-1.78, 1.84)
	% Variance	8.31 (0.01, 51.16)	2.84 (0.01, 22.46)	3.89 (0.01, 22.68)	2.74 (0.01, 16.13)	10.00 (0.01, 53.11)

Point Est. = Point estimate of treatment effect computed as the posterior mean and its Bayesian credible intervals (BCI).

% Variance = Proportion of variance (and BCI) associated with site to site variability in treatment (or dosage group) effect.

* Point estimates reliably different from 0.

[1] In addition to caregiver education, all service dose models contain the following covariates: child gender, race/ethnicity, caregiver immigrant status, health status, employment, caregiver cohabitation with spouse/partner, and number of children in the household.

APPENDIX

Table A1. Multivariate Mixed Effects Six- and Twelve-Month Caregiver Outcome Models Controlling for Caregiver Education, and Baseline Outcome

Outcomes Models		Six Month: Intervention Service Dose[1] N = 508			Six Month: Intent-to-Treat N = 508	Twelve Month: Intent-to-Treat N = 169
Caregiver Outcomes		Weak	Low	Medium/High	Intervention Group	Intervention Group
Resource Problems	Est (SE)	0.22 (0.73)	-0.34 (0.65)	0.04 (0.44)	-0.12 (0.40)	-0.58 (0.73)
	BCI	(-1.17, 1.65)	(-1.51, 1.03)	(-0.82, 0.92)	(-0.88, 0.70)	(-2.13, 0.87)
	Effect size	0.0132	-0.0233	0.0042	-0.0138	-0.0612
	% Variance	2.97 (0.01, 21.13)	2.35 (0.00, 17.87)	0.77 (0.00, 5.34)	0.59 (0.00, 3.93)	1.59 (0.00, 12.64)
Personal Problems	Est (SE)	0.18 (0.97)	-0.91 (0.87)	0.10 (0.64)	-0.14 (0.56)	0.80 (0.96)
	BCI	(-1.53, 2.11)	(-2.53, 0.91)	(-1.04, 1.46)	(-1.23, 1.04)	(-1.01, 2.73)
	Effect size	0.0082	-0.0466	0.0070	-0.0108	0.0651
	% Variance	1.45 (0.00, 11.58)	1.23 (0.00, 9.61)	1.37 (0.00, 9.70)	0.65 (0.00, 4.58)	5.37 (0.00, 31.06)
Parental Distress	Est (SE)	1.39 (1.18)	-1.29 (1.05)	-0.32 (0.71)	-0.24 (0.67)	-1.56 (1.15)
	BCI	(-0.81, 3.72)	(-3.24, 1.00)	(-1.64, 1.10)	(-1.62, 1.02)	(-3.76, 0.48)
	Effect size	0.0528	-0.0553	-0.0204	-0.0158	-0.1068
	% Variance	2.63 (0.00, 22.66)	1.62 (0.00, 13.67)	0.61 (0.00, 4.87)	0.66 (0.00, 5.13)	1.89 (0.00, 15.56)
Parent-Child Dysfunction	Est (SE)	0.65 (1.00)	-0.31 (0.89)	-0.07 (0.59)	-0.05 (0.56)	-0.49 (1.20)
	BCI	(-1.22, 2.65)	(-1.87, 1.60)	(-1.20, 1.02)	(-1.13, 1.06)	(-2.65, 1.86)
	Effect size	0.0293	-0.0157	-0.0054	-0.0044	-0.0321
	% Variance	2.64 (0.00, 19.08)	2.08 (0.00, 17.68)	0.63 (0.00, 4.69)	0.67 (0.00, 4.73)	2.10 (0.00, 16.35)
Child Difficulty	Est (SE)	0.74 (1.10)	-0.32 (1.04)	-0.86 (0.70)	-0.67 (0.63)	-0.04 (1.35)

	BCI	(-1.23, 3.05)	(-2.24, 1.79)	(-2.14, 0.46)	(-1.91, 0.59)	(-2.54, 2.64)
	Effect size	0.0304	-0.0137	-0.0551	-0.0470	-0.0022
	% Variance	1.26 (0.00, 10.33)	2.90 (0.00, 21.40)	0.85 (0.00, 6.28)	0.50 (0.00, 3.51)	6.28 (0.00, 33.57)
Total Parenting Stress	Est (SE)	2.96 (2.66)	-1.75 (2.39)	-1.38 (1.63)	-0.96 (1.54)	-2.28 (3.27)
	BCI	(-1.73, 8.26)	(-5.79, 2.98)	(-4.49, 1.63)	(-3.80, 2.00)	(-8.36, 3.84)
	Effect size	0.0502	-0.0330	-0.0381	-0.0280	-0.0550
	% Variance	1.22 (0.00, 10.51)	1.16 (0.00, 10.61)	0.45 (0.00, 3.86)	0.38 (0.00, 3.17)	3.51 (0.00, 24.83)
At least one DV incident	Est (SE)	-0.42 (0.76)	-0.37 (0.68)	-0.02 (0.34)	-0.05 (0.30)	-0.60 (1.08)
	BCI	(-1.97, 0.76)	(-1.83, 0.84)	(-0.70, 0.60)	(-0.67, 0.51)	(-3.20, 1.13)
	Effect size	-0.0248	-0.0241	-0.0031	-0.0078	-0.0435
	Odds	0.66 (0.14, 2.13)	0.69 (0.16, 2.32)	0.98 (0.50, 1.82)	0.95 (0.51, 1.66)	0.55 (0.04, 3.10)
	% Variance	11.92 (0.02, 69.97)	14.87 (0.02, 74.16)	1.59 (0.01, 10.88)	3.73 (0.03, 21.44)	35.53 (0.06, 88.21)

Est = Point estimate of treatment effect computed as the posterior mean and its standard error (SE).

BCI= Bayesian 95% credible intervals.

% Variance = Proportion of variance (and BCI) associated with site to site variability in treatment (or dosage group) effect.

[1] In addition to caregiver education, all service dose models contain the following covariates: child gender, race/ethnicity, caregiver immigrant status, health status, employment, caregiver cohabitation with spouse/partner, and number of children in the household.

A2. Multivariate Mixed Effects Six- and Twelve-Month Child Outcome Models Controlling for Caregiver Education, and Baseline Outcome

Outcomes Models		Six Month: Intervention Service Dose N = 508			Six Month: Intent-to-Treat N = 508	Twelve Month: Intent-to-Treat N = 169
Child Outcomes		Weak	Low	Medium/High	Intervention Group	Intervention Group
Any Violence Exposure	Est (SE)	0.14 (0.41)	0.24 (0.35)	0.26 (0.29)	0.17 (0.23)	0.36 (0.58)
	BCI	(-0.67, 0.96)	(-0.39, 0.92)	(-0.29, 0.89)	(-0.27, 0.63)	(-0.79, 1.49)
	Effect size	0.0152	0.0305	0.0407	0.0334	0.0475
	Odds	1.15 (0.51, 2.62)	1.27 (0.67, 2.50)	1.30 (0.75, 2.44)	1.19 (0.76, 1.88)	1.43 (0.45, 4.42)
	% Variance	2.70 (0.02, 18.97)	3.38 (0.02, 23.23)	4.25 (0.02, 24.92)	1.22 (0.02, 7.64)	18.95 (0.05, 69.67)
Violence Exposure Frequency	Est (SE)	-1.00 (1.23)	2.29 (1.45)	0.08 (0.80)	0.34 (0.75)	3.01 (1.63)
	BCI	(-3.39, 1.64)	(-0.52, 5.08)	(-1.45, 1.57)	(-1.12, 1.80)	(-0.40, 6.05)
	Effect size	-0.0372	0.0723	0.0046	0.0208	0.1500
	% Variance	1.36 (0.00, 10.95)	7.70 (0.00, 37.67)	1.00 (0.00, 7.80)	0.50 (0.00, 3.82)	6.20 (0.00, 31.59)
PTSD Symptoms	Est (SE)	-2.29 (2.07)	-1.29 (1.61)	-0.56 (1.26)	-0.32 (1.00)	2.75 (1.75)
	BCI	(-6.43, 1.72)	(-4.60, 1.66)	(-3.05, 1.74)	(-2.44, 1.57)	(-1.03, 6.08)
	Effect size	-0.0579	-0.0419	-0.0233	-0.0164	0.1391
	% Variance	2.43 (0.00, 21.58)	1.43 (0.00, 11.81)	1.90 (0.00, 14.29)	0.74 (0.00, 5.75)	2.03 (0.00, 18.08)
Externalizing Behavior	Est (SE)	-1.14 (1.19)	0.02 (1.48)	-0.52 (0.78)	-0.55 (0.61)	0.30 (1.14)
	BCI	(-3.30, 1.29)	(-2.80, 3.15)	(-2.01, 1.13)	(-1.77, 0.61)	(-1.98, 2.55)
	Effect size	-0.0509	0.0009	-0.0356	-0.0474	0.0232
	% Variance	2.81 (0.00, 24.14)	18.00 (0.02, 59.03)	3.84 (0.00, 21.33)	0.88 (0.00, 6.62)	4.22 (0.00, 30.80)
Depression	Est (SE)	-0.27 (0.67)	-0.34 (0.54)	-0.54 (0.36)	-0.19 (0.35)	-0.10 (0.61)
	BCI	(-1.66, 0.98)	(-1.43, 0.70)	(-1.25, 0.16)	(-0.90, 0.50)	(-1.27, 1.11)

	Effect size	-0.0207	-0.0323	-0.0777	-0.0274	-0.0147
	% Variance	2.46 (0.01, 19.85)	3.27 (0.01, 22.32)	1.29 (0.01, 9.30)	1.11 (0.01, 7.56)	4.79 (0.01, 28.99)
Total Behavior Problems	Est (SE)	-0.05 (0.13)	0.02 (0.16)	-0.08 (0.09)	-0.03 (0.07)	0.08 (0.12)
	BCI	(-0.31, 0.21)	(-0.28, 0.35)	(-0.26, 0.11)	(-0.17, 0.12)	(-0.16, 0.33)
	Effect size	-0.0164	0.0064	-0.0396	-0.0164	0.0511
	% Variance	4.54 (0.10, 24.38)	13.05 (0.21, 46.27)	3.89 (0.11, 17.85)	1.45 (0.10, 7.06)	4.14 (0.12, 21.65)
Assertion	Est (SE)	0.03 (0.15)	0.11 (0.12)	0.22 (0.09)*	0.10 (0.07)	0.07 (0.13)
	BCI	(-0.26, 0.33)	(-0.12, 0.34)	(0.05, 0.41)	(-0.05, 0.24)	(-0.18, 0.35)
	Effect size	0.0106	0.0439	0.1110	0.0593	0.0419
	% Variance	12.62 (0.23, 45.37)	3.17 (0.11, 17.20)	4.27 (0.14, 18.90)	2.56 (0.14, 12.13)	10.31 (0.20, 41.82)
Self-Control	Est (SE)	0.14 (0.14)	0.16 (0.13)	0.10 (0.10)	0.07 (0.08)	0.10 (0.13)
	BCI	(-0.13, 0.41)	(-0.10, 0.41)	(-0.10, 0.31)	(-0.08, 0.22)	(-0.15, 0.36)
	Effect size	0.0479	0.0580	0.0435	0.0409	0.0640
	% Variance	4.99 (0.12, 26.50)	4.01 (0.11, 21.22)	6.16 (0.18, 23.39)	2.28 (0.12, 10.28)	3.15 (0.11, 16.70)
Cooperation	Est (SE)	0.51 (0.91)	-0.03 (0.63)	1.02 (0.50)*	0.65 (0.43)	-0.01 (0.90)
	BCI	(-1.09, 2.35)	(-1.28, 1.19)	(0.14, 2.06)	(-0.16, 1.54)	(-1.78, 1.84)
	Effect size	0.0313	-0.0022	0.1140	0.0835	-0.0010
	% Variance	8.31 (0.01, 51.16)	2.84 (0.01, 22.46)	3.89 (0.01, 22.68)	2.74 (0.01, 16.13)	10.00 (0.01, 53.11)

Est = Point estimate of treatment effect computed as the posterior mean and its standard error (SE).

BCI= Bayesian 95% credible intervals.

% Variance = Proportion of variance (and BCI) associated with site to site variability in treatment (or dosage group) effect.

[1] In addition to caregiver education, all service dose models contain the following covariates: child gender, race/ethnicity, caregiver immigrant status, health status, employment, caregiver cohabitation with spouse/partner, and number of children in the household.