

Evaluating the Implementation of a Family-
Focused Prevention Program: Effectiveness of SAFE Children
Executive Summary

David B. Henry

University of Illinois at Chicago

Patrick H. Tolan

University of Virginia

Deborah Gorman-Smith

Michael E. Schoeny

Chapin Hall at the University of Chicago

Jack Zwanziger

Sage Kim

University of Illinois at Chicago

2006-JP-FX-0062

February 2012

Correspondence should be addressed to David Henry, Institute for Health Research and Policy,
University of Illinois at Chicago. 1747 W. Roosevelt Rd., Chicago, IL 60608. email:
Dhenry@uic.edu

Contents		page
A. <i>The Value of Prevention</i>		3
B. Methods		4
C. Modifications to or Problems with the Original Research Design		4
D. Findings		5
	Goal 1: To test the effectiveness of the SAFEChildren intervention program when implemented by community intervention providers and tutors.	5
	Goal 2: To examine the role of aspects of social network processes in the intervention effects.	7
	Goal 3: To examine the relation of implementation qualities to variation in effects	8
	Goal 4: To conduct a cost-benefit analysis of the SAFEChildren intervention.	10
E. Conclusions and Implications		12
References		14

A. The Value of Prevention.

Over the past twenty years a considerable body of knowledge has supported the value of prevention efforts regarding delinquency, violence, drug use, and school failure (Multisite Violence Prevention Project (MVPP), 2004; Tolan & Guerra, 1994; Wasserman & Miller, 1998). Developmental studies have clarified the most important risk factors to target in order to reduce youth violence (US Department of Health and Human Services (USDHHS) Public Health Service (PHS); Office of the Surgeon General, 2001; World Federation of Mental Health, 2002). As a result, a growing list of preventive interventions aimed at children and youth and targeting risk factors identified in basic developmental studies have been shown to be effective. The promise of these approaches as well as empirical findings from our prior studies led to development of the SAFEChildren intervention. Informed by developmental-ecological theory, we developed a family-oriented prevention program for those facing the exceptional challenges of raising children in inner-city communities (Mason, Cauce, Gonzales, & Hiraga, 1996). The intent in its development was to aid parents raising children within inner-city communities to:

1. promote positive adjustment and success during the first grade year,
2. promote strong parenting and engagement with other parents, the school, and other resources,
3. increase protection from risks associated with residence in these communities, and
4. build children's competencies as they grow up.

We reasoned that there might be protective benefits from working to engage parents to support children's academic adjustment and achievement during the first grade year. In addition to supporting parents, we thought it valuable to offer opportunities for children to improve basic reading skills. Inner-city parents often feel isolated and in need of information about normal child development and how to manage developmental and social setting challenges. We reasoned that focusing on these areas would build protection and impede risk trajectories for aggression, violence and school failure.

SAFEChildren is comprised of a reading tutoring program and a family-focused intervention (20 weeks duration) provided during the child's first grade year. The family-focused intervention is based in weekly multiple-family group meetings (about 5 families per group) that address issues of parenting, family relationships, child development, and parental involvement and investment in their child's schooling (McKay, Gonzales, Quintana, Kim, & Abdul-Adil, 1995; Tolan & McKay, 1996; Webster-Stratton, 1990). Multiple family groups are thought to build social support among participants, and improve parent-child interactions (McKay, et al., 1995).

The reading tutoring program emphasizes phonics as well as word recognition. It combines training in these skills with application to reading high-interest, relevant material with positive messages and activities. These materials are designed to increase parental involvement in their child's academic development (Coie & Krehbiel, 1984; Wallach & Wallach, 1976). This program is designed to complement the phonics-based reading program used by the Chicago Public Schools.

The results of an initial test of SAFEChildren (the *SAFE Efficacy Trial*) are reported in detail in Tolan et al. (2004). The intervention resulted in increased levels of academic achievement and parental involvement in school. Intervention participants increased reading skills at a rate approximating national norms and were actually a bit above the national average for reading skills by mid second grade (2 years, 6 months), with an average score grade equivalent (GE) of 2.9. In contrast, control condition students were just below the national average at the same point in time (GE = 2.4). Intervention families maintained levels of parental involvement in their children's schooling over the 2½ years of the study, but control families showed decreasing parental involvement. Intervention children in high-risk families showed decreased aggression over time but high-risk

controls showed essentially no change. Intervention children from high-risk families had positive slopes on Concentration and improved Social Competence while the control children showed no change. We also found that among high-risk families, the crucial skill of Parental Monitoring improved for those in the intervention, while pre-intervention levels were unchanged for high-risk controls. Children with initially high aggression who were assigned to the intervention condition had decreasing levels of Aggression and Hyperactivity compared to high-aggression controls. Aggressive children in the intervention also showed greater improvement in Social Competence and less of a decline in parent involvement than equally aggressive controls. In a longer-term booster intervention and follow-up study (*SAFE Booster Trial*), we were able to recruit and track 382 of the original 424 *SAFE Efficacy Trial* participants. There were several sustained effects on children with high initial aggression, including effects on Parental Monitoring, Parental Use of Effective Discipline Practices, and Parent Involvement in School. These results suggest maintenance of initial effects and emergence of new effects impacting those at greatest risk for later delinquency.

All of this evidence raises the question of the potential effectiveness of SAFEChildren when implemented with community providers of the family intervention and tutors recruited from the student bodies of the participating schools. Evaluating the effectiveness of SAFEChildren under such conditions (the *SAFE Effectiveness Trial*) is the primary focus of this report. For estimating the costs and potential benefits of SAFEChildren, we also report the results of the *SAFE Follow-up Study* that evaluated effects on the original participants after high school.

B. Methods

Our primary target sample was the children (and their parents/caregivers) enrolled in first grade in five elementary schools in disadvantaged neighborhoods in Chicago, IL. In addition, we enrolled teachers as sources of data about the children, tutors, and mental health service providers as participants in the study.

Individual children and their families were randomly assigned to treatment or control conditions. In addition to the first grade participants, we recruited 278 7th and 8th graders who had been nominated by teachers to be trained and participate as tutors for the first-grade children in the treatment condition. Each tutor worked with one or two first graders. Tutors averaged 14 years of age, and over two-thirds were female. Two-thirds reported their ethnicity as “African-American or Black” and 36.3% reported Hispanic or Latino ethnicity. The tutors received 40 hours of training prior to beginning the intervention and received 30 minutes of supervision each week during the course of the intervention.

We also recruited family intervention providers from two community agencies providing mental health services in the neighborhoods in which the schools were located. Fifteen individual therapists conducted groups. Two were male and 13 were female. Nine reported their ethnicity as African-American and 6 reported Latino/Hispanic ethnicity. Eight reported having a master’s degree, and 5 reported their level of education as associate degree or college degree. Most (9) reported their major field of study as social work

All research interviews were conducted using laptop computers and were checked for accuracy and completeness. At random, 5% of families were contacted and re-interviewed (briefly) to assure reliability and validity. Each child’s teacher was asked to fill out questionnaires on the child’s behavior, social competence, and self-regulation at each wave. Teachers were paid \$10 per child for each assessment.

C. Modifications to the Original Design

There were three important modifications to our original research design. The first was that the sample that could be recruited was 90% of the size originally intended. The second was that

participation rates were lower than in previous tests of SAFEChildren and similar interventions. Third, the results of the effectiveness trial did not permit us to conduct cost-benefit analysis in the same manner as proposed under Goal 4.

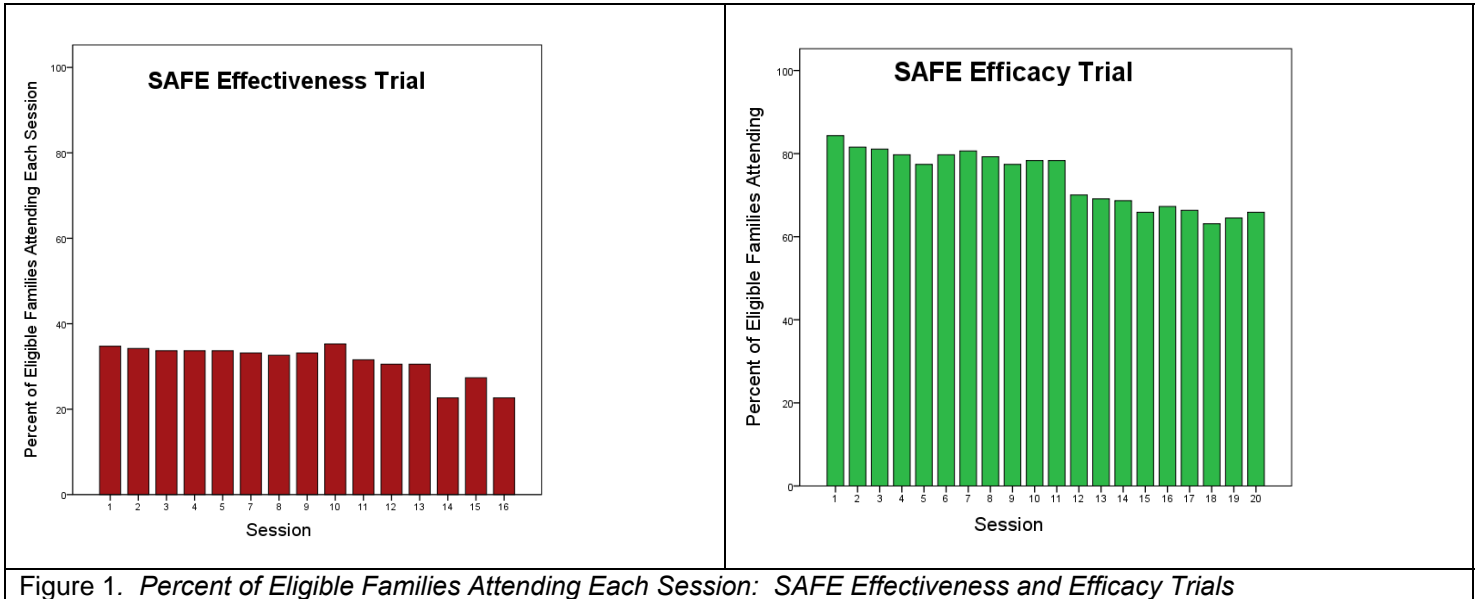


Figure 1. Percent of Eligible Families Attending Each Session: SAFE Effectiveness and Efficacy Trials

D. Findings

Overall, rates of participation in the intervention were considerably lower than the rates that had characterized the *SAFE Efficacy Trial*, as can be seen in Figure 1. We believe that lower rates of participation led to our finding fewer significant effects of the intervention than had been found previously. Specific findings relating to each goal of the proposal are reported below.

Goal 1: To test the effectiveness of the SAFEChildren intervention program when implemented by community intervention providers and tutors.

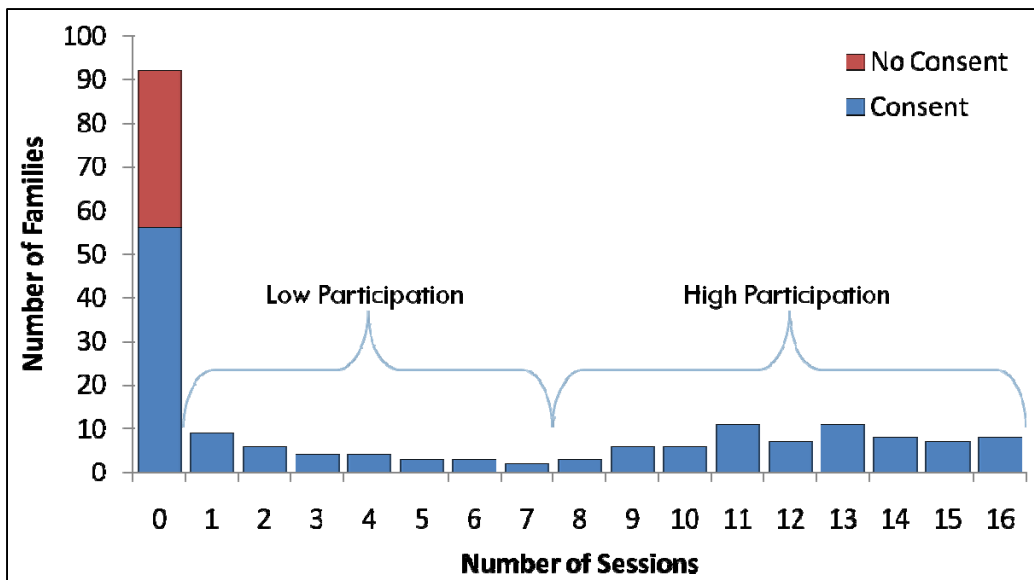


Figure 1. Participation in Family Sessions

1. Participation in Family Sessions Of the 191 students/families randomized to the treatment condition, 155 (81%) consented to participate in the intervention and 98 (53%) attended at least 1 family group session. Figure 2 is a histogram of family group attendance, divided into four participation groups: 1) Families who did not consent, 2) Families who consented but did not

participate, 3) Families who participated minimally in the sessions, and 4) Families who participated in at least half of the sessions.

2. Predicting Participation in the Family Groups. Because of the lower than expected level of participation and in order to better understand the process underlying participation, we modeled treatment participation as a function of baseline measures from four general domains: 1) characteristics of the primary caregivers, 2) characteristics of student participants, 3) characteristics of the families and family functioning, and 4) caregiver perceptions of their neighborhoods. Ultimately, a model in which participation was treated as four separate groups yielded the best fit and predictive accuracy. These four levels of participation were 1) no consent, 2) no participation, 3) low participation, and 4) high participation. Multinomial logistic regression was employed in which participation group was regressed on baseline measures. The final model included multiple measures from all four domains. The model demonstrated excellent predictive ability within the treatment group. Overall, the predicted group membership matched the actual group membership for 86.3% of cases.

Several variables emerged as predictors of attendance group. The strongest predictor of participation group was the child's initial reading ability. Parents of better readers were more likely to be in the high participation group. Parents who rated their communities as providing greater support were also more likely to be in one of the participation groups. Interestingly, being employed was associated with a higher likelihood of being in the no-consent group but a lower likelihood of being in the low attendance group, compared to the high participation group. Higher levels of parent education and more positive attitudes toward education were both associated with lower likelihood of low attendance compared to high attendance, as was fear of crime. Higher levels of parental monitoring were associated with lower probability of either no consent or no attendance, compared to the high attendance group.

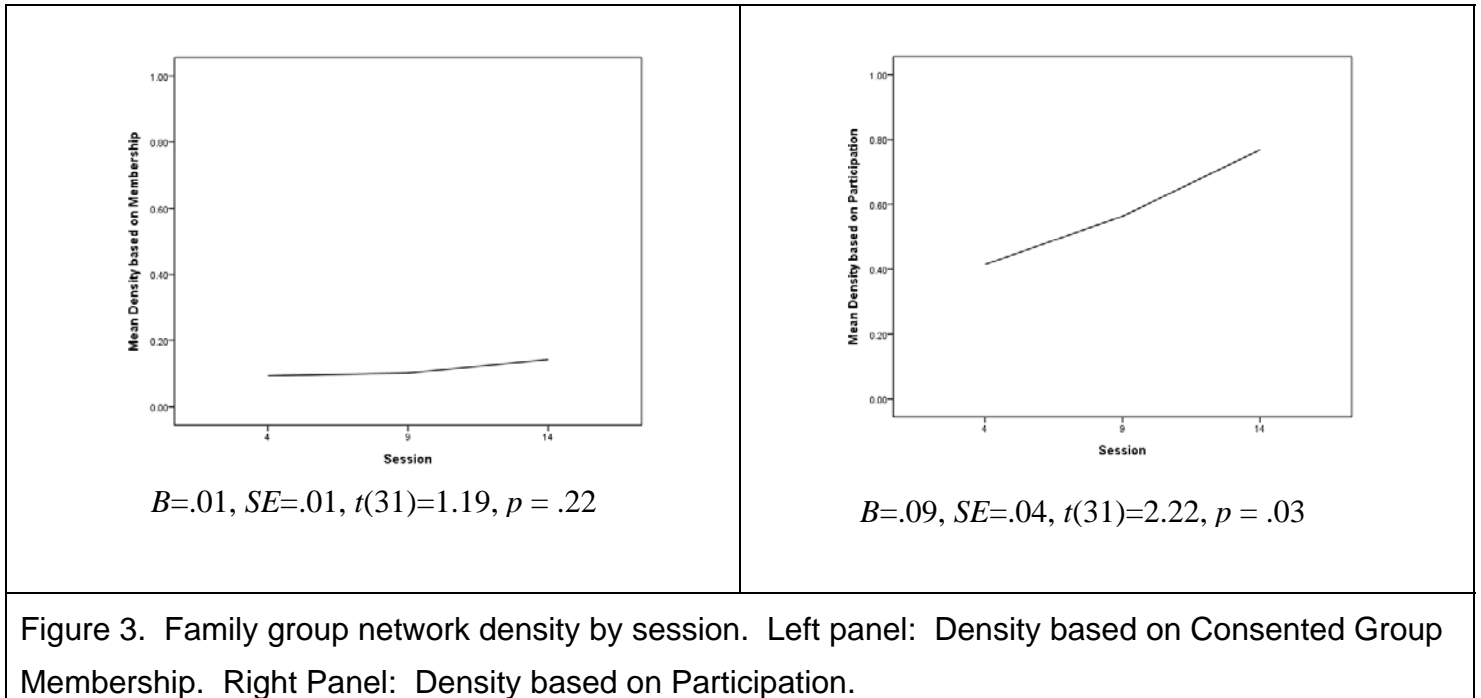
3. Overall Effects of the Intervention. No significant effects were found in the overall model academic achievement or school bonding. No significant effects were found on externalizing behaviors or adaptability. On parenting and family relationship characteristics, treatment participants showed slightly greater increases than controls in family organization ($p < .10$; $d = .11$). Attitudes of parents in the treatment group toward education declined modestly relative to the control group according to teacher reports ($p < .10$; $d = -.12$).

4. Effects for High-Participation Families. No significant differences were found in any of the measures of school functioning or on externalizing behaviors or adaptability. Among those in the high-participation group, treatment participants showed significantly greater increases than controls in social skills ($p < .05$; $d = .16$). Treatment participants also showed marginally greater increases in parental monitoring compared to controls ($p < .10$; $d = .12$), but family cohesion increased more slowly among treatment participants than among controls ($p < .05$; $d = -.13$).

Goal 2: To examine the role of aspects of social network processes in the intervention effects.

We tested two hypotheses about the network characteristics of the groups. The first hypothesis was that the groups would increase in their density as advice networks over the course of the intervention, and the second was that participation in the family groups would increase the likelihood that group members would contact other group members for help on parenting.

Density was calculated in two ways, one by dividing the number of contacts by the number of possible contacts between people assigned to the groups, and the other by dividing by the number participating in the groups. The increase in density was significant when calculated by the number participating, $B=.09$, $SE=.04$, $t(31)=2.22$, $p = .03$. Figure 3 illustrates these results.



We tested our second hypothesis using the item, “How likely would you be to contact each of the other parents/guardians in the group (excluding yourself) to ask for help with a parenting issue?” We found a significant session by previous acquaintance interaction, which is illustrated in Figure 4.

Parents who knew their fellow participants before the group started ($\beta = 0.42$, $t(657) = 3.14$, $p = .002$) were more likely to ask them for help with a parenting issue. As can be seen in Figure 4, the positive effect of time in the group on likelihood of contact for help with a parenting issue was limited to those who were previously acquainted.

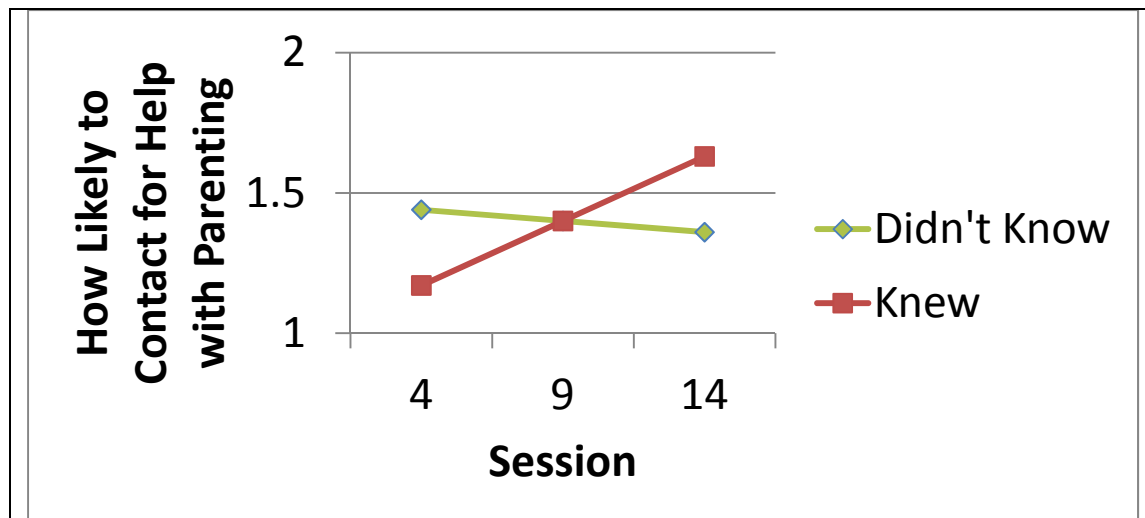


Figure 4. Likelihood of future contact for help with parenting as a function of group session and previous acquaintance.

Goal 3: To examine the relation of implementation qualities to variation in effects

Measures of Process Characteristics were collected from families and group leaders for 3 constructs: 1) Fidelity of Provision of Services (e.g., group activities); 2) Group Intervention Processes (e.g., attitudes about the group leader, family involvement); and 3) Provider Characteristics, (e.g., education, attitudes about prevention). All implementation data were scaled during the previous six months, and internal consistency reliabilities were found to be adequate for the process measures.

Fidelity. Measures of fidelity were completed by interventionists and parents. The extent to which session objectives were met was rated on five-point scales, with 1 = *not mentioned* and 5 = *fully completed* for items in the Activity Completion scale. Average percent of parents in complete agreement with therapists for each session were:

- Session 3 (5 items): 79.57%
- Session 8 (4 items): 78.12%
- Session 13 (5 items): 84.64%

Due to the high level of agreement between parents and therapists on the extent of activity completion, we averaged scores from the two reports for analysis of fidelity.

Parent and Therapist Intervention Process Characteristics. The 46 items administered to parents and the 36 items administered to therapists comprised eight subscales: (1) Trust/Positive Feelings of Parent, (2) Trust/Positive Feelings of Child, (3) Therapist Helpfulness to Parent, (4) Therapist Helpfulness to Child, (5) Program Helpfulness to Parent, (6) Parent Engagement, (7) Child Engagement, and (8) Effects of Program on Parent/Family.

Therapist attitudes about prevention. Six scales were administered to therapists to assess change in their attitudes about prevention in general and about the SAFEChildren intervention in particular: (1) *Attitudes about Prevention and Treatment*, (2) *Fit of the Program*, (3) *Benefits of the Program*, (4) *Agency Support*, (5) *Staff Support*, and (6) *Attitudes and Experiences with the SAFEChildren Program*.

Tutor process characteristics. Process measures were administered to tutors at four sessions. At the same four sessions, tutoring supervisors completed ratings of the tutors. The process

measures completed by tutors were *Tutees Relationship with the Tutor*, and *About the Program*. Tutoring supervisors rated tutors on *Quality of Relationship* and *Quality of Tutoring*.

The mean fidelity rating was 4.83 on a 5-point scale, indicating a very high degree of completion of the intended activities at each session. Attendance rates did not differ by the group in which parents attended the intervention sessions ($ICC = .07, Z < 1, ns$), but fidelity did differ significantly by intervention group ($ICC = .54, Z = 2.57, p < .01$). Fidelity in all groups was high, but three groups had average fidelity scores below 4.7 and others had average scores far in excess of 4.9.

Growth in Parent and Child Process Characteristics During the Intervention. Analysis of change in process characteristics over the course of the intervention revealed a strong and significant positive trend ($F(1, 74) = 12.30, p < .01$) with some variation by the specific measure ($F(10, 370) = 2.05, p = .05$). Figure 5 illustrates these results.

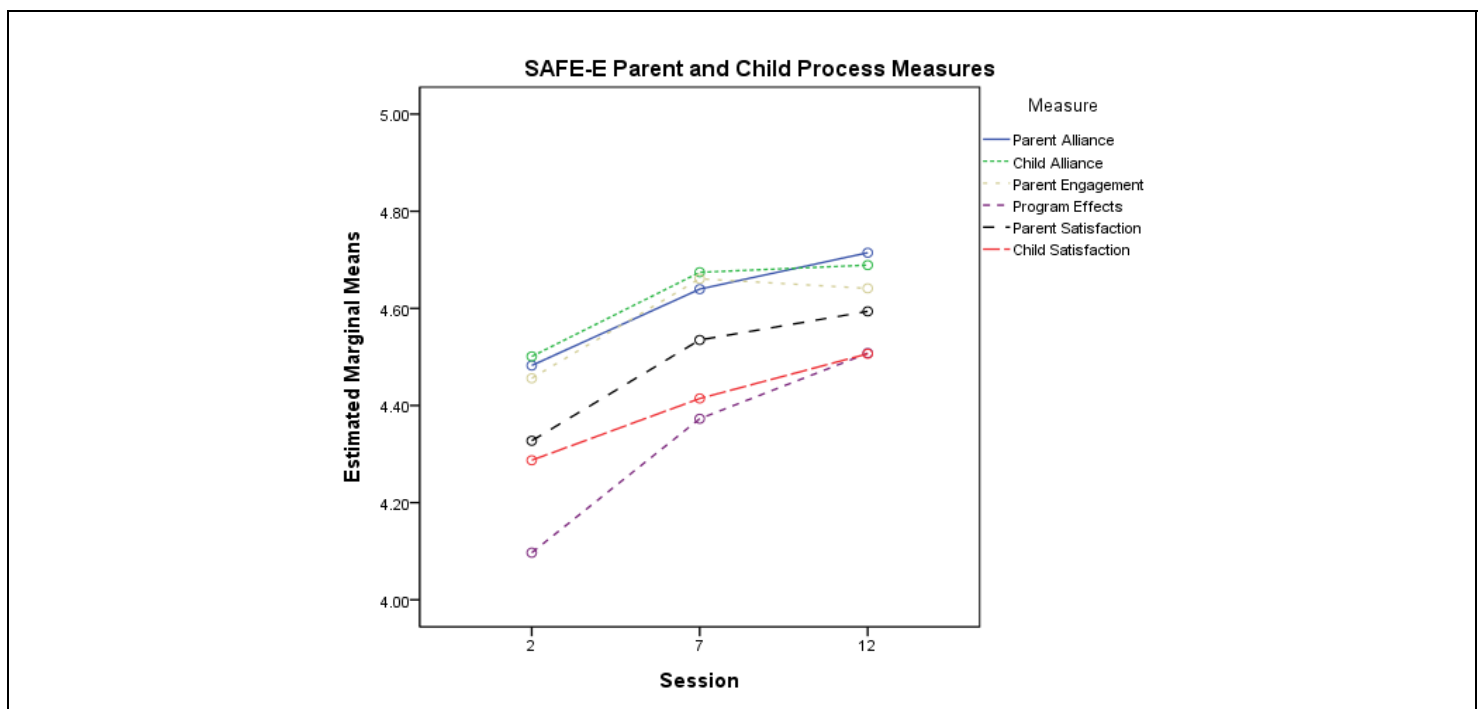


Figure 5. Parent and Parent Report of Child Mean Process Measures by Session.

Therapist Attitudes. Because of the small number of therapists, none of the pre-post comparisons on therapist attitudes was significant despite slight increases in each measure.

Tutor Process measures. Supervisors reported that tutor adherence to the “phonics” instructions in the manual increased significantly over the course of the sessions ($B = 0.02, SE = 0.01, t(519) = 2.41, p = .02$). Supervisors also reported greater adherence to the “Games” instructions ($B = 0.03, SE = 0.01, t(484) = 3.47, p = .001$) and to the “reading” instructions ($B = 0.04, SE = 0.01, t(448) = 4.59, p = .00$) in the manual as tutoring progressed. As tutoring progressed, the supervisors also reported increases in tutors maintaining the structure of the tutoring program (each segment 15 minutes, administers the appropriate segments) ($B = 0.03, SE = 0.01, t(521) = 3.41, p = .001$). No change over the course of tutoring was reported in maintaining a positive attitude, enthusiasm about tutoring, tutee interest in tutoring, tutor positive feedback to the tutee, behavior management, and session pacing.

Higher supervisors' ratings of the item, "tutee seems to like the tutor and tutoring" were significantly associated with higher weighted reading scores ($B = 2.62$, $SE=1.26$, $t(246)=2.07$, $p = .04$).

Goal 4: To conduct a cost-benefit analysis of the SAFEChildren intervention.

Our original plan had three objectives, but the final objective was modified:

Objective 1: To estimate the cost per child/family to provide the SAFEChildren program as a universal intervention to all first grade children in a school. Table 1 details the costs for implementing the SAFEChildren Program for 75 children in a school, divided by the tutoring and family intervention components of the program. These cost estimates are in 2011 dollars and are based on the actual costs to implement the *SAFE Effectiveness Trial*. Costs include all aspects of the intervention, including shared meals provided by the program. Implementation for 75 first graders would require 3 family interventionists and recruitment of up to 75 7th or 8th grade tutors.

Table 1		
<i>Costs for Administering the SAFE Intervention</i>		
Item	Cost details	Total Cost
Preparation and General Oversight		
	Initial Organizational Consultation	1.5 days
	Project Director/Clinical Supervisor	0.1 FTE/school
Tutoring Component (7 th -8 th grade tutors)		
	Tutoring Trainer / Supervisor	0.25 FTE/school
	Tutoring Manuals	\$20/ ea
	Tutoring Materials	\$75 (reproducible)
Family Component		
	Training for family interventionists	5 days
	Family Interventionists	3 @ 0.25FTE ea * \$50K
	Family Intervention Manuals	\$50/copy * 3 Interventionists
	Meals for 12 family groups	\$60/meal/group*16 sessions
Cost per School		\$92,495
Cost per Child for a single intervention		\$1,233.27
Cost per child for Initial and Booster Interventions		\$2,466.54

Objective 2: To estimate the societal benefits per child to age 18 in dollars. These estimates are the potential effects of the SAFEChildren program based on the long-term outcomes of the *SAFE Efficacy and Booster Trial* interventions found in the *SAFE Follow-up Study*. There were significant Type III tests and follow-up contrasts for the Booster vs. Initial Only effect on Combined Externalizing, and Elliot Level of Delinquency among those at high pre-existing family risk.

In addition, there were significantly fewer incidents of serious misconduct reported by the school among those assigned to the SAFE intervention than among controls, and marginally fewer violent incidents among those assigned to any intervention compare to control subjects.

There were no significant effects on high school graduation rates according to Chicago Public School records, but self reports coded to identify those who truly withdrew from school (i.e., did not enroll in GED programs or alternative high school) showed that, generally, youth who had been assigned to receive the initial intervention plus booster had lower rates of school withdrawal than those who were assigned to control or initial intervention only conditions.

Objective 3. To estimate the per child expenditure that would be justified by a long-term effect as large as that obtained in the SAFE Follow-up Study of the original interventions.

Monetizing the Costs of Outcomes

Regarding monetizing the benefits potentially associated with long-term effects of the SAFEChildren program, the *ex ante* methods of monetizing the potential social costs of crime, such as “willingness-to-pay” (WTP) and “contingent valuation” (CV) methodologies were developed in the environmental economics literature. These methods assess the public’s perception of value of issues or resources that may not be amenable to direct market valuation. Cohen and colleagues’ (2010; 2009) approach to the WTP method was to conduct a survey asking community respondents if they would be willing to vote for a proposal requiring each household in their community to pay a certain amount to be used to prevent one in ten incidents of certain crimes in their communities. Respondents were shown three different crimes, including burglary, serious assault, armed robbery, rape or sexual assault and murder. Cohen’s group looked at numerous crimes and the specific cost to each crime, and concluded that the typical “high risk” youth with 6+ police contacts (who collectively commit about 50% of all crimes), imposes between \$4.2 and \$7.2 million in costs on society. Discounted to present value at age 14, these costs total \$3.2–\$5.8 million. The bulk of these costs (\$2.7 million–\$ 4.8 million) are due to crimes, while an additional \$390,000–\$580,000 is estimated to be the value of lost productivity due to dropping out of high school. Juveniles that become career criminals imposes \$65,000 in costs through age 12 and \$230,000 through age 14.

Estimating the Potential Benefits of SAFE

Although the social costs of crime are not fully understood at present, for the purposes of this analysis, we are able to use Cohen’s (2010) estimates along with the *SAFE Follow-up Study* results around school reported misconduct and true withdrawal from high school (i.e., not attending alternative school or GED classes). Table 2 below reports the change in probability of a single offense, becoming a career criminal, and dropping out of high school, the upper and lower limits of estimates of social costs of each outcome, the expected lifetime change in the social costs, and, combined with our estimates of the costs of providing initial and booster SAFEChildren interventions, the potential benefit per child of the SAFE intervention.

Cohen and colleagues (2010) also provide estimates of the costs to society of single criminal offenses, crime careers, and high school dropout, their future values discounted to produce lifetime social costs, which are reported in Table 2. Cohen and Piquerro (2009) estimated that 4% of single offenders would become career criminals. We multiplied that by the 4% reduction in single offenses we found for SAFE III. We believe that this procedure produces a conservative estimate of the potential cost/benefit of SAFEChildren, because it is likely that schools did not always code a uniform discipline code offense for first offenders. It is possible that our estimates of reductions in single offenses are actually estimates of reductions in repeat offenders.

Putting together the monetized outcomes found to be significant in SAFE III with the actual costs of delivering the intervention, we produced the values in Table 2. As can be seen in Table 2, our most conservative estimate of the potential benefit of providing initial and booster SAFEChildren interventions is just over \$10,000 for every child provided with the intervention.

Table 2
Potential lifetime cost/benefit of the SAFEChildren Intervention

Behavior or outcome	Change in probability	Discounted Lifetime Cost		Expected change in lifetime social costs		Totals
		LL	UL	LL	UL	
Single criminal offense	0.04	\$53,000	\$78,500	\$2,120	\$3,140	
Career criminal ^a	0.0016	\$1,800,000	\$3,000,000	\$2,880	\$4,800	
High school dropout	0.039	\$200,000	\$450,000	\$7,800	\$17,550	
Lower Limit total lifetime benefits per dollar of intervention costs per child ^b						5.19
Upper limit total lifetime benefits per dollar of intervention costs per child ^b						10.34

^a Based on an estimated 4% of single offenders that will become career criminals (Cohen & Piquero, 2009)

^b Based on total costs of \$2,466 for delivering SAFE initial and booster interventions

E. Conclusions and Implications

The effects of a family-focused preventive intervention such as SAFEChildren may be less pronounced outside of the careful, close, supervision that characterizes research implementation. Careful examination of process and provider measures found no pronounced effects of differences in provider characteristics that could predict outcome. Although groups varied in average fidelity of implementation to a small but significant extent, participants and providers agreed that fidelity was uniformly high.

When SAFEChildren was implemented by university personnel (*SAFE Efficacy Trial* and *SAFE Booster Trial*), considerable effort was made to provide makeup sessions to families who did not attend group sessions. Makeup sessions were also provided in another large-scale trial of a family intervention that closely resembled SAFEChildren, namely the Multisite Violence Prevention Project, which found significant effects of the intervention on measures of violence. The mental health providers who delivered the family intervention in the *SAFE Effectiveness Trial* were not, as a rule, favorably disposed to visiting clients in their homes for makeup sessions. As a result, few makeup sessions were delivered, and none were delivered at the subject's homes.

Thus, an implication of this research is to emphasize the importance of home visitation for delivery of SAFEChildren and other similar interventions. Although it is impossible to determine from the *SAFE Effectiveness Trial* that home visitation for purposes of delivering intervention content in makeup sessions was the deciding factor in intervention effectiveness, the absence of home visitation in the *SAFE Effectiveness Trial* was a remarkable difference between it and the earlier efficacy trials. Moreover, this is consistent with substantial literature finding benefits of home visitation programs in the development of younger children in high risk settings (Farrington & Welsh, 2007).

As has been the case in prior SAFE studies and similar interventions, process measures indicate that the intervention can be delivered with high fidelity by community providers. With the exception of makeup sessions, as is noted above, measures of fidelity, alliance, and satisfaction

uniformly suggest that the characteristics associated with positive intervention impact in the literature, such as a strong initial alliance and satisfaction with the provider and the intervention were present in the *SAFE Effectiveness Trial*. Results of analyses of the tutoring process measures are consistent with high fidelity of tutoring program implementation by the seventh and eighth grade tutors. Both supervisor and tutor reports indicate that the program was delivered as intended.

Measures of social network processes within the groups suggest that group members increasingly came to rely on each other for parenting advice, and to incorporate each other into social support systems. These findings are consistent with the intent that the groups become sources of social support for members.

One of the central questions of this investigation concerned the extent to which it would be possible for community providers to deliver all components of the SAFEChildren intervention in a manner that is faithful to the intervention design and effective in changing key outcomes in early childhood. The foregoing evidence suggests that it is possible to deliver the SAFEChildren intervention with community providers, but that it may be difficult to acclimate mental health service providers to the assertive recruiting, follow-up, and provision of makeup sessions through home visits that characterizes successful preventive interventions.

The SAFE-E intervention was offered as a universal intervention to all families. However it is possible to conceive of SAFEChildren as a targeted intervention that would be offered to families with evidence of increased risk or whose children show early indications of risk. The analyses predicting participation in SAFE-E found no effects for child aggression but some evidence that better organized families in communities with more resources were more likely to participate fully in the groups. These findings suggest that offering SAFEChildren as a targeted intervention (or as a universal intervention) would require the expenditure of effort to recruit and maintain involvement in families for whom family organization deficits or neighborhood constraints form barriers to participation.

A final implication stems from the cost-benefit analysis. The social costs of youth risk are high enough that an expenditure equal to the cost of providing SAFEChildren universally is likely to be worthwhile. We conservatively estimate that it is possible to obtain at least five times the cost of providing SAFEChildren in societal benefits. However, such benefits will only be obtained if the intervention can be administered with levels of supervision, recruitment, and retention effort similar to those expended in the *SAFE Efficacy Trial*.

References

- Cohen, M., Piquero, A. & Jennings, W. (2010). Estimating the costs of bad outcomes for at-risk youth and the benefits of early childhood interventions to reduce them. *Criminal Justice Policy Review*, 21(4), 391-434.
- Cohen, M. A. & Piquero, A. R. (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, 25, 25-49.
- Farrington, D. P. & Welsh, B. C. (2007). *Saving children from a life of crime: Early risk factors and effective interventions*. New York, NY, US: Oxford University Press.
- Mason, C. A., Cauce, A. M., Gonzales, N. & Hiraga, Y. (1996). Neither too sweet nor too sour: Problem peers, maternal control, and problem behavior in African American adolescents. *Child Development*, 67, 2115-2130.

- McKay, M. M., Gonzales, J., Quintana, E., Kim, L. & Abdul-Adil, J. (1995). Multiple family groups: An alternative for reducing disruptive behavioral difficulties of urban children. *Research on Social Work Practice, 9*(5), 593-607.
- Multisite Violence Prevention Project (MVPP). (2004). The Multisite Violence Prevention Project: Background and overview. *American Journal of Preventive Medicine, 26*, 3-11.
- Tolan, P. & McKay, M. M. (1996). Preventing serious antisocial behavior in inner-city children: An empirically based family intervention program. *Family Relations, 45*, 148-155.
- Tolan, P. H., Gorman-Smith, D. & Henry, D. (2004). Supporting Families in a High-Risk Setting: Proximal Effects of the SAFEChildren Preventive Intervention. *Journal of Consulting and Clinical Psychology, 72*(5), 855-869.
- Tolan, P. H. & Guerra, N. G. (1994). Prevention of delinquency: Current status and issues. *Applied & Preventive Psychology, 3*, 251-273.
- US Department of Health and Human Services (USDHHS) Public Health Service (PHS); Office of the Surgeon General. (2001). *Youth Violence: A Report of the Surgeon General*. Rockville, MD: US Department of Health and Human Services.
- Wasserman, G. A. & Miller, L. (1998). The prevention of serious and violent juvenile offending. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks, CA: Sage.
- Webster-Stratton, C. (1990). Stress: A potential disrupter of parent perceptions and family interactions. *Journal of Clinical Child Psychology, 19*, 302-312.
- World Federation of Mental Health. (2002). *Toward a strategy for worldwide action to promote mental health and prevent mental and behavioral disorders*. Alexandria, VA: World Federation for Mental Health.

Evaluating the Implementation of a Family-
Focused Prevention Program: Effectiveness of SAFE Children
Final Technical Report

David B. Henry

University of Illinois at Chicago

Patrick H. Tolan

University of Virginia

Deborah Gorman-Smith

Michael E. Schoeny

Chapin Hall at the University of Chicago

Jack Zwanziger

Sage Kim

University of Illinois at Chicago

2006-JP-FX-0062

February 2012

Correspondence should be addressed to David Henry, Institute for Health Research and Policy,
University of Illinois at Chicago. 1747 W. Roosevelt Rd., Chicago, IL 60608. email:
Dhenry@uic.edu

Contents		page
I	Review of Relevant Literature	3
	A. <i>The Value of Prevention</i>	3
	B. <i>Developmental-Ecological Theory.</i>	3
	C. A Focus on Families Raising Children within the Inner-City	4
	D. The SAFEChildren Program.	4
	E. <i>Intervention Implementation Features of SAFEChildren.</i>	5
	F. <i>Efficacy of the SAFEChildren Intervention.</i>	5
	G. <i>Moving from Efficacy to Effectiveness</i>	7
II	Methods	8
	A. <i>Samples</i>	8
	B. <i>Procedures</i>	9
	C. <i>Variables and Measures</i>	10
	D. <i>Data Analysis Plan</i>	10
III	Modifications to or Problems with the Original Research Design	10
IV	Findings	11
	A. Goal 1: To test the effectiveness of the SAFEChildren intervention program when implemented by community intervention providers and tutors.	11
	1. <i>Participation in Family Sessions</i>	11
	2. <i>Predicting Participation in the Family Groups.</i>	12
	3. <i>Analysis Plan</i>	14
	4. <i>Overall Effects</i>	15
	5. <i>Effects for High-Participation Families</i>	16
	B. Goal 2: To examine the role of aspects of social network processes in the intervention effects.	23
	C. Goal 3: To examine the relation of implementation qualities to variation in effects	27
	1. <i>Measures of Process Characteristics.</i>	27
	2. <i>Sample</i>	27
	3. <i>Measures</i>	27
	4. <i>Analysis Plan</i>	28
	5. <i>Results for Attendance and Fidelity</i>	28
	6. <i>Growth in Parent and Child Process Characteristics During the Intervention.</i>	29
	7. <i>Therapist Attitudes.</i>	30
	8. <i>Tutor Process measures.</i>	30
	Goal 4: To conduct a cost-benefit analysis of the SAFEChildren intervention.	32
	Objective 1. <i>To estimate the cost per child/family to provide the SAFEChildren program as a universal intervention to all first grade children in a school.</i>	33
	Objective 2. <i>To estimate the societal benefits per child to age 18 in dollars</i>	33
	Objective 3. <i>To estimate the per child expenditure that would be justified by a long-term effect as large as that obtained in the SAFE-III follow-up study of the original SAFE-I and II interventions.</i>	45
V	Conclusions and Implications	48

I Review of Relevant Literature

There is now ample evidence that prevention affects later delinquency, drug use, school failure, and antisocial behavior (US Department of Health and Human Services (USDHHS) Public Health Service (PHS), 2001; World Federation of Mental Health, 2002). In recent years there has been a growing emphasis on the value of family-focused interventions to prevent violence for youth at high risk for violence (Metropolitan Area Child Study Research Group [MACS], 2002). Family-focused trials have suggested that efficacy is due to impact on parenting practices and family relationships (Lochman & Wells, 2002; Patterson, DeGarmo, & Forgatch, 2004). There is increasing interest in extending our understanding beyond the basic information that interventions can affect their intended outcomes (Farrington & Welsh, 2005). In particular, it is important to determine whether intervention effects that have been found in efficacy trials conducted under close supervision of university researchers can be obtained under conditions more closely approximating those that would exist if these interventions were implemented widely. This project aims to evaluate the effectiveness of the SAFEChildren program (Tolan, Gorman-Smith, & Henry, 2001), a two-component prevention program that has produced evidence for efficacy (Tolan, Gorman-Smith, & Henry, 2004) and conducting analyses related to its potential for larger scale implementation.

A. The Value of Prevention. Over the past twenty years a considerable body of knowledge has supported the value of prevention efforts regarding delinquency, violence, drug use, and school failure (Multisite Violence Prevention Project (MVPP), 2004; Tolan & Guerra, 1994; Wasserman & Miller, 1998). Developmental studies have clarified the most important risk factors to target in order to reduce youth violence (US Department of Health and Human Services (USDHHS) Public Health Service (PHS); Office of the Surgeon General, 2001; World Federation of Mental Health, 2002). As a result, a growing list of preventive interventions aimed at children and youth and targeting risk factors identified in basic developmental studies have been shown to be effective (Catalano, Arthur, Hawkins, Berglund, & Olson, 1998; Tolan & Guerra, 1994). Some have shown effects on long-term follow-up, while in others the promise is through impact on correlates or developmental predictors of delinquency and other antisocial behavior (US Department of Health and Human Services (USDHHS) Public Health Service (PHS); Office of the Surgeon General, 2001; World Federation of Mental Health, 2002). We know that prevention programs that consider developmental trajectories can be effective, reducing subsequent prevalence and seriousness of delinquency and antisocial behavior (Elliott & Tolan, 1999).

Meta-analytic studies (Aos, Lieb, Mayfield, Miller, & Penucci, 2004; Lipsey & Wilson, 1998) suggest that the most promising approaches have been efforts that emphasize parenting practices and improvement of family functioning (e.g., *Functional Family Therapy*, Alexander & Parsons, 1982; *Multi-systemic Therapy*, Henggeler, 2001; *The Incredible Years Program*, Webster-Stratton, 1998). These programs add to parent training by providing information about child development, support for parenting challenges and strain, and skill building and use of opportunities for better management of issues such as school performance, peer relations, and neighborhood conditions. They not only help improve potential deficits in parenting skills but also aid parents in effective use of skills, anticipation of child development tasks, better understanding of child needs, and more extensive and effective engagement with other systems influencing child outcomes (Tolan, 2002).

The promise of these approaches as well as empirical findings from our prior studies led to development of SAFEChildren (Metropolitan Area Child Study Research Group [MACS], 2002). Informed by developmental-ecological theory, we developed a family-oriented prevention program for those facing the exceptional challenges of raising children in inner-city communities (Mason, Cauce, Gonzales, & Hiraga, 1996).

B. Developmental-Ecological Theory. The content and theoretical basis of SAFEChildren is a developmental-ecological perspective on risk and prevention. This perspective is closely aligned with Bronfenbrenner's social ecological model of development (Bronfenbrenner, 1979, 1988). The

developmental-ecological model views time and growth as important considerations in understanding individual risk factors, how risk accumulates, how accumulated risk contributes to undesirable outcomes, and how interventions can affect risk. In essence interventions are meant to alter developmental course (Kellam & Rebok, 1992). Interventions are viewed as sensitive to age-related issues and the immediate developmental tasks facing the family. Thus, risk and intervention are understood within a conceptualization of development as trajectories toward or away from problematic outcomes (Gorman-Smith, Tolan, & Henry, 2000; Henry, Tolan, & Gorman-Smith, 2001).

The other central tenet of developmental-ecological theory is that individual development is influenced by the ongoing qualities of children's social settings. The same quality and style of family functioning may have different effects on risk depending on neighborhood residence (Furstenberg, 1993; Gorman-Smith, et al., 2000; Sampson, Raudenbush, & Earls, 1997). In addition, the challenges, resources, and opportunities for parenting and family management may vary depending on neighborhood characteristics (Sampson, et al., 1997).

C. A Focus on Families Raising Children within the Inner-City. There is considerable evidence of elevated rates of aggression and delinquency in inner-city communities. For example, Tolan & Henry (1996) found that rates of all types of psychopathology, as measured by the Teacher's Report Form of the Child Behavior Checklist, were above national averages among children living in inner-city neighborhoods. Aggression and delinquency rates, for example, were 2.5 and 2.8 times greater than would have been expected according to national norms. Similarly, Crane (1991) reported a sharp increase in risk of school dropout and teen pregnancy for adolescents living in inner-city neighborhoods over that found in other urban communities. These findings suggest a particularly risky developmental ecology associated with inner-city residence. Risk for children residing in the inner-city may be elevated apart from individual or family characteristics.

Characteristics of inner-city communities linked to increased risk include exposure to high rates of community violence (Richters & Martinez, 1993; Tolan & Gorman-Smith, 1996), absence of economic and social resources (McLoyd, 1990; Sampson & Laub, 1994), family disruption, economic homogeneity (Brooks-Gunn, Duncan, Klebanov, & Sealander, 1993), and lower levels of neighborhood support and involvement (Gorman-Smith, et al., 2000; Sampson, et al., 1997). In addition to the specific stressors associated with living in the inner city, families living in these communities are more likely to be burdened by chronic and serious health problems, but have less access to and familiarity with health care services (Aday, 1993).

D. The SAFEChildren Program. These ecological factors and developmental considerations led to development of the SAFEChildren preventive intervention. The intent in its development was to aid parents raising children within inner-city communities to:

1. promote good adjustment and success during the first grade year,
2. promote strong parenting and engagement with other parents, the school, and other resources,
3. increase protection from risks associated with residence in these communities, and
4. build children's competencies as they grow up.

We reasoned that there might be protective benefits from working to engage parents at the point of their children's first engagement with school to support children's academic adjustment and achievement during the first grade year. In addition to supporting parents, we thought it valuable to ensure opportunities for children to gain basic reading skills. From focus groups and prior studies we were aware that inner-city parents often feel isolated and in need of information about normal child development and how to manage developmental and social setting challenges. We reasoned that focusing on these areas would build protection and impede risk trajectories for aggression, violence and school failure.

As a result, SAFEChildren is comprised of a reading tutoring program and a family-focused intervention (20 weeks duration) provided during the child's first grade year. The family-focused

intervention is based in weekly multiple-family group meetings (about 5 families per group) that address issues of parenting, family relationships, child development, and parental involvement and investment in their child's schooling (McKay, Gonzales, Quintana, Kim, & Abdul-Adil, 1995; Tolan & McKay, 1996; Webster-Stratton, 1990). Multiple family groups are thought to build social support among participants, and improve parent-child interactions (McKay, et al., 1995). Although all family members meet together for the majority of sessions, there are some meetings in which parents and children meet separately to allow for concentrated focus on child social competence and self-control skills.

The reading tutoring program emphasizes phonics as well as word recognition. It combines training in these skills with application to reading high-interest, relevant material with positive messages and activities. These materials are designed to increase parental involvement in their child's academic development (Coie & Krehbiel, 1984; Wallach & Wallach, 1976). This program complements the phonics-based reading program used by the Chicago Public Schools. We work with the reading coordinator and the first grade teachers in each school to maximize the consistency of our tutoring with the primary work in the class. The program has been demonstrated to be effective with low-readiness children from poor, urban backgrounds (Coie & Krehbiel, 1984; Wallach & Wallach, 1976).

E. Intervention Implementation Features of SAFEChildren.

Our development of SAFEChildren was also influenced by concerns about issues of recruitment and retention for family-focused interventions (Dishion & McMahon, 1998). Attaining adequate participation and retention rates is not only important for valid evaluation but also for determining their likely utility (Tolan, 2002). As such, our approach has been to engage families in the positive potential of the group, to give careful consideration of practical and perceptual barriers to engagement, and to work to make groups accessible, useful, and engaging (McKay, et al., 1995).

Throughout this document we report on the effects of four studies. The *SAFE Efficacy Trial* tested the effects of the intervention with a cohort of first graders. The *SAFE Booster Trial* investigated longer-term effects of SAFE and the efficacy of a booster intervention. The *SAFE Follow-up Study* assessed the distal effects of SAFE on youth at the end of high school, and the *SAFE Effectiveness Trial* assessed the effects of SAFE when administered with community mental health providers and upper-grade children as tutors.

F. Efficacy of the SAFEChildren Intervention.

Initial Effects. In the *SAFE Efficacy Trial* we randomly assigning 424 families residing in inner-city neighborhoods with a child entering first grade to either a control or intervention condition. Outcomes were measured by assessments of children's reading ability and attitudes toward school, parent reports of family relationships and parenting practices, and both parent and teacher reports of parental involvement in school and child behavior. The first phase of the SAFEChildren project included four waves of parent and child interviews and five waves of teacher interviews.

The results are presented in detail in Tolan et al. (2004; in Appendix 8) and summarized in Table 1. The intervention resulted in increased levels of academic achievement and parental involvement in school. Intervention participants increased reading skills at a rate approximating national norms and were actually a bit above the national average for reading skills by mid second grade (2 years, 6 months), with an average score grade equivalent (GE) of 2.9. In contrast, control condition students were below the national average at the same point in time (GE = 2.4). Intervention families maintained levels of parental involvement in their children's schooling over the 2½ years, but control families showed decreasing parental involvement.

Effects Related to Pre-Intervention Risk Level. The *SAFE Efficacy Trial* also tested the proximal effects of the intervention on two high-risk sub-samples based on baseline risk: those with poorer family functioning and those with elevated child aggression. In each case about 25% of the sample met criteria for the specific designation of high risk.

Intervention children in high-risk families showed decreased aggression over time but high-risk

controls showed essentially no change. Intervention children from high-risk families had positive slopes on Concentration and improved Social Competence while the control children showed no change. We also found that among high-risk families, the crucial skill of Parental Monitoring improved for those in the intervention, while pre-intervention levels were unchanged for high-risk controls. Children with initially high aggression who were assigned to the intervention condition had decreasing levels of Aggression and Hyperactivity compared to high-aggression controls. Aggressive children in the intervention also showed greater improvement in Social Competence and less decline in parent involvement than equally aggressive controls.

Table 1: Summary of Key Proximal Outcomes of SAFE Children

Model	Outcome/Condition	Slope Est. (B)	SE	df	t
Overall Effects					
	WDRB Total Reading Composit (Grade Equivalent)				
	Control	1.27	0.07		
	Treatment	1.53	0.07	1067	2.76**
	Parent-rated Involvement in Child's Education				
	Control	-0.17	0.09		
	Treatment	0.04	0.08	1144	1.77†
Effects for High-Risk Families					
	Aggression				
	Control	-0.03	0.19		
	Treatment	-0.55	0.19	1106	-2.02*
	Concentration				
	Control	0.16	0.19		
	Treatment	0.68	0.18	1106	2.08*
	Social Competence (Adapatbility)				
	Control	0.05	0.20		
	Treatment	0.49	0.19	1107	1.70†
	Parental Monitoring				
	Control	0.14	0.19		
	Treatment	0.70	0.18	1097	2.26*
Effects for High-Risk Children					
	Aggression				
	Control	0.09	0.20		
	Treatment	-0.66	0.21	1106	-2.64*
	Hyperactivity				
	Control	0.18	0.19		
	Treatment	-0.29	0.20	1107	-1.72†
	Social Competence (Leadership)				
	Control	-0.03	0.19		
	Treatment	0.42	0.20	1107	1.67†
	Parent-rated Involvement in Child's Education				
	Control	-0.58	0.21		
	Treatment	0.14	0.22	1134	2.43*

Note. The slope estimates for the *t* values represent the coefficient for linear change over time (in years) from the growth model. Asterisks beside the *t* values for treatment indicate the level of significance for

$H_0: \text{Slope}_{\text{Treatment}} - \text{Slope}_{\text{Control}} = 0.$

† $p < .10$ * $p < .05$ ** $p < .01$

Table 2: Summary of Key Longer-Term Outcomes of SAFE Children

Model	Outcome	Intercept Difference Est. (B)	N	χ^2
Overall Effects				
	WDRB Total Reading Composite (Grade Equivalent)	0.26	331	4.50*
Effects for High-Risk Families				
	Parent-rated Involvement in Child's Education	0.48	331	2.75†
	Teacher-rated Parent Involvement in Child's Education	0.54	288	2.92†
	Family Organization	0.41	227	3.23†
Effects for High-Risk Children				
	Parental Monitoring	0.42	304	2.72†
	Parental Use of Effective Discipline Practices	0.54	304	3.36†
	Parent-rated Involvement in Child's Education	0.61	227	3.36†

Note. The intercept difference estimates for the χ^2 values represent the coefficient for the difference between treatment and control intercepts. Asterisks beside the χ^2 values for treatment indicate the level of significance for $H_0: \text{InterceptTreatment} - \text{InterceptControl} = 0$.

† $p < .10$ * $p < .05$ ** $p < .01$

Longer-Term Effects of the SAFEChildren Intervention. Longer term follow-up also provided support for the effectiveness trial. We were able to recruit and track 382 of the original 424 SAFEChildren participants for the second phase of the study that compared booster effects of an additional SAFEChildren intervention (during fourth grade) to those receiving the initial intervention only and to no-intervention controls. This study is labeled the *SAFE Booster Trial* in this report. By examining intercept differences between intervention and controls at pre-test for the second intervention, we were able to determine whether initial effects were sustained 3 years after intervention.

These analyses (see Table 2) indicated that those randomly assigned to SAFEChildren I maintained higher composite reading scores than controls. Although the effect on parental involvement was not sustained in these follow-up comparisons, it did remain in the comparisons of high-risk families on parent and teacher reports. Among high risk families, there was also evidence for sustained improvement in Family Organization.

There were several sustained effects on children with high initial aggression, including effects on Parental Monitoring, Parental Use of Effective Discipline Practices, and Parent Involvement in School. These results suggest maintenance of initial effects and emergence of new effects impacting those at greatest risk for later delinquency.

The results of the *SAFE Efficacy* and *SAFE Booster* trials suggest that SAFEChildren can promote stronger academic performance, child social competence, parental involvement, effective parenting practices, and reduced aggression; all predictors of lower risk for later delinquency and other antisocial behavior. This evidence raises the question of the potential effectiveness of SAFEChildren when implemented with community providers of the family intervention and tutors recruited from the student bodies of the participating schools.

G. Moving from Efficacy to Effectiveness.

Despite nearly 30 years of delinquency prevention research (Elliott, Huizinga, & Ageton, 1985; Hawkins & Weis, 1985), research findings on youth development and intervention have informed large scale programs and public policy to only a limited extent. There is need for strong evaluation of effectiveness and for greater knowledge about the key issues involved in moving interventions that demonstrate promise to being useful at full scale. There have been several formulations of key issues in effectiveness evaluation, all of which are oriented toward evidence-based prevention (e.g., Flay, et al., 2005; Sandler, et al., 2005; Spoth & Grenberg, 2005). Accompanying these conceptual formulations has been greater attention to practical considerations in undertaking such evaluations and valid measurement of key implementation issues in design of prevention trials (Tolan & Brown, 1998).

Consensus is emerging that effectiveness studies should test the practical utility and viability of promising interventions for “at-scale” or “real-world” implementation. As in efficacy trials, it is important that effectiveness trials employ strong evaluation designs with random assignment, longitudinal analysis, reliable and valid measurement, and sophisticated growth oriented data analytic methods. Effectiveness trials also can inform about issues involved in implementation and transition to typical practice. As is noted by Sandler et al. (2005), effectiveness trials function as “dress rehearsal” for going to scale. The Standards Committee of the Society for Prevention Research (Flay et al., 2005) suggests that effectiveness trials should provide good estimates of how the intervention can be implemented in actual practice, cost of such implementation, and understanding of for what population this intervention is intended/appropriate. These formulations guided the development of this effectiveness trial. This project incorporates measurement and analyses that fit with these desired qualities of effectiveness evaluations, and substantially increase the quality and extent of information yield from the study for prevention of delinquency and other antisocial behavior.

This final technical report consists of reports on research related to four separate goals, all related to the *SAFE Effectiveness Trial*. The first goal was to test the effects of the *SAFE Effectiveness Trial* intervention. Because of unexpected low participation rates, we include with this goal analyses aimed at understanding predictors of intervention participation. The second goal was to explore network processes within the groups and their relations to outcomes. The third was to explore pre-existing provider attitudes and process and fidelity measures, and their relations to outcome. The fourth and final goal was to conduct a cost-benefit analysis of the *SAFE Effectiveness Trial*. Specific methodological details are provided under the reports of each study goal below. Our overall study plan was to obtain a single baseline assessment before the intervention and then re-assess at post-test and at 6-, 12-, and 24-month post-intervention. Initial recruitment planning and collaborative relationships were formed prior to the beginning of the study, and were in place when initial recruitment began in the Spring of 2006, training the first set of providers began in the Summer of 2006, and recruitment of the first cohort in the Fall of 2006.

II Methods

A. Samples

Our primary target sample was the children (and their parents/caregivers) enrolled in first grade in five elementary schools in disadvantaged neighborhoods in Chicago, IL. In addition, we enrolled teachers as sources of data about the children, tutors, and mental health service providers as participants in the study. Table 3 reports the full number and ethnic distribution of all participants, children, parents/caregivers, 7th and 8th grade tutors, teachers, and community mental health providers. The sample for the *SAFE Follow-up Study* which contributed to the cost-benefit analysis is described under Goal 4.

Table 3: *SAFE Effectiveness Trial Samples*

Demographic	American Indian or Alaskan Native	Asian or Pacific Islander	Black, not of Hispanic origin	Hispanic	White, not of Hispanic origin	Other or unknown	Total
Females	8	1	405	129	1	6	550
Males	1	0	181	39	0	3	224
Unknown	0	0	174	2	0	228	404
Total	9	1	760	170	1	237	1178

Note: This table includes all participants: children, parents, teachers, community mental health providers, and tutors. Detailed demographic information was not collected on teachers and some tutors, resulting in substantial missing demographic information in this table.

Random assignment was at the individual level. Individual children and their families were randomly assigned to treatment or control conditions. In addition to the first grade participants, we recruited eighth-graders, and some seventh graders, who had been nominated by teachers to be trained and participate as tutors for the first-grade children in the treatment condition. Each tutor worked with one or two first graders who had been randomly assigned to receive intervention.

Two hundred seventy-eight (278) tutors were recruited. The median age of the tutors was 14 years, and the modal age was 13. Over two-thirds of the tutors were female (70.9%). Two-thirds (66.5%) reported their ethnicity as “African-American or Black” and 36.3% reported Hispanic or Latino ethnicity. Most reported that their mothers (87.4%) and fathers (90.3%) had at least graduated from high school. Nine out of 10 reported living with their mothers (90.6%) and 34.5% reported living with their fathers. The median number of sisters reported in each home was 2 and the median number of brothers reported was 1. Other family members reported in the home were grandmothers (21.2%), grandfathers (7.9%), aunts (11.9%), and uncles (11.5%).

We also recruited family intervention providers from two community agencies providing mental health services in the neighborhoods in which the schools were located. Nine therapists conducted groups in Cohort 1, and 8 therapists conducted groups in Cohort 2. Two therapists conducted groups in both cohorts, thus, 15 individual therapists conducted groups. Two therapists were male and 13 were female. Nine reported their ethnicity as African-American and 6 reported Latino/Hispanic ethnicity. Eight reported having a master’s degree, and 5 reported their level of education as associate degree or college degree. Most (9) reported their major field of study as social work

B. Procedures

1. *Securing Participation and Consent.* Because of our prior involvement with Chicago area inner-city communities, we had several schools interested in participating prior to the beginning of the project. The two community mental health agencies we recruited were agencies that served the communities in which the schools were located. We used a two-stage procedure for recruitment and consent of families. The participating schools provided us with the names, address and phone numbers of all children attending Kindergarten. Initial information was sent to all families. All information sent to families was translated into Spanish and included with English versions if the school served potentially Spanish speaking families. Following these letters, program staff contacted each family by phone or home visit. The research program was explained to families and informed consent was obtained. Families were asked to provide the names and phone numbers of two relatives who would always know where to find them in the event that the family moved without

providing forwarding information. Home visits were made to all families who could not be contacted by phone.

Eighth-grade tutors were nominated by their teachers. A consent procedure similar to that of the first-grade participants was followed to secure parent consent and child assent for participation. The tutors received 40 hours of training prior to beginning the intervention and received 30 minutes of supervision each week during the course of the intervention.

Providers were recruited and hired by their respective agencies, and trained by university personnel. They participated in an informed consent process for the research aspects of the study. In that process, providers were informed that they could refuse to participate in research-related assessments, but that assessments and observations normally required for supervision would be required.

2. *Data Collection.* We hired interviewers for data collection on a per-interview basis (\$45 per interview). All interviews were conducted using laptop computers and were checked for accuracy and completeness. At random, 5% of families were contacted and re-interviewed (briefly) to assure reliability and validity. Family interviews were scheduled around convenient times for the family. The interviewer completed the child interview prior to interviewing the caregiver(s) or parent(s). We interviewed two caregivers if present and agreeable.

Each child's teacher was asked to fill out questionnaires on the child's behavior, social competence, and self-regulation at each wave. Teachers were paid \$10 per child for each assessment. Parental consent and child assent were obtained for tutors' participation, for tutor completion of background and process questionnaires, and for gathering tutors' reading scores on the Iowa Test of Basic Skills. Family providers from the community mental health centers completed a questionnaire on demographic information and their experience in conducting preventive interventions similar to the intervention implemented in SAFEChildren. Tutors, family providers, and participants in the family intervention completed process assessments at four week intervals during the interventions for a total of six waves of assessment.

C. Variables and Measures

Each goal-related section below reports the specific measures used. Most measures to be completed by families were translated and back-translated in Spanish and evaluated for consistency. Fidelity and implementation measures for families, tutors, and providers were based on previous research and refined for this study. Most of the measures used in this research had been used extensively in our prior studies and in studies conducted by other researchers.

D. Data Analysis Plan

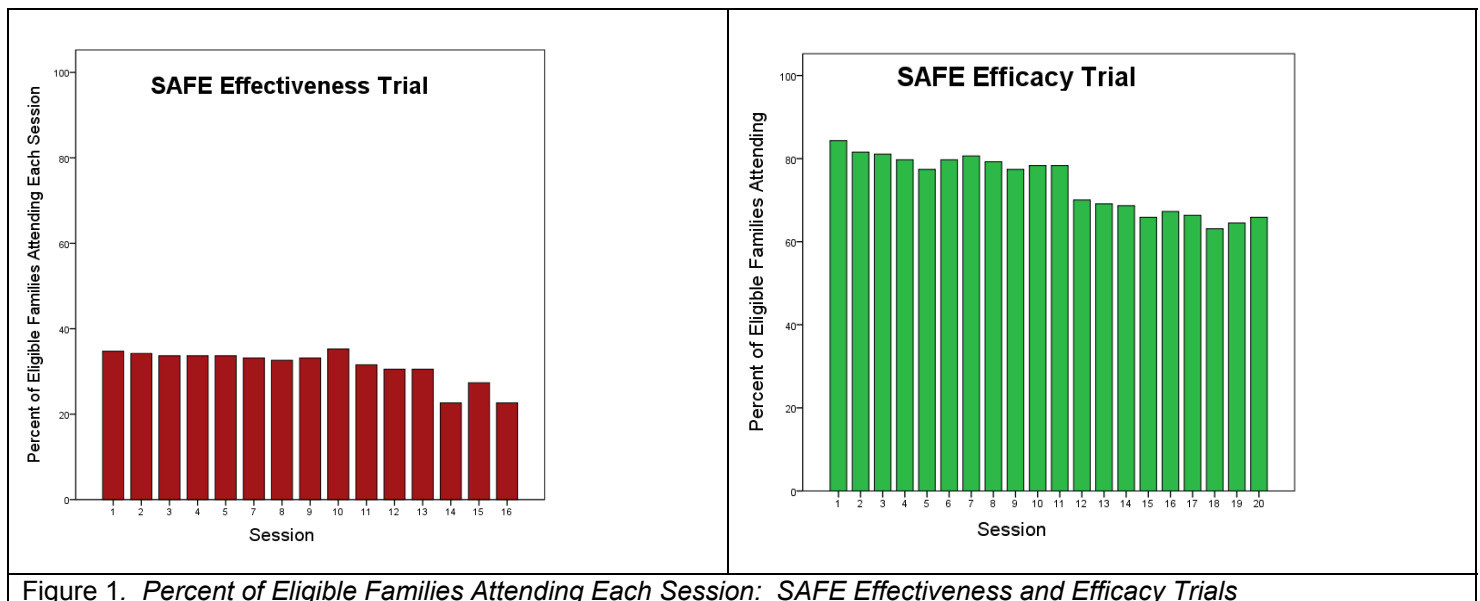
Each goal related to this proposal required different analysis approaches, thus, we include a general description here, and detailed explanation of the analyses under each goal of the study. For the first goal, namely to test the effectiveness of the *SAFE Effectiveness Trial* intervention program, we used multinomial logistic regression to predict participation in the family groups and random regression models (Bock, 1983, 1989; Gibbons, et al., 1993; Hedeker & Gibbons, 2006) for overall evaluation of the effects. For the second goal, namely to examine the role of social network processes in the intervention effects, we used social network analysis combined with mixed effects regression models to examine changes in network density and likelihood of contact for advice on parenting. Goal 3 was to examine the relation of implementation qualities to variation in effects. Growth models were used to investigate change in process characteristics and provider attitudes over the course of the intervention. We also investigated the relation of these characteristics to outcomes. Finally, the fourth goal was to conduct a cost-benefit analysis of the SAFEChildren intervention. Analyses related to this goal included estimation of dollar costs of the *SAFE Effectiveness Trial TRIAL* intervention, estimation of societal benefits using the *SAFE Follow-up Study* data of high school youth who had received the intervention as first graders, and estimating the ratio of costs and benefits per child.

III Modifications to or Problems with the Original Research Design

There were three important modifications to our original research design. The first involved the sample that could be recruited for the study, the second involved participation in the intervention, and the third involved the cost-benefit analysis proposed under Goal 4.

For reasons that included funding limitations and policy changes, the sample we were able to recruit was smaller than what we originally intended and proposed. Our original intention had been to solicit participation from 600 families. Based on the SAFEChildren efficacy trial, we expected 95% to agree to participate, and 84% of those who agreed to complete assessments, resulting in a sample of approximately 425 first-grade subjects who would complete the study. Ultimately, we were able to recruit 384 subjects who completed the study (90.3% of our intended final sample).

The second modification involved the rates of participation in the intervention. As can be seen in Figure 1, the proportion of participants who attended each of the intervention sessions in the *SAFE Effectiveness Trial* was substantially smaller than the rates that had characterized the original *SAFE Efficacy Trial*.



Lower rates of participation, combined with a smaller than anticipated sample, dampened the statistical power of the *SAFE Effectiveness Trial*, leading to finding fewer significant effects of the intervention, compared to the *SAFE Efficacy Trial*. Decreased power led to a third modification to our original design, concerning the intended cost-benefit analysis. We had proposed to base our cost-benefit analysis on the results of the *SAFE Effectiveness Trial* coupled with the association between immediate results of the original *SAFE Efficacy Trial* and the SAFE-III follow-up study. We did not obtain significant results on the same variables that were significant in the original *SAFE Efficacy Trial*, leading us to modify our strategy for the cost benefit analysis. Instead of our originally intended strategy, we will report under Goal 4 the outcomes of the SAFE-III follow-up study, and base our cost-benefit analysis on those long-term outcomes and the costs of administering the original SAFEChildren study. This will provide an estimate of the potential cost-benefit of SAFEChildren.

IV Findings

A. Goal 1: To test the effectiveness of the SAFEChildren intervention program when implemented by community intervention providers and tutors. Before turning to our intent-to-treat analysis of the effectiveness of the SAFE Children intervention, we describe the participation of the families in the intervention and report analyses used to predict participation.

1. Participation in Family Sessions Of the 191 students/families randomized to the treatment condition, 155 (81%) consented to participate in the intervention and 98 (53%) attended at least 1 family group session. Figure 2 is a histogram of family group attendance, divided into four participation groups: 1) Families who did not consent, 2) Families who consented but did not participate, 3) Families who participated minimally in the sessions, and 4) Families who participated in at least half of the sessions.

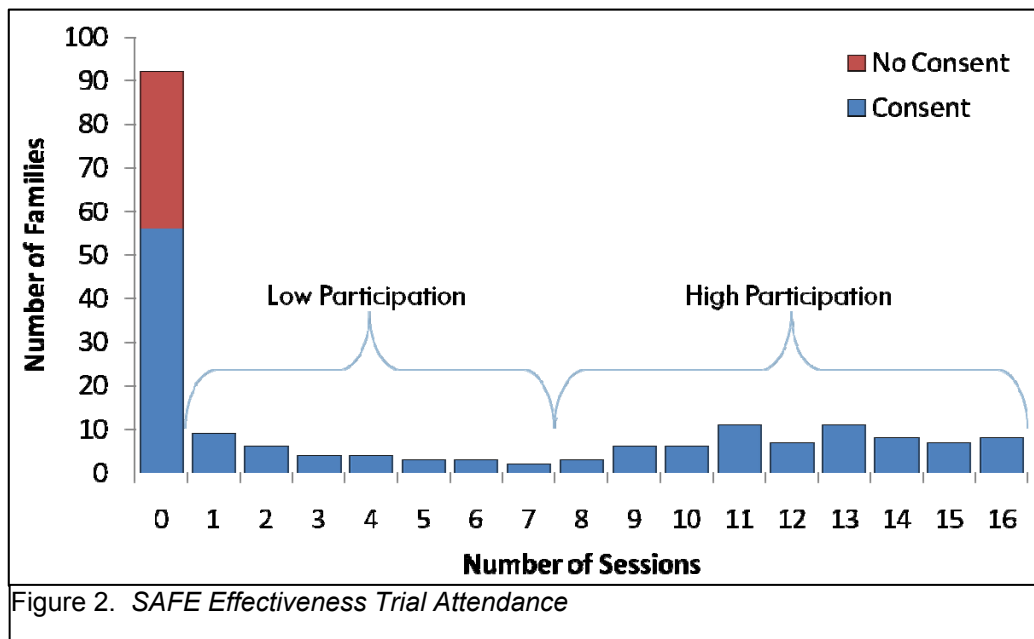


Figure 2. *SAFE Effectiveness Trial Attendance*

2. Predicting Participation in the Family Groups. Because of the lower than expected level of participation, and in order to better understand the process underlying participation, we modeled treatment participation as a function of baseline measures from four general domains: 1) characteristics of the

primary caregivers, 2) characteristics of student participants, 3) characteristics of the families and family functioning, and 4) caregiver perceptions of their neighborhoods. Ultimately, a model in which participation was treated as four separate groups yielded the best fit and predictive accuracy. These four levels of participation were 1) no consent, 2) no participation, 3) low participation, and 4) high participation (see Table 4 for a description of these groups). A multinomial logistic regression approach was employed in which participation group was regressed on baseline measures.

In a preliminary version of the model, there was an inordinately high level of misclassification between groups 2 (i.e., no participation) and 4 (i.e., high participation). Reviewing the results with some of our field staff generated the hypothesis that the employment status of the primary caregiver might explain some of the misclassification. Among families who might otherwise be predicted to participate, it was hypothesized that having a job might have interfered with their ability to attend family sessions. Similarly, among families who might otherwise be predicted to be non-participants, being unemployed might have enhanced the likelihood of participation. To test this hypothesis, we added interactions with employment status of the primary caregiver to the model. A number of significant interaction effects were found and were retained in the final model.

After removing non-significant predictors, the final model included multiple measures from all four domains (see Table 5). The model demonstrated excellent predictive ability within the treatment

group. Overall, the predicted group membership matched the actual group membership for 86.3% of cases (See Table 6).

Table 4: *Family Group Participation Levels*

Family Group Participation Level	Description	Group Size N (%)
No Consent	Consented for interviews and completed a baseline interview but either refused to participated in the intervention or could not be located by family group leaders.	36 (18.8)
No Participation	Consented to participate in the intervention but never attended the family group sessions. According to family group leaders, some of these families indicated that they wanted their children to receive the tutoring intervention, but did not have time or desire to attend family sessions.	57 (29.8)
Low Participation	Attended less than 50% of family group sessions.	31 (16.2)
High Participation	Attended 50% or more of family group sessions.	67 (35.1)

Table 5: Results of Multinomial Logistic Regression to Predict Level of Family Group Participation among Families Assigned to Treatment

Variable	Type III X ²	Parameter Estimates by Level of Participation ¹		
		No Consent	None	Low
Primary Caregiver Characteristics				
Caregiver Employment Status	12.27**	71.63*	.58	-73.82**
Caregiver Education	10.92*	2.49	-.61	-3.50*
Interaction with Employed	9.23*	-.74	-2.28**	1.92
Economic Stressful Events	11.61**	.97**	.13	.96**
Interaction with Employed	9.14*	-1.06*	-.16	-1.23*
Health Stressful Events	9.02*	.40*	.05	-1.35*
Fear of Crime	11.39**	4.34*	.55	-6.88*
Responses to Crime Fear	10.46*	16.72*	-.07	30.08**
Attitude toward Education	6.20	-3.48	-.74	-5.78*
Monitoring	13.15**	-3.62*	-.97**	.09
Child Characteristics				
Hispanic Ethnicity	12.83**	13.43*	.71	-25.14**
Social Skills	3.69	.09	.04	.16
Interaction with Employed	8.25*	-.03	-.20**	-.14
Child Temperament - Activity Lvl	1.61	.73	-.46	1.02
Interaction with Employed	9.97*	-1.10	1.82*	2.87*
Temperament (Fear)	9.12*	-2.31	.65	-2.90*
Adaptability	11.58**	.35	.06	.52**
Reading	19.42***	-6.85***	-.87**	-3.23**
Family/Household Characteristics				
Number of Children in the Home	2.48	1.43	.27	-.16
Interaction with Employed	6.70	-3.07	-.25	2.21
Number of Adults in the Home	3.80	3.00	.30	-.06
Interaction with Employed	9.23*	-.74	-2.28**	1.92
Beliefs about Family	14.67**	-20.41**	-3.52**	4.66
Cohesion	14.59**	21.10***	2.04*	4.19
Organization	9.39	-5.32	.80	-11.20**
Community Characteristics				
Community Belonging	8.58*	4.88**	.36	-.68
Community Involvement	13.11**	31.47***	2.97	23.06**
Interaction with CG				
Employed	9.88*	-47.43**	-.22	-18.03*
Community Problems	8.68*	-5.67*	-.15	2.25
Community Resources	9.19*	40.38	-1.13	-68.95**
Interaction with CG				
Employed	16.11**	-34.07	12.07**	54.03**
Community Support	15.28**	-8.35**	-1.44**	-1.86

* $p < .05$ ** $p < .01$ *** $p < .001$ ¹ High Participation is the Reference Category

Table 6: *Predicted vs. Actual Family Group Participation Levels*

Actual Group Membership	Predicted Group Membership ¹ N (% of total)				Total
	No Consent	No Participation	Low Participation	High Participation	
No Consent	34 (17.9)	0 (0.0)	0 (0.0)	2 (1.1)	36 (18.9)
No Participation	1 (0.5)	47 (24.7)	1 (0.5)	7 (3.7)	56 (29.5)
Low Participation	0 (0.0)	2 (1.1)	27 (14.2)	2 (1.1)	31 (16.3)
High Participation	1 (0.5)	8 (4.2)	2 (1.1)	56 (29.5)	67 (35.3)
Total	36 (18.9)	57 (30.0)	30 (15.8)	67 (35.3)	190 (100.0)

¹ Overall Prediction: 86.3% Correct

3. *Analysis Plan.* Our primary interest was in testing the effects of the intervention on growth trajectories of the intervention targets that are thought to affect or mark risk. Accordingly, our analysis plan centered on comparing the growth trajectories of participants assigned to intervention with those assigned to the control group. We included all participants who were randomly assigned to conditions in these analyses, regardless of attrition, dosage, or missing data. This intent-to-treat approach to analysis ensures that the randomness of assignment to conditions is preserved (Shadish, Hu, Glaser, Kownacki, & Wong, 1998) and that the comparisons reported reflect differences in growth trajectories associated with intervention.

We used a random regression approach to growth curve analysis to test hypotheses related to differential effects of the intervention on growth trends in the predictor variables. Growth curve models of multiple data points over time provide more reliable estimates of differential change because of intervention effects than do pre–post means comparisons (Muthén & Curran, 1997). In part this is because growth curve models allow for individual variation in measurement points within waves, such as those found in studies such as the present study, and can model random variation in effects (Bock, 1989; Gibbons et al., 1993). Random regression models are a class of mixed-effects linear models well suited to growth curve analysis of multiple repeated-measures data. Such models have been developed by Bock (1983, 1989), Gibbons et al. (1993), and Bryk and Raudenbush (1992). Random effects regression models assume that the available data at any given point of measurement estimate the group growth trend and each individual’s deviation from the group trend at that point of measurement. This assumption permits valid estimates of slopes and intercepts with cases that have missing waves of data, regardless of whether the data are missing at random (for reasons unrelated to the variables under study). They were developed, in part, because of the unreliability and limited precision of focusing only on pre–post intercept differences in evaluating intervention effects (Gibbons et al., 1993). They have been lauded as suited to prevention because they are sensitive to effects that emerge over time and consistent with a developmental approach to intervention effects (Muthén & Curran, 1997). The focus of the analysis and interpretation of effects is on the difference in growth attributable to intervention, rather than on the relative level between groups at any given point in time. We used 2 two-level models in these outcome analyses. The Level-1 models predicted an outcome variable from an individual intercept and linear slopes terms for wave of measurement (expressed in years), which was centered at the date of the baseline assessment. The Level-1 equations also included terms for family income and parental marital status at each wave of measurement. Thus, parental marital status and income were treated as time-varying covariates in these analyses.

Quadratic terms for slope were included as fixed effects to model any overall curvature in the shape of the growth curves. We limited interpretation of intervention effects to linear growth differences because our interest was in group differences in the linear rate of change associated with the intervention (see Barnes, Reifman, Farrell, & Dintcheff, 2000; Bryk & Raudenbush, 1992). The Level-2 equations predicted the Level-1 intercepts and linear slope terms by intervention condition, gender, ethnicity, and the child's school at the time of random assignment to conditions.

Analysis of participation level. Given the low levels of participation in the family groups, we then refit the models with terms for level of participation. Using the results from the participation prediction models, we created a binary variable for level of participation. High participation was defined as attending half or more of the family groups and low participation was defined as attending less than half of the family group sessions (i.e., collapsing the non-consent, non-participation, and low-participation groups). Using a propensity score method (Rosenbaum & Rubin, 1983), we created a variable reflecting the actual level of participation for the treatment group members and the predicted level of participation for the control group members. The participation models included Level-2 terms for level of participation and the interaction between that participation and intervention condition. Because a significant Intervention x participation interaction might not necessarily indicate a slope difference between intervention and control conditions among high participation families, we constructed planned comparisons evaluating the differences in linear slopes by intervention condition within the high-participation group. These comparisons were linear contrasts with one degree of freedom.

4. *Overall Effects.* Table 7 reports means and standard deviations by condition and wave for all participants. Table 8 summarizes the initial comparisons that focused on overall effects of the intervention (the entire intent-to-treat sample without consideration of differential impact by family group participation level). These models were fit without terms for participation level and without interactions between participation level and intervention condition. The slope coefficients and standard errors reported in the control rows represent the linear growth trend for control participants, and thus the expected developmental trend, in the absence of intervention of each variable. The slope estimates and standard errors in the intervention rows report the linear growth trends for participants randomly assigned to the intervention. The degrees of freedom, significance tests, and effect sizes relate to differences in linear growth from the control condition.¹

Child's school functioning. No significant slope differences were found in the overall model academic achievement. Both groups displayed steep growth in reading ability. There was no significant difference by intervention group in these comparisons on school bonding.

Child's behavior and social competence. No significant differences were found in the overall model for measures of a child's externalizing behaviors or adaptability. Both treatment and control participants showed increases in social skills over time, although the treatment participants showed marginally greater increases over time ($p < .10$; $d = .11$).

Parenting and family relationship characteristics. Control and treatment participants showed relatively stable levels of family organization over the course of the study. Treatment participants showed slight increases in family organization over time ($p < .10$; $d = .11$). We had separate parent and teacher measures of parental attitudes toward education. By parent reports attitudes for both treatment and control went down over time whereas teacher reports indicated that they went up over time. In both cases, attitudes for parents in the treatment group got worse relative to the control

¹ Slope difference effect sizes cannot be directly equated to the more commonly reported intercept difference effect sizes. The former represent relative difference in growth per unit time, whereas the latter represent differences at a given point in time, albeit controlling for prior differences in level. Because slope differences are measures of expectable increases in differences over time, the same magnitude effect size does not equate to the same impact. Typically, slope effect sizes are smaller than what might be found in intercept effect size estimates. A modest or small slope difference can translate to a large impact over time. We report linear estimates here, controlling for any nonlinear effects.

group (i.e., went down more quickly by parent report and went up more slowly by teacher report). These differences between treatment and control were marginally significant ($p < .10$; $d = -.12$) for both measures. Overall comparisons did not show any significant effects for any of the other measures of parenting and family-relationship characteristics.

5. *Effects for High-Participation Families.* Table 9 reports the linear slope estimates and standard errors for participants who attended 50% or more of the family sessions (high participants) assigned to control and intervention conditions. The table also reports significance tests and effect sizes of differences in linear growth between those assigned to control and intervention conditions within the high-participation families. These families ($N = 123$; 56 control and 67 intervention) were defined by the participation prediction equation described previously. For families in the treatment group, this represents their actual participation whereas for those in the control group, this represents the predicted level of participation based on baseline characteristics and the results of the prediction model. Effects found here represent those specific to high-participation families if no significant difference was found for the overall comparison or represent differentiated (heightened) effects of high-participation families when a significant difference was found in the overall comparison.

Child's school functioning. No significant slope differences were found in any of the measures of school functioning.

Child's behavior and social competence. No significant differences were found in the overall model for measures of a child's externalizing behaviors or adaptability. Among those in the high-participation group, control participants showed small increases in social skills over time. In contrast, the treatment participants showed significantly greater increases in social skills over time ($p < .05$; $d = .16$).

Parenting and family relationship characteristics. Control and treatment participants showed increases in parental monitoring over the course of the study with treatment participants showing marginally higher increases over time ($p < .10$; $d = .12$). Family cohesion increased at a significantly slower rate among treatment participants than among control participant ($p < .05$; $d = -.13$). The remaining comparisons between treatment and control participants within the high-participation group did not show significant treatment effects for any of the other measures of parenting and family-relationship characteristics.

Table 7: Means, Standard Deviations, and Numbers Assessed by Wave of Measurement and Condition

	Base line			Post-test			1-Year Follow-up			2-Year Follow-up		
	M	SD	N	M	SD	N	M	SD	N	M	SD	N
Child School Functioning												
Total Reading												
Control	442.6	20.5	189	456.4	19.9	157	472.0	18.2	154	483.5	16.7	157
Treatment	444.1	24.0	188	457.3	23.1	161	472.1	19.6	152	482.2	18.1	161
CSI Feelings about School Scale - transformed												
Control	1.4	0.6	187	1.5	0.6	154	1.6	0.5	154	1.6	0.6	157
Treatment	1.4	0.6	189	1.5	0.5	161	1.6	0.6	153	1.6	0.5	161
CSI Feelings about Teacher Scale - transformed												
Control	1.2	0.6	187	1.3	0.6	154	1.4	0.6	154	1.4	0.6	157
Treatment	1.3	0.6	189	1.3	0.7	161	1.4	0.7	153	1.4	0.6	161
Child Behavior and Social Competence												
parent-teacher BASC externalizing - transformed composite												
Control	3.0	0.5	193	3.0	0.5	191	3.0	0.5	179	3.0	0.5	179
Treatment	3.0	0.5	190	3.1	0.5	189	3.1	0.6	179	3.1	0.5	178
parent-teacher BASC adaptability - composite												
Control	48.8	8.6	193	48.4	9.3	191	50.1	8.4	179	49.0	8.3	179
Treatment	47.0	8.4	190	47.0	9.4	189	47.4	9.3	179	47.2	9.1	179
parent-teacher BASC social skills - composite												
Control	49.9	11.6	192	50.0	11.9	160	48.7	11.2	158	49.6	11.4	161
Treatment	46.9	11.9	189	48.7	12.0	163	48.3	12.4	157	48.9	11.7	161
Parenting and Family Relationships												
parenting practices - clear rules - transformed												
Control	1.5	0.3	192	1.4	0.3	160	1.5	0.3	158	1.5	0.3	161
Treatment	1.4	0.3	190	1.5	0.3	163	1.5	0.3	157	1.5	0.3	161

 Table 7: Means, Standard Deviations, and Numbers Assessed by Wave of Measurement and Condition

	Base line			Post-test			1-Year Follow-up			2-Year Follow-up		
	M	SD	N	M	SD	N	M	SD	N	M	SD	N
Control	2.5	0.9	189	2.5	1.0	191	2.8	1.0	169	2.7	1.0	172
Treatment	2.4	0.9	190	2.6	1.0	187	2.5	1.0	168	2.5	1.1	174

Table 8: *Estimates, Significance Tests, and Effect Sizes by Treatment Condition for Overall Growth Models*

	Slope Estimate	SE	df	t	Effect Size (d)
Child School Functioning					
CWDRTOTW: Total Reading - Average W Score					
Control	23.41	1.14			
Treatment	22.55	1.14	923	-1.03	-0.07
CCSISCHR: CSI Feelings about School Scale - transformed					
Control	0.14	0.05			
Treatment	0.13	0.05	920	-0.20	-0.01
CCSITCHR: CSI Feelings about Teacher Scale - transformed					
Control	0.15	0.06			
Treatment	0.12	0.06	920	-1.10	-0.07
Child Behavior and Social Competence					
PTBEXTTR: parent-teacher BASC externalizing - transformed composite					
Control	0.09	0.03			
Treatment	0.11	0.03	950	1.38	0.09
PTBADPT: parent-teacher BASC adaptability - composite					
Control	0.97	0.57			
Treatment	0.85	0.57	950	-0.35	-0.02
PTBSOCT: parent-teacher BASC social skills - composite					
Control	1.42	0.59			
Treatment	2.01	0.59	950	1.72~	0.11
Parenting and Family Relationships					
PPPQRULR: parenting practices - clear rules - transformed					
Control	0.03	0.03			
Treatment	0.03	0.03	948	-0.15	-0.01
PPPQCONR: parenting practices - consequences - transformed					
Control	0.04	0.03			
Treatment	0.02	0.03	946	-1.06	-0.07
PPPQINVR: parenting practices - involvement - transformed					
Control	0.00	0.03			
Treatment	-0.01	0.03	948	-0.49	-0.03
PPPQMON: parenting practices - monitoring w1					
Control	0.41	0.09			
Treatment	0.42	0.09	948	0.26	0.02
PPPQPOSR: parenting practices - positive parenting -					

Table 8: *Estimates, Significance Tests, and Effect Sizes by Treatment Condition for Overall Growth Models*

	Slope Estimate	SE	df	t	Effect Size (d)
transformed					
Control	0.00	0.03			
Treatment	0.00	0.03	948	-0.31	-0.02
PPPQHARR: parenting practices - harshness - transformed					
Control	0.02	0.02			
Treatment	0.02	0.02	945	-0.11	-0.01
PFRSCOHR: family cohesion - transformed					
Control	0.04	0.03			
Treatment	0.02	0.03	949	-0.78	-0.05
PFRSORGR: family organization - transformed					
Control	0.01	0.03			
Treatment	0.04	0.03	949	1.74~	0.11
PFTPINV: Parent-report Parent Initiated Involvement					
Control	0.25	0.05			
Treatment	0.24	0.05	948	-0.39	-0.03
PFTPPAER: Parent-report Parent Attitude toward Education - transformed					
Control	-0.07	0.04			
Treatment	-0.11	0.04	948	-1.87~	-0.12
TTININV: Teacher-report Parent-initiated Parent Involvement					
Control	0.25	0.05			
Treatment	0.23	0.05	911	-0.50	-0.03
TTINATE: Teacher-report Parent Attitude toward Education					
Control	0.23	0.08			
Treatment	0.14	0.08	911	-1.81~	-0.12

Table 9: *Estimates, Significance Tests, and Effect Sizes by Treatment Condition for High Participation Families*

	Slope Estimate	SE	df	t	Effect Size (d)
Child School Functioning					
CWDRTOTW: Total Reading - Average W Score					
Control	21.18	1.45			
Treatment	20.69	1.37	921	-0.34	-0.02
CCSISCHR: CSI Feelings about School Scale - transformed					
Control	0.10	0.06			
Treatment	0.14	0.06	918	0.75	0.05
CCSITCHR: CSI Feelings about Teacher Scale - transformed					
Control	0.15	0.07			
Treatment	0.10	0.07	918	-0.88	-0.06
Child Behavior and Social Competence					
PTBEXTTR: parent-teacher BASC externalizing - transformed composite					
Control	0.10	0.03			
Treatment	0.11	0.03	948	0.28	0.02
PTBADPT: parent-teacher BASC adaptability - composite					
Control	0.49	0.69			
Treatment	1.05	0.66	948	0.91	0.06
PTBSOCT: parent-teacher BASC social skills - composite					
Control	0.72	0.70			
Treatment	2.24	0.67	948	2.54*	0.16
Parenting and Family Relationships					
PPPQRULR: parenting practices - clear rules - transformed					
Control	0.01	0.03			
Treatment	0.03	0.03	946	0.84	0.05
PPPQCONR: parenting practices - consequences - transformed					
Control	0.04	0.03			
Treatment	0.04	0.03	944	0.16	0.01
PPPQINVR: parenting practices - involvement - transformed					
Control	0.00	0.03			
Treatment	0.00	0.03	946	-0.08	-0.01
PPPQMON: parenting practices - monitoring w1					
Control	0.31	0.10			

Table 9: *Estimates, Significance Tests, and Effect Sizes by Treatment Condition for High Participation Families*

	Slope Estimate	SE	df	t	Effect Size (d)
Treatment	0.47	0.10	946	1.79~	0.12
PPPQOSR: parenting practices - positive parenting - transformed					
Control	0.01	0.03			
Treatment	-0.01	0.03	946	-0.70	-0.05
PPPQHARR: parenting practices - harshness - transformed					
Control	0.02	0.03			
Treatment	0.03	0.03	943	0.37	0.02
PFRSCOHR: family cohesion - transformed					
Control	0.07	0.03			
Treatment	0.02	0.03	947	-2.05*	-0.13
PFRSORGR: family organization - transformed					
Control	0.01	0.03			
Treatment	0.03	0.03	947	0.79	0.05
PFTPINV: Parent-report Parent Initiated Involvement					
Control	0.19	0.06			
Treatment	0.26	0.06	946	1.15	0.07
PFTPPAER: Parent-report Parent Attitude toward Education - transformed					
Control	-0.08	0.04			
Treatment	-0.11	0.04	946	-0.77	-0.05
TTININV: Teacher-report Parent- initiated Parent Involvement					
Control	0.20	0.06			
Treatment	0.23	0.06	909	0.56	0.04
TTINATE: Teacher-report Parent Attitude toward Education					
Control	0.23	0.10			
Treatment	0.11	0.09	909	-1.50	-0.10

B. Goal 2: To examine the role of aspects of social network processes in the intervention effects.

One of the objectives under this goal was to evaluate the extent to which group participants would be likely to seek help from other participants as a result of group participation. The findings presented in this section were also part of a presentation given at the Biennial Meeting of the Society for Community Research and Action in Chicago in June 2011.

With very slight change, consented membership in the family groups stayed constant throughout the sessions at a mean size of 6.5 members (SD=2.57). Participation, however changed somewhat over the course of the sessions. The average number of participants at Session 4 was 3.05 (SD=1.43), or just under half of the number who had consented to participate in the intervention. The average number participating increased slightly to 3.10 (SD=1.37), by Session 9, and then decreased markedly to 2.47 (SD=1.23) in Session 14.

We tested two hypotheses about the network characteristics of the groups, and how they would relate to intervention effects. The first hypothesis was that the groups would increase in their density as advice networks over the course of the intervention. We defined the density of the groups as the number of contacts between group members outside of the group, divided by the number of possible contacts, which is a function of the group size:

$$Density = \frac{Contacts}{N(N - 1)}$$

The number of contacts was measured by the question, “How many times have you talked with <Group Member> in the past 2 weeks?” This question was asked of each group member with reference to every other group member. Possible responses were “0 times” to “4+ times.” For calculation of density, we coded each instance of any contact as a 1, and absence of contact was coded 0. We defined group size in two different ways, resulting in two different density measures. The first used the number assigned to the group in the denominator, and the second used the number participating in the group at each session as the denominator.

We tested this first hypothesis by fitting a mixed effects regression model of group network density on session. We used the “Pre” value (how well participants knew each other before the intervention) as a covariate in these models, and entered the interaction between previous acquaintance and time to determine whether the level of previous acquaintance within a group moderated the effect of group participation on density. As it turned out, the “Pre” value, though a significant predictor when density was calculated using the number consenting, was not significant when density was calculated with the number participating as the denominator. Interactions between time and previous acquaintance were not significant in either model, so we removed them from the final models.

The results of the models are reported in Tables 10 and 11, and illustrated in Figure 3. When density was calculated using the number consenting as the denominator in each group, the change over time was not significantly different from zero, $B=.01$, $SE=.01$, $t(31)=1.19$, $p = .22$, but prior acquaintance was a significant predictor, $B=.70$, $SE=.22$, $t(20)=3.12$, $p = .005$. However, when density was calculated with the number participating as the denominator, prior acquaintance was not significant, $B=1.32$, $SE=0.79$, $t(13)=1.66$, $p = .12$, but linear change was found to be positive and significant, $B=.09$, $SE=.04$, $t(31)=2.22$, $p = .037$.

Table 10

Mixed effects regression model predicting network density, calculated using the number consenting as the denominator as a function of prior acquaintance and session. N=19 groups

Effect	<i>Estimate</i>	<i>SE</i>	<i>df</i>	<i>t or Z</i>	<i>p</i>
Fixed Effects					
Intercept	0.04	0.03	22	1.62	.12
Prior Acquaintance	0.70	0.22	20	3.12	.005
Session	0.016	0.01	32	1.57	.13
Random Effects					
Family Group	0.004	0.001		3.66	<.001
Residual	0.006	0.002		1.80	.07

Note: Estimates of random effects are variances, and significance tests are Z-tests.

Table 11

Mixed effects regression model predicting network density, calculated using the number participating as the denominator as a function of prior acquaintance and session. N=19 groups

Effect	<i>Estimate</i>	<i>SE</i>	<i>df</i>	<i>t or Z</i>	<i>p</i>
Fixed Effects					
Intercept	0.38	0.10	19	3.72	.001
Prior Acquaintance	1.32	0.79	13	1.66	0.12
Session	0.09	0.04	32	2.18	0.037
Random Effects					
Family Group	.02	0.04		0.51	0.61
Residual	0.18	0.05		3.57	<.001

Note: Estimates of random effects are variances, and significance tests are Z-tests.

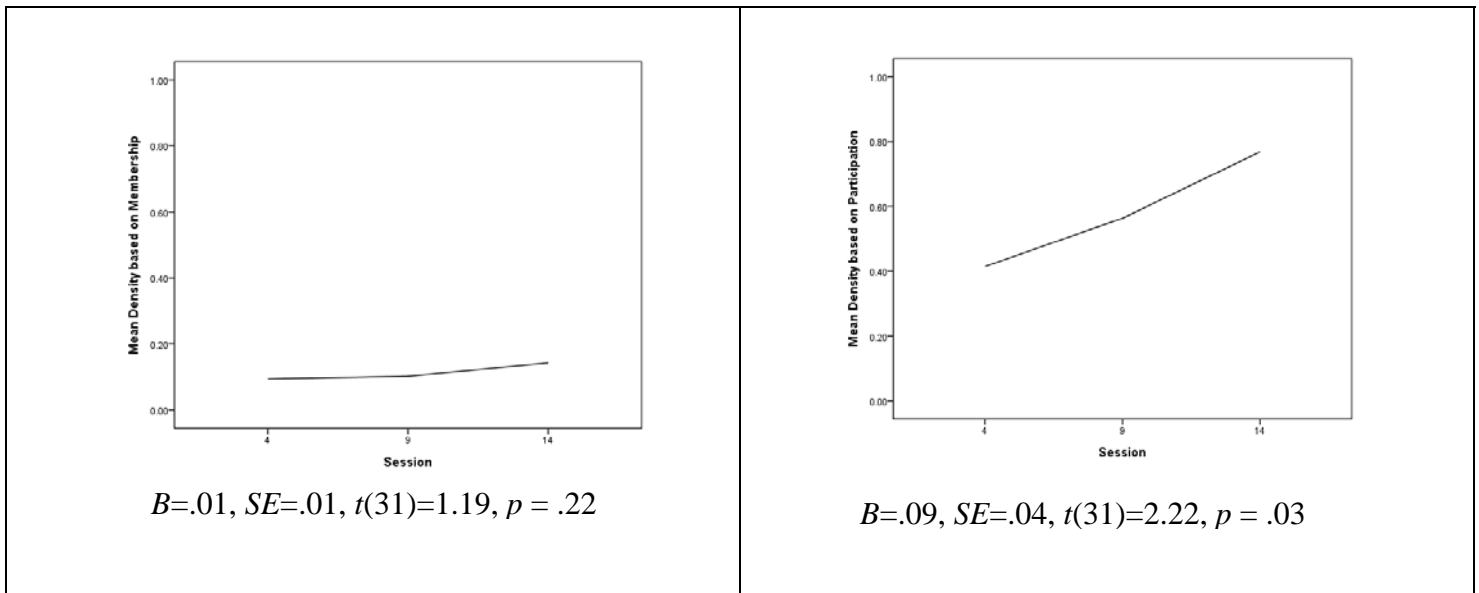


Figure 3. Family group network density by session. Left panel: Density based on Consented Group Membership. Right Panel: Density based on Participation.

Our second hypothesis involving network characteristics was that participation in the family groups would increase the likelihood that group members would contact other group members for help on parenting. The likelihood of future contact was assessed with four items at the end of the group network measure: 1) How likely would you be to contact each of the other parents/guardians in the group (excluding yourself) to ask for help with a parenting issue? 2) How likely would you be to contact each of the other parents/guardians in the group (excluding yourself) to ask for help/advice for another reason? 3) How likely would you be to talk to each of the other parents/guardians in the group (excluding yourself) at a school function? and 4) How likely would you be to talk to each of the other parents/guardians in the group (excluding yourself) at social function such as a party, cookout, church social?

Because our primary interest was that group members should become resources for help with parenting, we tested for the effect of time on the likelihood of future contact for help with a parenting issue. Because the data were at the level of the session, and session was nested within other persons in the group, each of which was nested within the individual respondent, who were nested within groups, we created a four-level mixed effects regression model predicting likelihood of contact from session. As in the test of the first hypothesis, we included prior acquaintance as a covariate, and tested the interaction of prior acquaintance with the time effect. We also included contact in the prior two weeks as a covariate. This analysis returned a significant session by previous acquaintance interaction, which is reported in Table 12 and illustrated in Figure 4.

Table 12
Mixed Effects Regression Model predicting likelihood of contact for help with a parenting issue as a function of prior acquaintance, session, and previous contact. N=98 participants in 19 groups

Effect	Estimate	SE	df	t or Z	p
Fixed Effects					
Intercept	1.85	0.29	18	6.29	<.0001
Prior Acquaintance	0.42	0.14	657	3.14	0.002
Contact - preceding two weeks	0.08	0.03	657	2.96	0.003
Session	0.36	0.19	657	1.92	0.055

Session*Acquaintance	0.25	0.09	657	2.6	0.01
Random Effects					
Family Group	.08	.10		0.83	0.204
Child Intercept(Family Group)	1.82	0.44		4.16	<.0001
Correlation between Child Intercept and Child Slope(Family Group)	-0.74	0.22		3.38	0.001
Child Time Slope(Family Group)	0.43	0.12		3.65	0.0001
Other Family(Child*Family Group)	0.06	0.02		2.81	0.0025
Residual	0.31	0.02		12.41	<.0001

Note: Estimates of random effects are variances, and significance tests are Z-tests.

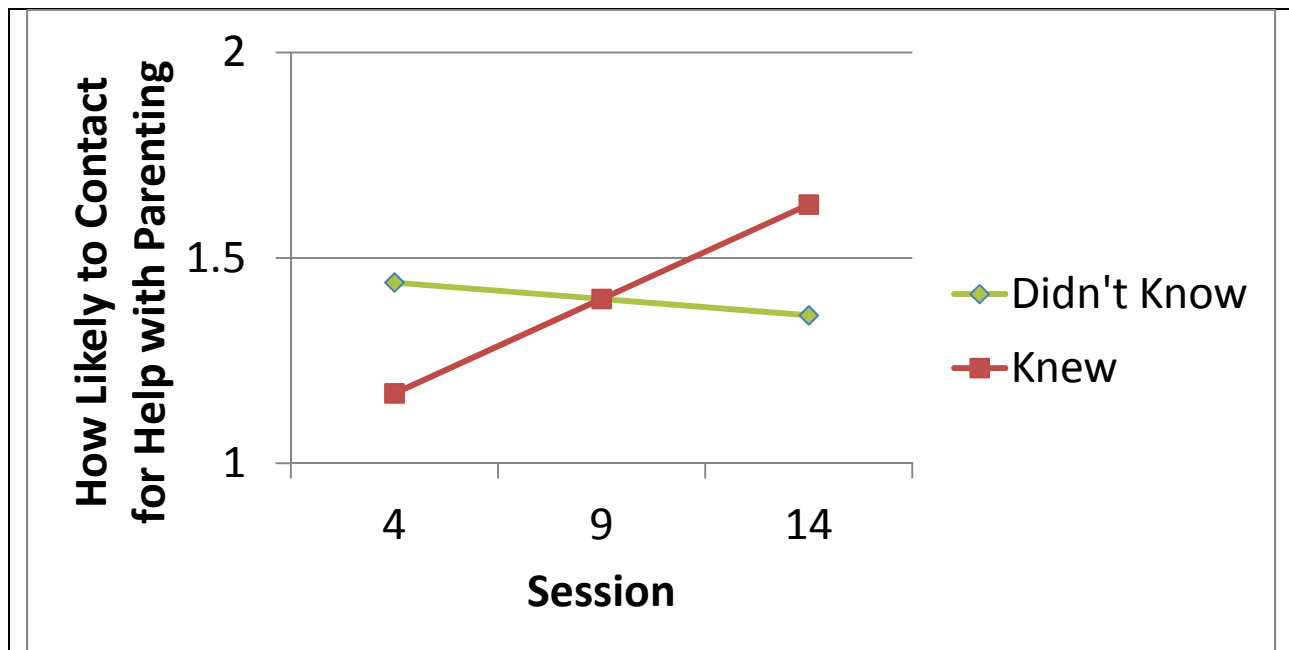


Figure 4. Likelihood of future contact for help with parenting as a function of group session and previous acquaintance.

Parents who knew their fellow participants before the group started ($\beta = 0.42$, $t(657) = 3.14$, $p = .002$) were more likely to ask them for help with a parenting issue. Contact in the two weeks preceding the session also predicted contact for help with a parenting issue ($\beta = 0.08$, $t(657) = 2.96$, $p = .003$). Controlling for those factors, the likelihood of contact increased as a linear function of time, ($\beta = 0.36$, $t(657) = 1.92$, $p = .05$), but this effect was moderated by a significant interaction between time and previous acquaintance ($\beta = -0.25$, $t(657) = 2.6$, $p = .01$). As can be seen in Figure 4, the positive effect of time in the group on likelihood of contact for help with a parenting issue was limited to those who were previously acquainted.

The results suggest that the intervention groups had, to some extent, their desired effects on the social contexts of the participating parents. However, the results also suggest that the process of change in social networks due to intervention groups is likely to be a relatively slow process. It does appear that the groups promoted the giving and receiving of parenting advice, particularly among those who knew each other prior to the beginning of the intervention. This cautiousness is consistent with previous findings that inner-city parents must be more vigilant than other parents in monitoring children and other parenting issues (Mason, et al., 1996). However, these results seem to clarify that family groups can increase trust and strengthen social bonds among parents. Future research should follow up on group members to determine the extent to which such strengthened social bonds are sustained over time.

C. Goal 3: To examine the relation of implementation qualities to variation in effects

1. *Measures of Process Characteristics.* Data were collected from families and group leaders for 3 constructs: 1) Fidelity of Provision of Services (e.g., group activities); 2) Group Intervention Processes (e.g., attitudes about the group leader, family involvement); and 3) Provider Characteristics, (e.g., education, attitudes about prevention). All implementation data were scaled during the previous six months, and internal consistency reliabilities were found to be adequate for the process measures.

2. *Sample.* The sample for the study of Process and Fidelity Characteristics consisted of the 97 parents and their children, 17 therapists, and 278 youth tutors enrolled in the study. Among the therapists, 88.1% were female and 56% reported African-American ethnic identification. 64.2% reported having a college degree and 47.1% reported a masters degree. Among the tutors, the median and modal age was 13 years. 72% of the tutors were female and 67.3% listed their ethnic identification as "African-American or Black."

3. *Measures.*

Fidelity. The provider self-report measure of implementation fidelity derived from the GREAT Schools and Families fidelity measures (Miller-Johnson, Sullivan, Simon, & The Multisite Violence Prevention Project, 2004) and was adapted from the Fidelity of Implementation Rating System (FIMP; Forgatch, Patterson, & DeGarmo, 2005). It assesses the extent to which session objectives were met, specified activities were completed, and adults and children were engaged in the session on three scales. Items are rated on five-point scales, with 1 = *not mentioned* and 5 = *fully completed* for items in the Activity Completion scale.

The parent fidelity measure also was derived from the GREAT Schools and Families fidelity measures (Miller-Johnson et al., 2004). It assesses parent perceptions of the extent to which specific activities were completed. Items are rated on a five-point scale, with 1 = *didn't come up* and 5 = *completed* for each intervention activity. Fidelity ratings were uniformly consistent across therapist and parent reports. Average percent of parents in complete agreement with therapists for each session were:

- Session 3 (5 items): 79.57%
- Session 8 (4 items): 78.12%
- Session 13 (5 items): 84.64%

Additionally, therapists and parents did not disagree in any session on whether participants were informed of weekly homework assignments or whether participants were paid for participation. Due to the high level of agreement between parents and therapists on the extent of activity completion, we averaged scores from the two reports for analysis.

Parent and Therapist Intervention Process Characteristics. The Intervention Process Measures were derived from similar measures used in the GREAT Schools and Families project (Miller-Johnson, et al., 2004). These measures assess impressions from parents and therapists about the parent-therapist relationship, parent satisfaction with the program, and program effects. Items are rated on five-point scales, with 1 as the low anchor and 5 as the high anchor; response categories vary from item to item.

The 46 items administered to parents in the SAFE-E study comprised eight subscales whose item count and internal consistency reliabilities of the scores were as follows: Trust/Positive Feelings of Parent (5 items, $\alpha = .82-.85$); Trust/Positive Feelings of Child (3 items, $\alpha = .78-.93$); Therapist Helpfulness to Parent (7 items, $\alpha = .84-.92$); Therapist Helpfulness to Child (2 items, $\alpha = .57-.78$); Program Helpfulness to Parent (8 items, $\alpha = .79-.92$); Parent Engagement (6 items, $\alpha = .79-.81$); Child Engagement (3 items, $\alpha = .76-.83$); Effects of Program on Parent/Family (12 items, $\alpha = .88-.93$).

The items administered to the therapists in the *SAFE Effectiveness Trial* comprised the same eight subscales with a different item count (36) than the corresponding parent scales. The internal consistency reliabilities of the therapist scores were as follows: Trust/Positive Feelings of Parent (5 items, $\alpha = .85-.86$); Trust/Positive Feelings of Child (3 items, $\alpha = .70-.82$); Therapist Helpfulness to Parent (7 items, $\alpha = .85-.90$); Therapist Helpfulness to Child (2 items, $\alpha = .26-.59$); Program Helpfulness to Parent (6 items, $\alpha = .87-.90$); Parent Engagement (6 items, $\alpha = .80-.85$); Child Engagement (3 items, $\alpha = .76-.83$); Effects of Program on Parent/Family (4 items, $\alpha = .78-.83$).

Therapist attitudes about prevention. Six scales were administered to therapists to assess change in their attitudes about prevention in general and about the SAFEChildren intervention in particular. *Attitudes about Prevention and Treatment* was a 7-item scale ($\alpha = .76$) on which higher scores indicated attitudes favoring clinical judgment and individualized treatment over manualized prevention programs. *Fit of the Program* (7 items, $\alpha = .90$) assesses the extent to which the prevention program fits with the values of the therapist and his or her perceptions of clients and the community. *Benefits of the Program* (10 items, $\alpha = .93$) assesses perceived benefits to the therapist of learning how to conduct the prevention program. *Agency Support* (4 items, $\alpha = .93$) assesses the extent to which the therapist believed his/her agency supported the program, and *Staff Support* (8 items, $\alpha = .81$) assessed the extent to which the therapist sensed the support of other staff members for the program. Finally, *Attitudes and Experiences with the SAFEChildren Program* (5 items, $\alpha = .81$) assessed approval of specific aspects of the SAFEChildren Program.

Tutor process characteristics. Process measures were administered to tutors at four sessions. At the same four sessions, tutoring supervisors completed ratings of the tutors. The process measures completed by tutors comprised scales for *Tutees Relationship with the Tutor* (5 items, $\alpha = .79$) and *About the Program* (8 items, $\alpha = .87$). Tutoring supervisors rated tutors on *Quality of Relationship* (5 items, $\alpha = .85$) and *Quality of Tutoring* (6 items, $\alpha = .91$).

4. *Analysis Plan.* First, we conducted analyses to determine the distribution of fidelity ratings among those who participated in the intervention. Next, we conducted analyses to determine the patterns of growth in parent and child intervention process characteristics, therapist attitudes, and tutor process characteristics. Finally, we investigated the relations among therapist attitudes, child and family process characteristics, and intervention outcome, using the single outcome (Social skills) that was found to be significant under Goal 1.

5. *Results for Attendance and Fidelity.* Figure 5 shows the distribution of averaged fidelity ratings for the entire study. The mean fidelity rating was 4.83 on a 5-point scale, indicating a very high degree of completion of the intended activities at each session. Attendance rates did not differ among the different groups in which the intervention was conducted ($ICC = .07, Z < 1, ns$). A mixed effects regression model of fidelity with intervention group as a random effect found that fidelity did differ significantly by group ($ICC = .54, Z = 2.57, p < .01$) either. Fidelity in all groups was high, but

three groups had average fidelity scores below 4.7 and others had average scores far in excess of 4.9.

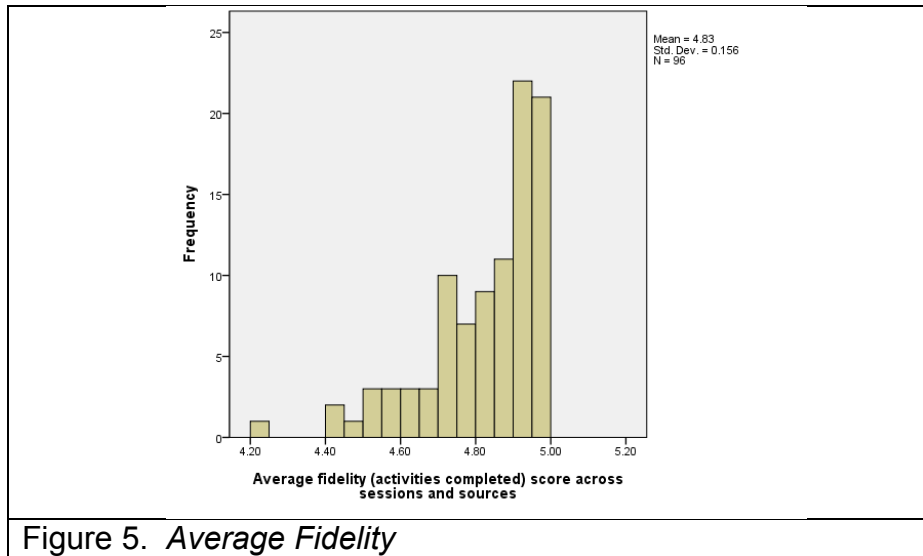


Figure 5. Average Fidelity

6. *Growth in Parent and Child Process Characteristics During the Intervention.* Analysis of change in process characteristics over the course of the intervention revealed a strong and significant positive trend ($F(1, 74)=12.30, p < .01$) with some variation by the specific measure ($F(10, 370)=2.05, p = .05$). Figure 6 illustrates these results.

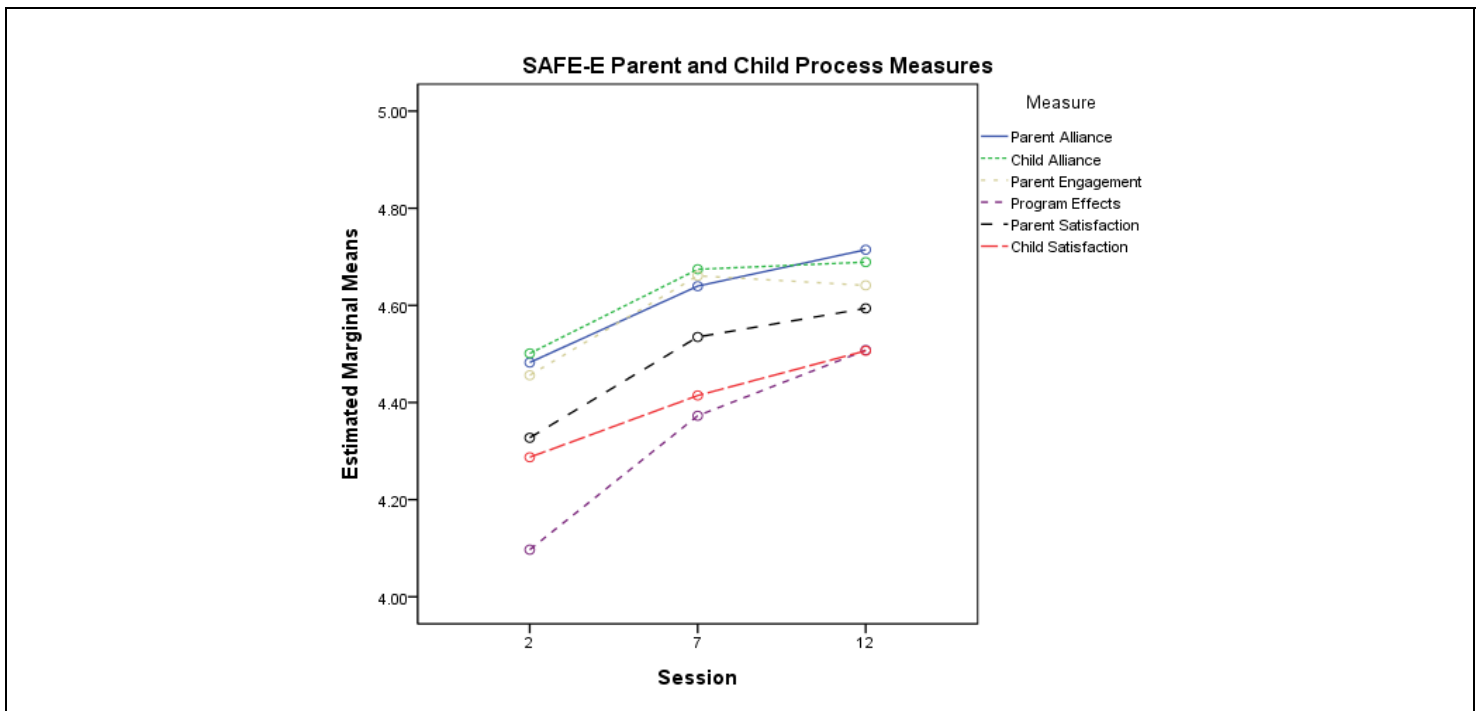


Figure 6. Parent and Parent Report of Child Mean Process Measures by Session.

Clearly, there was evidence of positive linear change in all process measures. However, exploratory analyses found no evidence that positive change in process characteristics was associated with total attendance.

7. *Therapist Attitudes.* Table 13 reports the pretest and posttest scores on the 6 scales measuring therapist attitudes and within-subject t-tests of pre-post differences. Because of the small sample size, none of the pre-post comparisons was significant despite slight increases in each measure.

	Mean	Std. Deviation	Mean	Std. Deviation	t (13)
Attitudes about Prevention and Treatment	4.00	0.69	4.12	0.73	0.81
Fit of the Program	4.89	0.72	4.99	0.74	0.59
Benefits of the Program	4.72	0.80	5.01	0.73	1.08
Agency Support	5.06	0.86	5.30	0.68	1.49
Staff Support	5.30	0.54	5.25	0.72	0.61
Attitude toward SAFE	4.80	0.43	5.30	0.64	0.98

8. *Tutor Process measures.* Table 14 reports the mean ratings of the tutoring supervisors, by session. Mixed effects regression models tested for significant change in these ratings over the three supervisory sessions in which measures were collected. Supervisors reported that tutor adherence to the “phonics” instructions in the manual increased significantly over the course of the sessions ($B = 0.02$, $SE = 0.01$, $t(519) = 2.41$, $p = .02$). Supervisors also reported greater adherence to the “Games” instructions ($B = 0.03$, $SE = 0.01$, $t(484) = 3.47$, $p = .001$) and to the “reading” instructions ($B = 0.04$, $SE = 0.01$, $t(448) = 4.59$, $p = .00$) in the manual as tutoring progressed. As tutoring progressed, the supervisors also reported increases in tutors maintaining the structure of the tutoring program (each segment 15 minutes, administers the appropriate segments) ($B = 0.03$, $SE = 0.01$, $t(521) = 3.41$, $p = .001$). No change over the course of tutoring was reported in maintaining a positive attitude, enthusiasm about tutoring, tutee interest in tutoring, tutor positive feedback to the tutee, behavior management, and session pacing.

Tutors also completed process measures detailing their perceptions of the sessions, and, as with the tutor supervisory ratings, we used mixed effects regression models to assess change. These measures were completed at Sessions 4, 8, 12, 16, and 20. These measures indicated that tutors saw increases in the tutee’s levels of trust ($B = 0.02$, $SE = 0.005$, $t(713) = 3.02$, $p = .003$), the tutee looking forward to the next session ($B = 0.01$, $SE = 0.005$, $t(711) = 2.08$, $p = .04$), and the tutee’s belief that tutoring was helping ($B = 0.01$, $SE = 0.005$, $t(708) = 2.21$, $p = .03$). Interestingly, the tutors also reported decreases in the tutor showing the tutee respect ($B = -0.01$, $SE = 0.003$, $t(730) = 3.42$, $p = .001$).

There were no differences in supervisor ratings by gender, age of the tutor, or mother’s educational attainment. We also tested whether tutor demographics or supervisory ratings were associated with outcomes on the Woodcock Diagnostic Reading Battery. This regression analysis returned a single significant effect. At the scale level, there were no significant effects either of intercept or slope of the

supervisory or tutor-report measures on change in reading scores. However, at the item level, higher supervisory ratings of the tutee seeming to like the tutor and tutoring were associated with higher weighted reading scores for the tutee ($B = 2.62$, $SE=1.26$, $t(246)=2.07$, $p = .04$).

	Session					
	4		8		12	
	Mean	SD	Mean	SD	Mean	SD
The tutor maintains a positive attitude during the session.	3.85	.978	3.69	1.037	3.81	1.089
The tutor is enthusiastic about tutoring (seems to enjoy working with the tutee).	3.78	1.002	3.50	1.087	3.72	1.119
The tutee seems to like the tutor and seems interested in the tutoring.	3.99	.840	3.90	1.064	4.04	1.012
The tutor gives the tutee compliments during the session.	3.76	1.113	3.66	1.044	3.79	1.096
The tutor is able to manage the tutee's behavior.	3.98	1.048	3.85	1.115	4.03	1.106
The tutor adheres to the "Phonics" instructions in the manual.	3.62	.913	3.54	.970	3.86	1.055
The tutor adheres to the "Games" guidelines in the manual.	3.79	.903	3.76	.872	4.13	1.005
The tutor adheres to the "Reading" guidelines in the manual.	3.75	.763	3.78	.630	4.18	1.000
The tutor maintains the structure of the tutoring program (each segment 15 minutes, administers the appropriate segments).	3.83	.912	3.84	.912	4.17	1.021
The tutor is well prepared and organized with materials.	3.75	1.039	3.58	1.084	3.91	1.131
The tutor paces the tutoring in response to the tutee's skill.	3.96	.884	3.69	1.002	4.06	1.043

	Session									
	4		8		12		16		20	
	M	SD	M	SD	M	SD	M	SD	M	SD
My tutee likes me.	4.15	.934	4.20	.940	4.25	.946	4.24	.970	4.38	1.21
I like my tutee.	4.30	.903	4.19	.946	4.34	.927	4.29	.898	4.53	0.90
My tutee trusts me.	3.95	1.05	4.14	.951	4.17	.945	4.14	1.03	4.47	1.02
I understand my tutee.	4.28	.922	4.28	.855	4.31	.929	4.27	.962	4.50	.707
I show my tutee respect.	4.80	.512	4.77	.498	4.66	.669	4.64	.698	4.82	.521
I am well-organized.	4.21	.818	4.16	.767	4.15	.899	4.12	.847	4.29	.719
My tutee is satisfied with the tutoring.	4.19	1.02	4.19	.964	4.18	.923	4.20	.931	4.47	.961
My tutee believes the tutoring is helping him/her.	4.16	1.02	4.24	.979	4.22	1.02	4.32	.919	4.44	.991
In the program, my tutee tries hard to improve.	4.25	.995	4.25	.988	4.23	1.01	4.22	.994	4.32	.976
Even when frustrated, my tutee keeps on trying to achieve her/her goals.	4.15	1.06	4.03	1.07	4.14	1.11	4.07	1.11	4.15	1.16
My tutee usually looks forward to the next session.	4.06	1.08	4.14	1.07	4.25	.969	4.18	1.04	4.26	1.11
Changes my tutee has made	3.78	.962	4.03	.940	4.17	.953	4.24	.915	4.44	.960
Tutee is trying to reach his/her goals	4.07	.991	4.14	.897	4.11	.953	4.08	1.01	4.24	1.02

Goal 4: To conduct a cost-benefit analysis of the SAFEChildren intervention.

As is noted above in the section on Modifications to the Original Research Design, the effects of the *SAFE Effectiveness Trial* did not permit us to do the cost-benefit analysis as planned. Our original plan had three objectives. Each objective and the modifications necessary are detailed below:

(1) *To estimate the cost per child/family to provide the SAFEChildren program as a universal intervention to all first grade children in a single school, based on the dollar costs accumulated by the SAFEChildren effectiveness study. These were to include costs of recruiting, training, and supervising community service providers for the family and tutoring components, as well as costs of service provision itself, but excluding the costs of the research components of the study, other than those that would be required for ongoing evaluation of the program if implemented. This objective remains unchanged.*

(2) *To estimate the societal benefits per child to age 18 in dollars. These estimates were*

to be based on the published costs of providing educational remediation and behavioral services, as well as juvenile justice services for children who require such involvement. Estimates of the benefits to age 18 were to have come from an analysis of significant effects of the *SAFE Effectiveness Trial* intervention and the association between similar effects on the SAFE-I intervention and SAFE-III measures collected at age 18. In relation to this goal we are able to report the effects of the original SAFE-I intervention that were found to be significant in the SAFE-III study.

(3) *To estimate the ratio of costs and benefits per child.* We intended to link the probabilities of desirable and undesirable events over time and estimate the effects of intervention on their associated societal costs. The modest effects of the *SAFE Effectiveness Trial* intervention make it impossible for us to do this. Therefore, we revise this objective as follows: *To estimate the per child expenditure that would be justified by a long-term effect as large as that obtained in the SAFE-III follow-up study of the original SAFE-I and II interventions.*

Objective 1: To estimate the cost per child/family to provide the SAFEChildren program as a universal intervention to all first grade children in a school. Table 16 details the costs for implementing the SAFEChildren Program for 75 children in a school, divided by the tutoring and family intervention components of the program. These cost estimates are in 2011 dollars and are based on the actual costs to implement SAFEChildren in the *SAFE Effectiveness Trial*. Costs include all aspects of the intervention, including shared meals provided by the program. Implementation for 75 first graders would require 3 family interventionists and recruitment of up to 75 7th or 8th grade tutors.

<i>Costs of Providing the SAFE Intervention</i>		
Item	Cost details	Total Cost
Preparation and General Oversight		
Initial Organizational Consultation	1.5 days	\$2,500
Project Director/Clinical Supervisor	0.1 FTE/school	\$7,000
Tutoring Component (7 th -8 th grade tutors)		
Tutoring Trainer / Supervisor	0.25 FTE/school	\$14,000
Tutoring Manuals	\$20/ ea	\$1,500
Tutoring Materials	\$75 (reproducible)	\$75
Family Component		
Training for family interventionists	5 days	\$3,250
Family Interventionists	3 @ 0.25FTE ea * \$50K	\$52,500
Family Intervention Manuals	\$50/copy * 3 Interventionists	\$150
Meals for 12 family groups	\$60/meal/group*16 sessions	\$11,520
Cost per School		\$92,495
Cost per Child for a single intervention		\$1,233.27
Cost per child for Initial and Booster Interventions		\$2,466.54

Objective 2: To estimate the societal benefits per child to age 18 in dollars. These estimates are the potential effects of the SAFEChildren program based on the long-term outcomes of the SAFEChildren Trials (*SAFE Efficacy Trial* and *SAFE Booster Trial*) found in the *SAFE Follow-up Study*. A summary of the *SAFE Follow-up Study* methods and results follows, after which we

calculate the potential per-child benefit based on these results.

Method

Participants

Participants in the *SAFE Follow-up Study* were youth who had participated in phases 1 and 2 of the original *SAFEChildren* study, as well as their parents and teachers. All families who participated in the *SAFE Efficacy Trial* study (424 families) were eligible to be recruited for participation in this study. Families were originally recruited through seven Chicago Public Schools that were selected based on their school and neighborhood demographic characteristics (e.g., 40% or more of families at the poverty level, crime level significantly above average for Chicago, predominantly African-American and/or Latino/Hispanic enrollment). Details of *SAFEChildren* I and II may be found in Tolan, Gorman-Smith and Henry (2004) and Tolan, Gorman-Smith, Henry, & Schoeny (2009).

Of the 424 recruited for *SAFE Efficacy Trial*, 401 (95%) completed the first five waves of assessment. Forty-five percent (45%) of the families reported African-American ethnic identification and 55% reported Latino ethnic identification. The sample included approximately equal numbers of boys and girls. Forty-four percent (44%) of families recruited had a primary caregiver who had not completed high school. Fifty-nine percent (59%) had family incomes of less than \$20,000 per year and 85% had family incomes of less than \$30,000 per year. Using existing archival records, comparisons were made between the 424 recruited families and their schools and communities for average income, ethnicity distribution, and marital status distribution. No significant differences were found. Among those randomized to the treatment condition ($n=225$), 79% ($n=178$) completed the intervention (defined as attending at least half of the sessions). Many of those attended all or almost all sessions. The entire sample ($n=424$) was followed throughout the first phase of the study (even if they moved).

Despite this, we were able to locate 348 of the 424 children for *SAFEChildren* II, a booster intervention study that took place in 4th-6th grades. Children who had moved from their original communities were not eligible for *SAFEChildren* II because the interventions involved multiple family groups of people living in the same geographic area. The *SAFEChildren* II study included four assessments (waves 6 - 9): pre-test, mid-intervention, post-test, and 6- and 12-month follow-up assessments. For the long-term follow-up study, all 424 families were eligible to participate, regardless of their residence location at the time of data collection, or their eligibility for *SAFEChildren* II. Most mobility in the sample was within the communities included in the study. Of the 424 families contacted for recruitment, 393 had not previously refused further participation in *SAFEChildren*, and were considered available for recruitment to this study. Three-hundred twenty-three youth (323, 76.2% of eligible participants) and one parent of 327 youth (77.1% of eligible participants) consented and completed assessments at either Wave 10 or 11 of the present study.

Figure 7 is a CONSORT chart detailing subject recruitment, assignment, and retention through all phases of the *SAFEChildren* studies. As can be seen in Figure 7, 318 of the original 424 children and families (75.0%) had sufficient data for inclusion in the follow-up study (parent and youth data). The sample was nearly half female (47.5%) and that percentage did not differ significantly from the 424 subjects in the *SAFE Efficacy Trial* study, $\chi^2(1) < 1$, *ns*. Similarly the ethnic distribution of the *SAFE*

Follow-up Study sample did not differ significantly from the ethnic distribution of the *SAFE Efficacy Trial* sample (45.3% African-American in the *SAFE Follow-up Study*, 42.5% African American in *SAFE Efficacy Trial*, $\chi^2(1) < 1$, *ns.*). The *SAFEIII* sample did not differ significantly from those not included on pre-existing family risk, $\chi^2(1) = 0.02$, *ns.*, or pre-existing externalizing risk, $\chi^2(1) = 0.56$, *ns.*

Relationship status (married or living as if married) of the parents changed considerably during the *SAFE* studies. In order to incorporate accurate information in family structure that might impact the analysis of long-term effects, we fit a growth mixture model (Muthén & Muthén, 2000) of partner status at multiple waves of measurement. The best fitting model was a four class solution. Close to half were married or living as married consistently throughout the 12 years of the *SAFE*Children studies (43.4%), and over one-fourth (27.7%) were not partnered during the same 12 years. Fewer partnered early and separated (11.6%) than those who partnered later in the study years (17.3%). The median and modal family income level reported was \$20,000 - \$24,999.

As can be seen in Figure 7, there was no significant difference in the proportion of those assigned to *SAFE Efficacy* and *Booster Trial* interventions who were assessed in the *SAFE Follow-up Study* between the Booster (95.9%), Initial Only (85.3%), or Control (92.0%) conditions, $\chi^2(2, N=424) = 4.09$, *ns.*

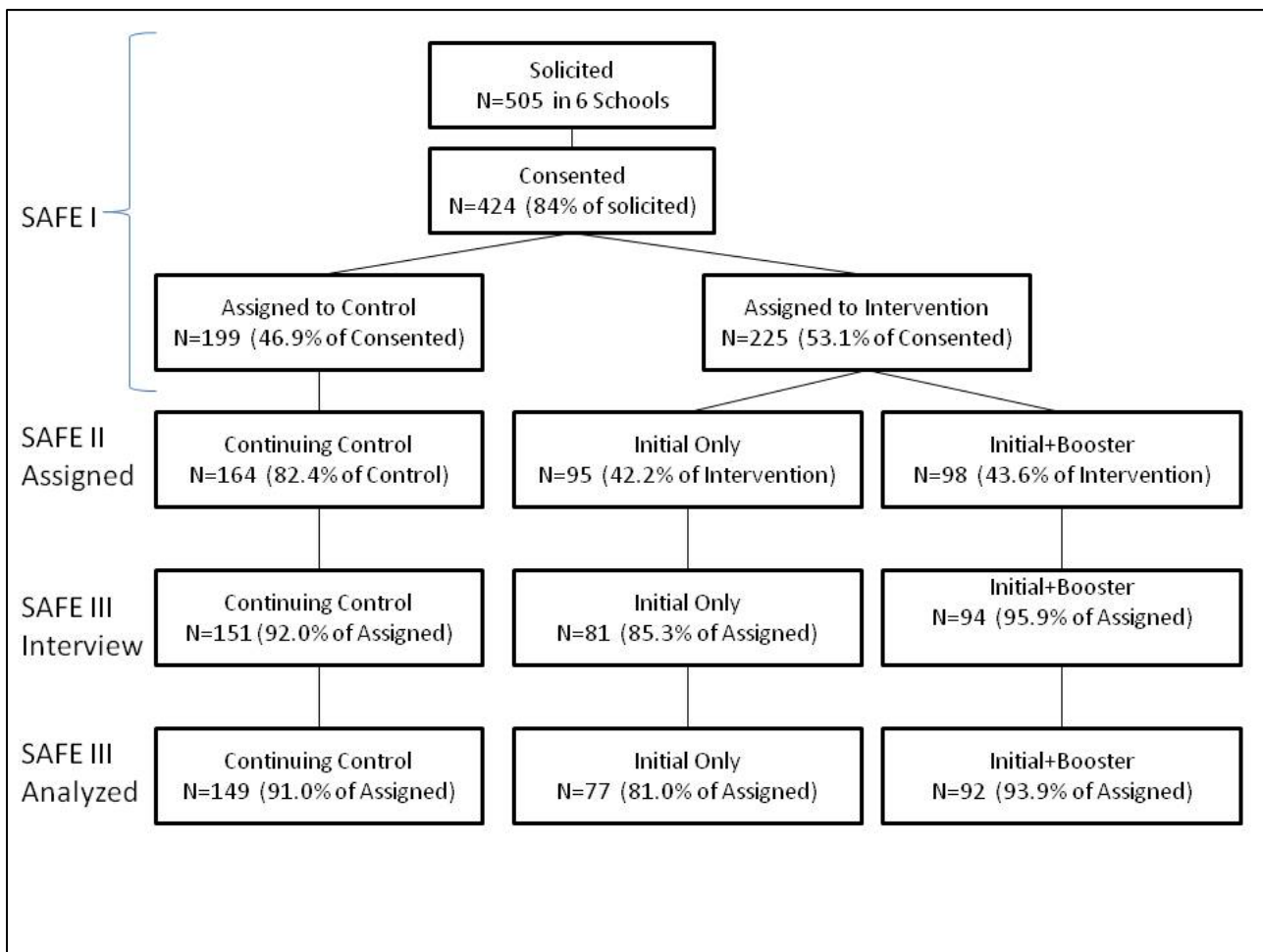


Figure 7. *SAFE Follow-up Study* CONSORT Chart

Measures

Youth externalizing. Externalizing refers to generally disruptive behaviors such as aggression. We assessed externalizing using parent and youth reports on the *Child Behavior Checklist* (CBCL; Achenbach, 1991). Specifically, we used a young adult version of this measure developed by Achenbach and colleagues (Achenbach & Rescorla, 2003) for parent and young adult reporters. The Externalizing subscale had internal consistency of 0.93 in the parent version and 0.88 in the youth version. The mean of the two versions was used as a composite score.

Youth delinquency. The Self-Report of Delinquency (SRD; Elliott, et al., 1985) provided estimates of criminal activity. We calculated a frequency score weighted by legal seriousness for this study. This ordinal scale has five levels, ranging from no involvement through minor status offenses and misdemeanors (2-3) to multiple serious misdemeanors and felonies (4-5). We refer to it as the Elliott Level as it was first suggested by Elliott and colleagues (Elliott, Huizinga & Ageton, 1985). In the Chicago Youth Development Study (Tolan, Gorman-Smith, & Henry, 2003) this score correlated positively and significantly with concurrent mother's reports of adolescent delinquent involvement ($r=0.38, p<0.01$) and with official police records of arrests ($r=0.19, p<0.01$).

Youth depression. Depression was assessed using the Center for Epidemiologic Studies Depression Scale (Radloff, 1977). This 20-item scale has been used in numerous studies of depression in the population. Its internal consistency in a general population sample was 0.85.

Youth intimate partner violence. Perpetration and victimization of relationship violence were assessed using the The Victimization and Perpetration in Dating Relationships Scales (Wolfe, Scott, Reitzel-Jaffe, & Wekerle, 2001). This measure includes 26 items assessing physical ($\alpha=0.87$) and psychological ($\alpha=0.82$) victimization, and 24 items tapping physical ($\alpha=0.89$) and psychological ($\alpha=0.83$) perpetration of aggression.

Substance use. Binge drinking over the past 30 days was assessed with a single item from the Monitoring the Future Scale (Bachman, Johnston, & O'Malley, 1986), namely, "Think back over the past 30 days. How many times have you had five or more drinks in a row of any type of alcoholic beverage?" Responses were on a scale ranging from "None" to "10 or more times." Tobacco use over the past 30 days also was assessed with the Monitoring the Future Scale, "During the PAST 30 DAYS, how frequently have you smoked cigarettes (used tobacco)?" Responses were "Not at all" to "Two packs or more per day." Finally, we created a binary variable representing any use of alcohol, tobacco, or other drugs from the Monitoring the Future scale.

Sexual risk. Items adapted from the National Longitudinal Study of Adolescent Health (Add Health; Harris & Udry, 2011) were used to assess sexual risk. These items assessed sex without a condom and sex with multiple partners. These items were combined into a single binary variable assessing any high risk sex (1) vs. not sexually active or low risk sex only (0).

Pre-existing Risk. Externalizing risk and family risk existing at the baseline of *SAFE Efficacy Trial* were taken from the *SAFE Efficacy Trial* data (Tolan, Gorman-Smith, & Henry, 2005) and used as moderator variables in this study.

Highest Grade Completed, School Withdrawal, and High School Graduation. Highest Grade Completed and School Withdrawal were assessed in the child interview. School withdrawal was coded if the child had not attended school in two years by age 18. High School Graduation was derived from archival records provided by the Department of Research and Evaluation of the Chicago Public Schools. Three-hundred four (304) of the original 424 SAFEChildren participants had CPS archival records that indicated either that they had graduated or that they had remained in the system until the *SAFE Follow-up Study* data collection began.

Archival Misconduct Records. Data from the Chicago Public Schools on misconduct (Chicago Board of Education, 2009, 2010) was collected for all SAFEChildren participants for the period between the *SAFE Booster Trial* and *SAFE Follow-up Study* data collections. Two hundred sixty-one (261) of the original 424 participants remained in the Chicago Public Schools for the entire period. Misconduct records were obtained from data files provided to Chapin Hall at the University of Chicago by the Chicago Public Schools Bureau of Safety and Security (BSS) for each school year. The data contains information on the number and type of minor to major disciplinary incidents for individual students during each school year. Because youth can be involved in multiple incidents during the school year, multiple records can exist for each student. We extracted four types of variables for analysis from these records. One was an indicator variable for serious misconduct (serious misdemeanors/felonies). Included were violent offenses (simple and aggravated battery, simple and aggravated assault, fighting, bullying), substance use and sale, extortion, disorderly conduct, gang activity, sex crimes, intimidation, robbery, murder, and attempted murder. In addition, we created indicators for violent offenses alone, substance use offenses alone, and offenses resulting in arrest. Only 4 substance-related incidents were recorded and only 28 incidents resulting in arrest were recorded. Thus, we did not include substance-related incidents in analyses.

Procedures

In SAFEChildren III, we aimed to interview youth and at least one parent during the child's 10th grade year and once during the 12th grade year. Teams of community interviewers selected for their experience living or working in similar neighborhoods completed the assessments. Interviewers had backgrounds in social service or related fields, and were able to bridge the academic interests of the study and the practical concerns of the families in order to successfully engage families and complete interviews. Project staff contacted families by telephone or in person to arrange for an initial interview. The child and at least one parent were asked to participate. At the initial interview, staff explained the research project to the parent(s) and children, and obtained written informed consent and child assents if parents and child agreed to participate. When convenient, the first interview included collection of the first wave of data (Wave 10). For the second wave, interviews were scheduled at a convenient time and location for each family. The parent interview included questionnaires about the child's behavior, social competence, parenting practices, family relationship characteristics, peer relationships, and stressful life experiences. The youth interview included similar questionnaires, with the addition of measures of academic achievement and self-reports of delinquent involvement, substance use, and sexual behavior. These more sensitive questionnaires were administered using audio-computer assisted self-interviewing (Audio-CASI), which allowed youth to respond without the interviewer or the parent knowing the response. Less sensitive sections of the parent and child

interviews were conducted using laptop computers and computer assisted personal interviewing (CAPI), in which the interviewer read the questions from a computer screen and entered the responses. The program assured that a valid response was entered, highlighted the response for multiple choice items, and managed interview branching logic (i.e., skip rules). Parent interviews took approximately 1½ hours to complete, and youth interviews required approximately one hour. Each participating family was paid \$50 for each interview or \$100 for completing both interviews. During weekly supervision, the assessment coordinator collected all completed interviews and verified completion. Twenty-five percent (25%) of interviewees were contacted by the assessment coordinator to ensure that the interviewers maintained fidelity to the interview protocol.

Analytic Approach

Mixed effects regression models were fit for outcomes on which there were numeric measures and multiple waves. These intercept-only models used both waves of SAFE-III data to estimate the outcome. Binary logistic regression models were used for binary variables such as substance use class and high risk sex.

The analyses proceeded in three major phases. First, the Type III effects of the intervention on each outcome were estimated. Separate models were fit for overall effects and for moderation by pre-existing externalizing risk, pre-existing family risk, and gender. Next, for each significant Type III effect, planned comparisons determined whether the significant effect was between levels of the intervention. Finally, for those effects with significant Type III tests whose differences were between levels of the intervention, we examined the pattern of means to interpret the effects.

Each model included terms for multiple covariates, including youth gender, ethnicity, parent partner trajectory, and parental stress during the *SAFE Efficacy Trial* intervention. Two Helmert contrasts represented Initial intervention vs. Control, and Booster vs. Initial intervention only (nested within the *SAFE Efficacy Trial* intervention), as was done in the analyses of the *SAFE Booster Trial* study (Tolan, et al., 2005; Tolan, et al., 2009):

	Contrast 1	Contrast 2
Control	-1	0
Initial Only	.5	-1
Initial + Booster	.5	1

Results

Type III Significance Tests

As is indicated above, the first phase of the analysis involved evaluating the Type III hypothesis tests for effects of the initial and booster SAFE interventions. Separate models were fit to test the overall effects and the effects moderated by family risk status and individual child externalizing risk status existing in Wave 1 of SAFE, as well as by gender. Table 17 reports the Type III hypothesis tests. As can be seen there, there were 5 significant Type III tests. No overall effects were significant. In the moderated analysis by family risk, there were significant Type III tests for the Booster vs. Initial Only

	Overall		By Family Risk		By Externalizing Risk		By Gender	
	Initial vs control	Booster vs Initial	Initial vs control	Booster vs Initial	Initial vs control	Booster vs Initial	Initial vs control	Booster vs Initial
Combined Externalizing	2.9	1.16	0.6	7.68**	1.87	0.04	0.81	0.26
Elliot Level	0.02	2.32	0.43	5.28**	1.02	0.48	0.38	0.77
Depression	0.09	0.14	0	0.26	0.61	0.37	1.22	1.59
IPV Victimization	1.85	2.39	0.02	1.76	0.29	0.1	0.17	2.18
IPV Perpetration	0.37	3.07	0.15	0.19	0	0.57	0.03	0.82
Highest Grade Completed	0.14	0.51	0.19	0.01	0	0.11	0.05	7.73**
Age of Sexual Debut	0.02	0.09	0.49	3.49+	3.32	0.08	0.46	0.88
30-day Binge Drinking	1.56	1.45	0.52	0.86	0.2	0.76	0.89	4.8
30-Day Tobacco Use	1.15	1.35	0.72	0.59	0.08	0.04	1.5	0.03
Sexual risk	0.57	0.78	0.29	12.3**	2.62	0.07	0.09	0.17
Any Substance Use	0	0.54	1.18	1.15	1	0.06	0.29	5.62**

* p < .05. ** p < .01

Note: Type III tests are F tests with 2 numerator and 315 denominator degrees of freedom.

effect on Combined Externalizing, Elliot Level of Delinquency, and Sexual Risk. Within the analysis moderated by gender, there were significant Type III tests for Highest Grade Completed and for Any Substance Use.

Contrast Results

The Type III tests only suggest that there may be a significant comparison within an overall effect. The next step was to evaluate the comparisons of interest (Any Intervention vs. Control and Booster vs. Initial Only) within levels of the moderators. These results are shown in Tables 18-20.

The Booster vs. Initial Only contrast was significant for Combined Externalizing within families at low pre-existing risk, and for Elliott Level of Delinquency among families at both high and low levels of pre-existing externalizing. As can be seen in Table 18, this effect appears to have been somewhat stronger among those at high pre-existing family risk, which is, no doubt, responsible for the significant interaction between family risk and Booster vs. Initial intervention for Elliott Level of Delinquency. The contrasts testing intervention effects on Sexual Risk were not significant.

As can be seen in Table 18, the significant moderated effect on highest grade completed was the result of a significant difference between Booster and Initial Only participants. Those assigned to the Booster intervention, on average, completed approximately one third of a year less schooling than those assigned to the Initial Only condition. The contrasts testing intervention effects on Any Substance Use did not indicate significant mean differences.

Table 18

Estimated Means and Contrasts, By SAFE Efficacy Trial Family Risk Levels

	Low Family Risk						High Family Risk					
	Control		Initial		Booster		Control		Initial		Booster	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Combined												
Externalizing	8.78	0.58	11.00 ^a	0.88	6.57 ^b	0.88	9.45	0.74	11.66	1.16	7.23	1.12
Elliott Level	1.44	0.14	2.07 ^a	0.22	0.81 ^b	0.23	1.32	0.16	1.95 ^a	0.29	0.69 ^b	0.28
Depression	11.68	0.81	10.93	1.27	12.42	1.28	11.00	0.95	10.26	1.65	11.75	1.60
IPV Victimization	2.12	0.46	1.92	0.75	2.31	0.72	2.34	0.53	2.14	0.97	2.54	0.91
IPV Perpetration	1.28	0.43	0.13	0.70	2.43	0.67	1.75	0.50	0.60	0.90	2.90	0.85
Highest Grade	12.03	0.08	12.00	0.11	12.06	0.11	11.72	0.11	11.69	0.15	11.75	0.14
Age of Sexual debut	15.88	0.15	15.72	0.21	16.04	0.20	15.93	0.20	15.77	0.29	16.09	0.27
30-day Binge												
Drinking Frequency	0.35	0.08	0.30	0.12	0.40	0.11	0.32	0.11	0.27	0.16	0.37	0.15
30-day Tobacco Use												
Frequency	0.18	0.08	0.20	0.11	0.17	0.11	0.22	0.11	0.24	0.15	0.21	0.15
Sexual Risk	0.23	0.20	0.26	0.28	0.20	0.27	0.23	0.20	0.27	0.27	0.20	0.26
Substance User	0.17	0.20	0.19	0.30	0.16	0.27	0.20	0.27	0.21	0.42	0.19	0.39

Note: Within rows and levels of family risk, means with different superscripts are significantly different at $p < .01$.

Table 19

Estimated Means and Contrasts, By SAFE Efficacy Trial Externalizing Risk Levels

	Low						High					
	Control		Initial		Booster		Control		Initial		Booster	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Combined Externalizing Elliot Level	9.01	0.6	9.49	0.93	8.53	0.91	12.57	1.19	13.05	1.39	12.10	1.37
Depression	1.38	0.1	1.68	0.241	1.09	0.23	1.62	0.24	1.91	0.31	1.32	0.30
IPV Victimization	11.42	0.83	11.20	1.36	11.63	1.33	11.74	1.41	11.52	1.77	11.95	1.75
IPV Perpetration	2.10	0.48	1.17	0.81	3.03	0.75	1.45	0.77	0.53	1.00	2.38	0.96
Highest Grade	1.44	0.44	0.35	0.75	2.53	0.69	0.93	0.72	-0.16	0.94	2.02	0.89
Age of Sexual debut	11.90	0.08	11.87	0.12	11.94	0.11	11.68	0.18	11.64	0.21	11.71	0.20
30-day Binge Drinking Frequency	15.91	0.15	15.96	0.22	15.85	0.21	15.57	0.35	15.63	0.39	15.51	0.38
30-day Tobacco Use Frequency	0.33	0.08	0.25	0.12	0.42	0.12	0.68	0.20	0.59	0.22	0.77	0.22
Sexual Risk	0.21	0.08	0.28	0.12	0.13	0.11	0.56	0.20	0.63	0.22	0.48	0.21
Substance User	0.23	0.15	0.20	0.21	0.27	0.20	0.25	0.31	0.21	0.35	0.29	0.34
	0.18	0.22	0.17	0.33	0.20	0.27	0.19	0.50	0.18	0.56	0.21	0.53

Note: Within rows and levels of externalizing risk, means with different superscripts are significantly different at $p < .01$.

Table 20

Estimated Means and Contrasts, By Gender

	Female						Male					
	Control		Initial		Booster		Control		Initial		Booster	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Combined Externalizing	8.69	0.66	9.10	1.10	8.28	1.13	10.09	0.70	10.50	1.14	9.68	1.13
Elliot Level	1.02	0.15	1.43	0.28	0.61	0.29	1.77	0.16	2.18	0.29	1.36	0.29
Depression	11.79	0.88	10.26	1.58	13.31	1.63	10.80	0.92	9.28	1.62	12.33	1.63
IPV Victimization	1.42	0.50	1.32	0.92	1.52	0.91	2.64	0.52	2.54	0.94	2.74	0.92
IPV Perpetration	1.66	0.47	0.29	0.86	3.03	0.85	1.14	0.49	-0.23	0.88	2.51	0.86
Highest Grade	12.06	0.09	12.20	0.14	11.92	0.14	11.68	0.10	11.82 ^a	0.14	11.54 ^b	0.14
Age of Sexual debut	16.65	0.18	16.57	0.26	16.73	0.27	14.99	0.19	14.90	0.28	15.07	0.27
30-day Binge Drinking												
Frequency	0.24	0.10	0.30	0.15	0.17	0.15	0.52	0.11	0.58	0.16	0.46	0.15
30-day Tobacco Use												
Frequency	0.08	0.09	0.18	0.14	-0.03	0.14	0.43	0.10	0.54	0.15	0.33	0.15
Sexual Risk	0.89	0.28	0.90	0.43	0.87	0.42	0.55	0.22	0.59	0.39	0.52	0.38
Substance User	0.86	0.26	0.80	0.37	0.90	0.44	0.75	0.26	0.67	0.41	0.82	0.40

Note: Within rows and levels of gender, means with different superscripts are significantly different at $p < .01$.

Low Base Rate Comparisons

Archival School Records of Misconduct. Overall, there were 133 instances of serious misconduct recorded over five school years, committed by 60 individuals. Ninety-four incidents were violent in nature. Arrests were made in 28 instances. Using the counts of serious misconduct and violent misconduct, we fit generalized linear models using a negative binomial distribution, which is appropriate for sparse count data. The results of analysis by random assignment condition are reported in Table 21. As can be seen there, the effect of the first Helmert contrast testing the effect of any SAFE intervention vs. no SAFE intervention was significant for serious misconduct. The same effect was marginally significant for the violent offenses. The difference between any SAFE intervention and controls on arrests was not significant, although there was a marginal tendency for those assigned to the booster intervention to have been more likely to be arrested.

High School Graduation and School Withdrawal. Of those with existing archival records, 60.5% were listed as graduated from high school. A generalized linear model of high school graduate status (1, 0) including only those with CPS archival records until the time of the *SAFE Follow-up Study* found non-significant differences in high school graduation rates between those assigned to any SAFEChildren intervention and those assigned to control ($B=-.30$, $SE=.21$, $\chi^2(1, N=246)=2.0$, $p = .16$), and between those assigned to Booster and those assigned to Initial Only ($B=-.14$, $SE=.20$, $\chi^2(1, N=246) = .45$, $p = .50$).

Logistic regression and survival analysis (treating the cases that had not yet reached age 18 as censored) were attempted with the self-report data on school withdrawal, but these models did not converge because of the sparseness of the outcome variable. Thus, we used a simple series of cross-tabulations with Fisher Exact Tests (Fisher, 1922) to evaluate the effects on school withdrawal. The Fisher Exact Test tests for association between the rows and columns of a table, in this case, between levels of each of the two Helmert contrasts used to indicate intervention levels and the binary variable indicating withdrawal from school. The Fisher Exact Test is appropriate for use with very sparse data such as this, providing exact probabilities even when one or more cells contain fewer than five cases.

Table 22 reports the results for true school withdrawals from self-reports. We report the numbers of true withdrawals for each condition overall, and within subgroups defined by family risk, eternalizing risk, and gender. As can be seen there, the Fisher Exact Tests indicated significant differences for Booster + Initial vs. Initial Intervention Only overall, and within low and high family risk and for males. In each case save one, youth who had been assigned to receive the booster intervention had lower rates of school withdrawal than those who were not assigned as such. The single exception was within high family risk, where there were no withdrawals for Initial Intervention Only and one for Booster + Initial.

Table 21

SAFE Children Effects on Archival School Reports of Misconduct

Outcome	<i>p</i> (at least one)			Any SAFE vs. No		Booster vs. Initial	
	Control	Initial Only	Initial + Booster	SAFE		SAFE	
				B	SE	B	SE
Serious Misconduct	.16 (.03)	.12 (.03)	.13 (.04)	-.46**	.16	.05	.18
Violence	.14 (.02)	.07 (.02)	.12 (.04)	-.34+	.18	.22	.20
Arrests	.03 (.01)	.04 (.02)	.07 (.03)	.18	.28	.48+	.27

+ $p < .10$. * $p < .05$. ** $p < .01$.

Table 22

Results for School Withdrawal by Assigned Intervention Condition, Risk, and Gender

	Control		Initial Only		Booster		Odds Ratios		Fisher Exact Tests	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	I v C	B v I	I v C	B v I
Overall	8	5.23	4	4.21	1	1.28	0.81	0.35	0.13	0.03
By Family Risk										
Low	3	3.70	4	7.55	0	0.00	2.16	0.14	0.30	0.02
High	5	7.04	0	0.00	1	3.33	0.19	10.66	0.09	0.04
By Externalizing Risk										
Low	5	4.07	3	3.80	1	1.41	1.07	0.41	0.22	0.06
High	3	10.00	1	6.67	0	0.00	0.67	0.25	0.34	0.24
By Gender										
Female	2	2.63	0	0.00	0	0.00	0.18	-	0.20	0.20
Male	6	7.79	4	9.30	1	2.86	1.22	0.38	0.23	0.05

Notes: I v C = Helmert contrast comparing controls with *SAFE Efficacy Trial* Intervention

B v I = Helmert contrast comparing Booster with initial *SAFE Efficacy Trial* Intervention only

The odds ratio for Booster vs. Initial within female could not be calculated because of zeros in both cells.

Summary of Objective 2:

The long term benefits of SAFEChildren may be summarized as follows. For those who were assigned to receive both the initial and 4th grade booster SAFE interventions, the risk of withdrawal from school was reduced to 25% of the risk among controls and 30% of the risk among those receiving only the initial intervention. Serious delinquent offenses appear also to be reduced, and school records indicate a reduction in violent offenses to 70% of the level reported for those assigned to the control condition.

Regarding monetizing the benefits potentially associated with long-term effects of the SAFEChildren program, the typical “high risk” youth with 6+ police contacts (who collectively commit about 50% of all crimes), imposes between \$4.2 and \$7.2 million in costs. Discounted to present value at age 14, costs total \$3.2–\$5.8 million. The bulk of these costs (\$2.7 million–\$ 4.8 million) are due to crimes, while an additional \$390,000–\$580,000 is estimated to be the value of lost productivity due to dropping out of high school. Those juveniles that become career criminals impose \$65,000 in costs through age 12 and \$230,000 through age 14. Cohen admits that further research is needed to understand the full lifetime costs of crime – including lost freedom for the incarcerated offender, child development concerns for children with incarcerated parents, and lifetime earning capacity (Cohen, Piquero, & Jennings, 2010; Cohen & Piquero, 2009).

Objective 3. To estimate the per child expenditure that would be justified by a long-term effect as large as that obtained in the SAFE-III follow-up study of the original SAFE-I and II interventions.

Monetizing the Costs of Outcomes

The *ex ante* methods of monetizing the potential social costs of crime, such as “willingness-to-pay” (WTP) and “contingent valuation” (CV) methodologies were developed in the environmental economics literature. These methods assess the public’s perception of value of issues or resources that may not be amenable to direct market valuation.

Cohen and colleagues’ (2010; 2009) approach to the WTP method was to conduct a survey asking community respondents if they would be willing to vote for a proposal requiring each household in their community to pay a certain amount to be used to prevent one in ten incidents of certain crimes in their communities. Respondents were shown three different crimes, including burglary, serious assault, armed robbery, rape or sexual assault and murder. Cohen’s group looked at numerous crimes and the specific cost to each crime, and concluded that the typical “high risk” youth with 6+ police contacts (who collectively commit about 50% of all crimes), imposes between \$4.2 and \$7.2 million in costs on society. Discounted to present value at age 14, these costs total \$3.2–\$5.8 million. The bulk of these costs (\$2.7 million–\$ 4.8 million) are due to crimes, while an additional \$390,000–\$580,000 is estimated to be the value of lost productivity due to dropping out of high school. Juveniles that become career criminals imposes \$65,000 in costs through age 12 and \$230,000 through age 14.

Estimating the Potential Benefits of SAFE

Although the social costs of crime are not fully understood at present, for the purposes of this analysis, we are able to use Cohen’s (2010) estimates along with the *SAFE Follow-up Study* results around school reported misconduct and true withdrawal from high school (i.e., not attending alternative school or GED classes). Table 23 below reports the change in probability of a single offense, becoming a career criminal, and dropping out of high school, the upper and lower limits of estimates of social costs of each outcome, the expected lifetime change in the social costs, and,

combined with our estimates of the costs of providing initial and booster SAFEChildren interventions, the potential benefit per child of the SAFE intervention.

Cohen and colleagues (2010) provide estimates of the costs to society of single criminal offenses and crime careers, their future value discounted to produce a lifetime social cost, which is reported in Table 23. Cohen and Piquero (2009) estimated that 4% of single offenders would become career criminals. We multiplied that by the 4% reduction in single offenses we found for the *SAFE Follow-up Study*. We believe that this procedure produces a conservative estimate of the potential cost/benefit of SAFEChildren, because it is likely that schools did not always code a uniform discipline code offense for first offenders. It is possible that our estimates of reductions in single offenses are actually estimates of reductions in repeat offenders.

Putting together the monetized outcomes found to be significant in the *SAFE Follow-up Study* with the actual costs of delivering the intervention, we produced the values in Table 23. As can be seen in Table 23, our most conservative estimate of the lifetime potential benefit of providing initial and booster SAFEChildren interventions is \$5.19 per dollar of intervention cost.

Table 23

Potential lifetime cost/benefit of the SAFEChildren Intervention

Behavior or outcome	Change in probability	Discounted Lifetime Cost		Expected change in lifetime social costs		Totals
		LL	UL	LL	UL	
Single criminal offense	0.04	\$53,000	\$78,500	\$2,120	\$3,140	
Career criminal ^a	0.0016	\$1,800,000	\$3,000,000	\$2,880	\$4,800	
High school dropout	0.039	\$200,000	\$450,000	\$7,800	\$17,550	
Lower Limit total lifetime benefits per dollar of intervention costs per child ^b						5.19
Upper limit total lifetime benefits per dollar of intervention costs per child ^b						10.34

^a Based on an estimated 4% of single offenders that will become career criminals (Cohen & Piquero, 2009)

^b Based on total costs of \$2,466.54 for delivering SAFE initial and booster interventions

V Conclusions and Implications

This document reports the results of four aspects of an investigation into the effects of a multi-component intervention aimed at modifying early markers of risk for delinquency and substance use. The first aspect was evaluation of the overall effects of the SAFEChildren intervention when the group-based family intervention was administered by community mental health providers and the reading tutoring program was provided by upper grade children from the same schools attended by the participants. We found that we were able to predict participation in the intervention with very high accuracy using propensity score methods, but despite the use of such methods we found few significant effects of the intervention, even in analyses of subgroups defined by pre-existing risk.

The second aspect was analysis of social network characteristics of the intervention groups and the relation of change in such characteristics to intervention effects. These analyses produced evidence that the intervention groups had functioned as we hoped, namely reducing parents' sense of isolation and creating opportunities for social support related to parenting. We saw evidence for

change in group characteristics such as network density and evidence for change in individual reports of the likelihood of making use of other group members for advice later on.

The third aspect was examination of process characteristics, therapist and tutor attitudes, and their relations to outcomes. In these analyses, careful examination of process and provider measures found no pronounced effects of differences in provider characteristics that could predict outcomes. Although groups varied in average fidelity of implementation to a small but significant extent, participants and providers largely agreed that fidelity was uniformly very high. Despite some variation and one very modest association with outcome, evaluation of tutor performance was generally positive. There were too few therapists for reliable assessment of change in attitudes over time. However, analysis of the effects of pre-existing therapist attitudes found no evidence that they in any way impeded implementation of the intervention.

The fourth and final aspect of this research was estimation of the potential costs and benefits associated with the SAFEChildren Program. Because of the modest effects of the *SAFE Effectiveness Trial*, we made use of the results of the *SAFE Follow-up Study* results which found significant long-term effects for delinquency and aggression, as well as long-term evidence from archival records for effects on disciplinary referrals for disruptive and violent behavior at school. The estimates of costs were gathered from the actual costs of implementing the SAFE Effectiveness Trial intervention, less the research-related costs.

The results lead to a number of conclusions, questions, and directions for future research and for efforts to implement this type of family-focused prevention program on a large scale. First, the results indicate that the effects of a family-focused preventive intervention such as SAFEChildren may be less pronounced outside of the careful, close, supervision that characterizes implementation in the context of research. A prominent example of this is the implementation of makeup sessions for families who missed regular intervention sessions. When SAFEChildren was implemented by university personnel (*SAFE Efficacy Trial* and *SAFE Booster Trial*), considerable effort was made to provide makeup sessions to families who did not attend group sessions. Makeup sessions were also provided in another large-scale trial of a family intervention that closely resembled SAFEChildren, namely the Multisite Violence Prevention Project, which found significant effects of the intervention on measures of violence (Multisite Violence Prevention Project, 2009). Anecdotal evidence suggests that the mental health providers who delivered the family intervention in the *SAFE Effectiveness Trial*, despite considerable enthusiasm about the intervention itself, were not, as a rule, favorably disposed to visiting clients in their homes for makeup sessions. As a result, few makeup sessions were delivered, and none were delivered at the subject's homes.

Thus, we believe that an important implication of this research is to emphasize the centrality of home visitation in the delivery of SAFEChildren and other similar interventions. Although it is impossible to determine from the *SAFE Effectiveness Trial* that home visitation for purposes of delivering intervention content in makeup sessions was the deciding factor in intervention effectiveness, the absence of home visitation in the *SAFE Effectiveness Trial* was a remarkable difference between it and the earlier trials. Moreover, the notion that home visitation is pivotal to effects of an intervention such as SAFEChildren is consistent with substantial literature finding benefits of home visitation programs in the development of younger children in high risk settings

(Farrington & Welsh, 2007), and the current inclusion of such programs in the national budget and on rosters of effective evidence-based interventions (Center for the Study and Prevention of Violence, 1999; Coalition for Evidence-based Policy, 2007).

As has been the case in prior SAFEChildren studies and similar interventions, process measures indicate that the intervention can be delivered with high fidelity by community providers. With the exception of makeup sessions, as is noted above, measures of fidelity, alliance, and satisfaction uniformly suggest that the characteristics associated with positive intervention impact in the literature, such as a strong initial alliance and satisfaction with the provider and the intervention were present in the *SAFE Effectiveness Trial*. Tutoring process measures are consistent with high fidelity of tutoring program implementation by the seventh and eighth grade tutors. Both supervisor and tutor reports indicate that the program was delivered as intended.

Measures of social network processes within the groups suggest that group members increasingly came to rely on each other for parenting advice, and to incorporate each other into social support systems. These findings are consistent with the intent that the groups become sources of social support for members.

One of the central questions of this investigation concerned the extent to which it would be possible for community providers to deliver all components of the SAFEChildren intervention in a manner that is faithful to the intervention design and effective in changing key outcomes in early childhood. The foregoing evidence suggests that it is possible to deliver the SAFEChildren intervention with community providers, but that it may be difficult to acclimate mental health service providers to the assertive recruiting, follow-up, and provision of makeup sessions through home visits that characterizes successful preventive interventions.

The SAFE-E intervention was offered as a universal intervention to all families. However it is possible to conceive of SAFEChildren as a targeted intervention that would be offered to families with evidence of increased risk or whose children show early indications of risk. The analyses predicting participation in SAFE-E found no effects for child aggression but some evidence that better organized families in communities with more resources were more likely to participate fully in the groups. These findings suggest that offering SAFEChildren as a targeted intervention (or as a universal intervention) would require the expenditure of effort to recruit and maintain involvement in families for whom family organization deficits or neighborhood constraints form barriers to participation.

A final implication stems from the cost-benefit analysis. The social costs of youth risk are high enough that an expenditure equal to the cost of providing SAFEChildren universally is likely to be worthwhile. We conservatively estimate that it is possible to obtain at least five times the cost of providing SAFEChildren in societal benefits. However, such benefits will only be obtained if the intervention can be administered with levels of supervision, recruitment, and retention effort similar to those expended in the *SAFE Efficacy Trial*.

References

- Achenbach, T. M. (1991). *Integrative Guide to the 1991 CBCL/4-18, YSR, and TRF Profiles*. Burlington, VT: University of Vermont.
- Achenbach, T. M. & Rescorla, L. A. (2003). *Manual for the ASEBA Adult Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Aday, L. A. (1993). *At risk in America: The health and health care needs of vulnerable populations in the United States*. San Francisco, CA: Jossey-Bass.
- Alexander, J. F. & Parsons, B. V. (1982). *Functional family therapy*. Monterey, CA: Brooks/Cole.
- Aos, S., Lieb, R., Mayfield, J., Miller, M. & Pennucci, A. (2004). Benefits and costs of prevention and early intervention programs for youth.
. Olympia, WA: Washington State Institute for Public Policy.
- Bachman, J. G., Johnston, L. D. & O'Malley, P. M. (1986). Recent findings from Monitoring the Future: A continuing study of the lifestyles and values of youth. In F. M. Andrews (Ed.), *Research on the quality of life* (pp. 215-234). Ann Arbor, MI: Institute for Social Research.
- Bock, R. D. (1983). Within-subject experimentation in psychiatric research. In R. D. Gibbons & M. W. Dysken (Eds.), *Statistical and methodological advances in psychiatric research* (pp. 59-90). New York, NY: Spectrum Books.
- Bock, R. D. (1989). Measurement of human variation: A two-stage model. In R. D. Bock (Ed.), *Multilevel analysis of educational data* (pp. 319-342). Orlando, FL: Academic Press.
- Bronfenbrenner, U. (1979). Contexts of child rearing: Problems and prospects. *American Psychologist*, 34, 844-850.
- Bronfenbrenner, U. (1988). Interacting systems in human development. research paradigms: Present and future. In A. C. N. Bolger, G. Downey & M. Moorehouse (Ed.), *Persons in context: Development processes* (pp. 25-49). New York, NY: Cambridge University Press.
- Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K. & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology*, 99, 353-395.
- Catalano, R. F., Arthur, M. W., Hawkins, J. D., Berglund, L. & Olson, J. J. (1998). Comprehensive community- and school-based interventions to prevent antisocial behavior. In R. Loeber & D. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks, CA: Sage Publications. .
- Center for the Study and Prevention of Violence. (1999). Nurse-Family Partnership. *Blueprints for Violence Prevention* Retrieved January 31, 2012
- Chicago Board of Education. (2009). *Student Code of Conduct*. Chicago, IL: Chicago Public Schools.
- Chicago Board of Education. (2010). *Student Code of Conduct for the 2010-2011 School Year*. Chicago, IL: Chicago Public Schools.
- Coalition for Evidence-based Policy. (2007). Nurse-family Partnership. *Social Programs that Work* Retrieved 01/31/2012, 2007, from <http://www.evidencebasedprograms.org/Default.aspx?tabid=154>
- Cohen, M., Piquero, A. & Jennings, W. (2010). Estimating the costs of bad outcomes for at-risk youth and the benefits of early childhood interventions to reduce them. *Criminal Justice Policy Review*, 21(4), 391-434.
- Cohen, M. A. & Piquero, A. R. (2009). New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, 25, 25-49.
- Coie, J. D. & Krehbiel, G. (1984). Effects of academic tutoring on the social status of low-achieving, socially rejected children. *Child Development*, 55, 1465-1478.
- Crane, J. (1991). The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American Journal of Sociology*, 96(5), 1226-1259.

- Dishion, T. J. & McMahon, R. J. (1998). Parental monitoring and the prevention of child and adolescent problem behavior: A conceptual and empirical formulation. *Clinical Child and Family Psychology Review*, 1(1), 68-75.
- Elliott, D., Huizinga, D. & Ageton, S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage.
- Elliott, D. & Tolan, P. H. (1999). Youth, violence prevention, intervention and social policy: An overview. In D. Flannery & R. Hoff (Eds.), *Youth violence: A volume in the psychiatric clinics of North America* (pp. 3-46). Washington, DC: American Psychiatric Press, Inc.
- Farrington, D. P. & Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1(1), 9-38.
- Farrington, D. P. & Welsh, B. C. (2007). *Saving children from a life of crime: Early risk factors and effective interventions*. New York, NY, US: Oxford University Press.
- Fisher, R. A. (1922). On the interpretation of chi-square from contingency tables, and the calculation of p. *Journal of the Royal Statistical Society*, 85(1), 87-94.
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S. (2005). Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*, 6(3), 151-175.
- Forgatch, M. S., Patterson, G. R. & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of Parent Management Training. *Behavior Therapy*, 36(1), 3-13.
- Furstenberg, F. (1993). How families manage risk and opportunity in dangerous neighborhoods. In W. J. Wilson (Ed.), *Sociology and the Public Agenda* (pp. 231-258). Newbury Park, CA: Sage Publications.
- Gibbons, R. D., Hedeker, D., Elkin, I., Waternaux, C., Kraemer, H. C., Greenhouse, J. B. (1993). Some conceptual and statistical issues in analysis of longitudinal psychiatric data. Application to the NIMH treatment of Depression Collaborative Research Program dataset. *Arch Gen Psychiatry*, 50(9), 739-750.
- Gorman-Smith, D., Tolan, P. H. & Henry, D. (2000). A developmental-ecological model of the relation of family functioning to patterns of delinquency. *Journal of Quantitative Criminology*, 16, 169-198.
- Harris, K. M. & Udry, J. R. (2011). National Longitudinal Study of Adolescent Health (Add Health), 1994-2008 ICPSR21600-v4 (Computer file) (Publication no. 10.3886/ICPSR21600). Retrieved July 6, 2011, from Ann Arbor, MI: Inter-university Consortium for Political and Social Research
- Hawkins, J. D. & Weis, J. G. (1985). The social development model: An integrated approach to delinquency prevention. *The Journal of Primary Prevention*, 6(2), 73-97.
- Hedeker, D. & Gibbons, R. D. (2006). *Longitudinal Data Analysis (Wiley Series in Probability and Statistics)*. New York, NY: Wiley Blackwell.
- Henggeler, S. W. (2001). Multisystemic therapy. *Residential Treatment for Children & Youth. Special Issue: Innovative mental health interventions for children: Programs that work*, 18(3), 75-85.
- Henry, D., Tolan, P. H. & Gorman-Smith, D. (2001). Longitudinal family and peer group effects on violence and non-violent delinquency. *Journal of Clinical Child Psychology*, 30, 172-186.
- Kellam, S. G. & Rebok, G. W. (1992). Building developmental and etiological theory through epidemiologically based preventive intervention trials. In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 162-195). New York, NY: Guilford Press.
- Lipsey, M. & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.
- Lochman, J. E. & Wells, K. C. (2002). The Coping Power program at the middle-school transition: Universal and indicated prevention effects. *Psychology of Addictive Behaviors, Vol 16*((4, Suppl)), S40-S54.
- Mason, C. A., Cauce, A. M., Gonzales, N. & Hiraga, Y. (1996). Neither too sweet nor too sour: Problem peers, maternal control, and problem behavior in African American adolescents. *Child Development*, 67, 2115-2130.

- McKay, M. M., Gonzales, J., Quintana, E., Kim, L. & Abdul-Adil, J. (1995). Multiple family groups: An alternative for reducing disruptive behavioral difficulties of urban children. *Research on Social Work Practice, 9*(5), 593-607.
- McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61*, 311-346.
- Metropolitan Area Child Study Research Group [MACS]. (2002). A cognitive-ecological approach to preventing aggression in urban settings: Initial outcomes for high risk children. *Journal of Consulting and Clinical Psychology, 70*, 179-194.
- Miller-Johnson, S., Sullivan, T. N., Simon, T. R. & The Multisite Violence Prevention Project. (2004). Evaluating the impact of interventions in the Multisite Violence Prevention Study: Samples, procedures, and measures. *American Journal of Preventive Medicine, 26*, 48-61.
- Multisite Violence Prevention Project. (2009). The ecological effects of universal and selective violence prevention programs for middle school students: A randomized trial. *Journal of Consulting and Clinical Psychology, 77*(3), 526-542.
- Multisite Violence Prevention Project (MVPP). (2004). The Multisite Violence Prevention Project: Background and overview. *American Journal of Preventive Medicine, 26*, 3-11.
- Muthén, B. & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research, 24*, 882-891.
- Patterson, G. R., DeGarmo, D. & Forgatch, M. S. (2004). Systematic Changes in Families Following Prevention Trials. *Journal of Abnormal Child Psychology, 32*(6), 621-633.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385-401.
- Richters, J. E. & Martinez, P. (1993). The NIMH community violence project: I. Children as victims of and witnesses to violence. *Psychiatry, 56*, 7-21.
- Rosenbaum, P. A. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.
- Sampson, R. J. & Laub, J. H. (1994). Urban poverty and the family context of delinquency: A new look at the structure and process in a classic study. *Child Development, 65*(2), 523-539.
- Sampson, R. J., Raudenbush, S. W. & Earls, F. (1997). Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science, 277*, 918-924.
- Sandler, I., Ostrom, A., Bitner, M. J., Ayers, T., Wolchik, S. & Daniels, V.-S. (2005). Developing effective prevention services for the real world: a prevention service development model. *American Journal of Community Psychology, 35*(3/4), 127-142.
- Spoth, R. L. & Grenberg, M. T. (2005). Toward a comprehensive strategy for effective practitioner-scientist partnerships and large-scale community health and well-being. *American Journal of Community Psychology, 35*(3/4), 107-126.
- Tolan, P. & Gorman-Smith, D. (1996). Families and the development of urban children. In H. Walberg, O. Reyes & R. P. Weissberg (Eds.), *Children and youth: Interdisciplinary perspectives* (Vol. 7, pp. 67-90). Chicago, IL: The University of Illinois at Chicago.
- Tolan, P. & McKay, M. M. (1996). Preventing serious antisocial behavior in inner-city children: An empirically based family intervention program. *Family Relations, 45*, 148-155.
- Tolan, P. H. (2002). Family-focused prevention research: Tough but tender with family intervention research. In H. Liddle, J. Bray, D. Santeban & R. Levant (Eds.), *Family psychology intervention science* (pp. 197-214). Washington, D.C.: American Psychological Association.
- Tolan, P. H., Gorman-Smith, D. & Henry, D. (2001). New study focuses on efficacy of 'whole school' prevention approaches. *Emotional and Behavior Disorders in Youth, 2*, 5-6, 22.
- Tolan, P. H., Gorman-Smith, D. & Henry, D. (2003). The developmental-ecology of urban males' youth violence. *Developmental Psychology, 39*, 274-291.

- Tolan, P. H., Gorman-Smith, D. & Henry, D. (2004). Supporting Families in a High-Risk Setting: Proximal Effects of the SAFEChildren Preventive Intervention. *Journal of Consulting and Clinical Psychology, 72*(5), 855-869.
- Tolan, P. H., Gorman-Smith, D. & Henry, D. (2005). Supporting Families in a High-Risk Setting: Proximal Effects of the, SAFEChildren Preventive Intervention. *Journal of, Consulting and Clinical Psychology, 72*, 855-869.
- Tolan, P. H., Gorman-Smith, D., Henry, D. & Schoeny, M. (2009). The benefits of booster interventions: Evidence from a family-focused prevention program. *Prevention Science, 10*(4), 287-297.
- Tolan, P. H. & Guerra, N. G. (1994). Prevention of delinquency: Current status and issues. *Applied & Preventive Psychology, 3*, 251-273.
- Tolan, P. H. & Henry, D. B. (1996). Patterns of psychopathology among urban poor children: Comorbidity and aggression effects. *Journal of Consulting and Clinical Psychology, 64*, 1094–1099.
- US Department of Health and Human Services (USDHHS) Public Health Service (PHS), O. o. t. S. G. (2001). *Youth Violence: A Report of the Surgeon General*. Rockville, MD: US Department of Health and Human Services.
- US Department of Health and Human Services (USDHHS) Public Health Service (PHS); Office of the Surgeon General. (2001). *Youth Violence: A Report of the Surgeon General*. Rockville, MD: US Department of Health and Human Services.
- Wallach, M. A. & Wallach, L. (1976). *Teaching all children to read*. Chicago, IL: University of Chicago Press.
- Wasserman, G. A. & Miller, L. (1998). The prevention of serious and violent juvenile offending. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks, CA: Sage.
- Webster-Stratton, C. (1990). Stress: A potential disrupter of parent perceptions and family interactions. *Journal of Clinical Child Psychology, 19*, 302-312.
- Webster-Stratton, C. (1998). Preventing conduct problems in head start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology, 66*(5), 715-730.
- Wolfe, D. A., Scott, K., Reitzel-Jaffe, D. & Wekerle, C. (2001). Development and Validation of the Conflict in Adolescent Dating Relationships Inventory. *Psychological Assessment, 13*(2), 277-293.
- World Federation of Mental Health. (2002). *Toward a strategy for worldwide action to promote mental health and prevent mental and behavioral disorders*. Alexandria, VA: World Federation for Mental Health.