

The Strengthening Washington D.C. Families Project: A Randomized Effectiveness Trial of Family-Based Prevention

Denise Gottfredson,^{1,4} Karol Kumpfer,² Danielle Polizzi-Fox,¹ David Wilson,³
Veronica Puryear,¹ Penny Beatty,¹ and Myriam Vilmenay¹

The Strengthening Washington DC Families (SWFP) Project examined implementation fidelity and effectiveness when a selective, evidence-based prevention program was implemented with a sample of 715 predominantly African American families across multiple settings in an urban area. Using a true experimental design, this study reports on the differential effectiveness of four conditions (child skills training only, parent skills training only, parent and child skills training plus family skills training, and minimal treatment controls) in reducing child antisocial behavior and its precursors. Major challenges with recruitment and retention of participants and uneven program coverage were documented. No statistically significant positive effects for any of the program conditions were observed, and a statistically significant negative effect on child reports of Negative Peer Associations was observed for children of families assigned to the family skills training condition. Two marginally significant findings were observed: Child's positive adjustment favored families assigned to family skills training condition relative to minimal treatment and child training only, and family supervision and bonding was lower for children in family skills training than in the other three conditions. Hypotheses about potential explanations for the weaker than expected effects of this program are offered, as are thoughts about the infrastructure necessary to successfully implement family strengthening programs and the future of prevention science.

KEY WORDS: family-based prevention; effectiveness trial; randomized design; African American population; parent training; parent education.

INTRODUCTION

The past 25 years have witnessed major advances in prevention science. Research has demonstrated that preventive intervention in families, schools, and communities can reduce subsequent levels of youth problem behavior, including substance use and crime. Programs aimed at modifying parenting practices and increasing child social skills

have proven particularly efficacious (Gottfredson, 2001; Kumpfer & Alvarado, 2003; McMahon, 2000; Mooney, 1995; Webster-Stratton & Taylor, 2001). But several attempts to implement research-based programs outside of the context of the original research have yielded disappointing results, and Lipsey and Wilson's (1998) meta-analysis of interventions for juvenile offenders shows that greater researcher involvement in the implementation of programs is associated with larger effect sizes for those programs.

One example of an efficacious prevention program for which an effectiveness trial yielded disappointing results is reported in Henggeler *et al.* (1997). This study examined the effectiveness of multisystemic treatment (MST) in community mental health settings without the intense clinical oversight provided in previous clinical trials. In the effectiveness trial testing the generalizability of initial

¹Department of Criminology and Criminal Justice, University of Maryland, College Park, Maryland.

²Department of Health Promotion and Education, University of Utah, Salt Lake City, Utah.

³Department of Public and International Affairs, George Mason University, Manassas, Virginia.

⁴Correspondence should be directed to Denise Gottfredson PhD, 2220D LeFrak Hall, Department of Criminology and Criminal Justice, University of Maryland, College Park, Maryland 20742; e-mail: dgottfredson@crim.umd.edu.

efficacy results, MST did not reduce self-reported delinquency in a sample of violent and chronic juveniles and their families. The researchers attributed the mediocre study findings to poor treatment adherence. Delinquency outcomes were substantially better when adherence to the principles of MST, as reported by parents, adolescents, and therapists, was high (Henggeler *et al.*, 1997).

Alper (2002) reported on several attempts to implement a research-based model of nurse home visits outside the context of the original research reported by Olds *et al.* (1998). In one of the replications the program was implemented with less oversight by program developers. In another, paraprofessionals were hired to implement the program rather than certified nurses as required by the protocol. These replications did not remain true to the original design of the program, and neither produced the same positive results as the original research. The variation in positive outcomes appears to be related to adherence to the program curriculum and the use of qualified program staff.

The present study adds to the growing body of evidence suggesting that as the prevention field attempts to translate positive findings from efficacy research into practice, it may face challenges that are likely to diminish the returns on prevention investments. It reports on a randomized effectiveness trial that tested a model family skills program—the Strengthening Families Program (SFP; Kumpfer *et al.*, 1989)—and its separate components with a substantially different population than in the original research and under far more challenging conditions than in any prior test.

SFP: Rationale, Description, and Prior Research

Several characteristics of families predict the levels of problem behavior of the family's children. Families with negative parent/child relationships characterized by higher levels of conflict (Brook *et al.*, 1990; Dembo *et al.*, 1988; Simcha-Fagan *et al.*, 1986) and families in which parents fail to monitor and supervise their children (Kandel & Andrews, 1987; Loeber & Stouthamer-Loeber, 1986) are more likely to have children who engage in problem behavior. Families in which discipline practices are clear and consistent without being harsh, punitive, or lax are less likely to have children who engage in problem behavior. Parental warmth and support are also related to lower levels of adolescent problem behavior (Brook *et al.*, 1990; Kandel &

Andrews, 1987; Simcha-Fagan *et al.*, 1986; Vicary & Lerner, 1986). Programs aimed at altering one or more of these family characteristics have been successful at reducing levels of child drug use and antisocial behavior in carefully controlled research studies (Kumpfer & Alvarado, 2003; McMahon, 1999; Mooney, 1995; Serketich & Dumas, 1996; Webster-Stratton & Taylor, 2001).

SFP is one of these programs. Designed to target elementary school-aged children, SFP is based on cognitive-behavioral social learning theory and family systems theory. In this program (described in more detail later), parents receive training in parenting skills (PT), children receive training primarily in social skills (CT), and the entire family receives training in family skills (FT). SFP was originally tested and found effective in reducing parent, family, and youth risk factors for substance use and later youth substance use in children of drug abusers in treatment (Kumpfer & DeMarsh, 1985). This initial research contrasted the FT condition to CT and PT and found evidence that the full program was more beneficial than either of its components alone. Subsequently, independent investigators developed and tested culturally adapted versions of the full SFP program with several cultural groups (Aktan, 1995; Aktan *et al.*, 1996; Harrison *et al.*, 1995; Kameoka, 1996; Whitbeck & Smith, 2001). Most of these replications or quasi-replications reported positive effects, especially when the programs were implemented with high-risk families by a single sponsoring agency (Kumpfer *et al.*, 2004; National Institute on Drug Abuse, 1997).

SFP has been identified as an effective prevention program by several federal agencies interested in reducing substance use and delinquency. For example, it is one of several “model programs” disseminated by the Center for Substance Abuse Prevention (CSAP) and is now being implemented statewide in several states. It has, however, never been tested in a rigorous effectiveness trial, and research subsequent to the initial study has tested only the full program. Given the large body of research (cited earlier) supporting PT programs, and emerging evidence that grouping high-risk youths together to receive programming may have iatrogenic effects (Dishion *et al.*, 1999), there is a need to replicate the initial research that demonstrated additional benefits accruing for families receiving the FT program, which involves grouping high-risk youths together for skills training and is, of course, more costly than either CT or PT training.

This paper reports on a randomized effectiveness trial designed to assess the outcomes of SFP among children of primarily African American families in the Washington DC metropolitan area. It assesses the separate effects of three different experimental conditions—FT, PT, and CT—compared to a minimal treatment (MT) control group. The primary hypothesis tested in the research is that the FT condition produces more positive outcomes on the measures targeted by the program than the MT, PT, or CT conditions as was found in the original SFP research. Secondary hypotheses are that both the CT and the PT conditions produce more positive outcomes than the MT condition. The study seeks to further the understanding of how research-based practices can be applied on a larger scale by documenting program effects under natural conditions.

METHODS

Design

The study used an experimental design in which entire families were randomly assigned to one of the four study conditions (FT, PT, CT, or MT). Eligible families were randomly assigned when they arrived at a pre-testing session by graduate assistants, using a table of random numbers. Following randomization, family members completed a pre-test. An immediate post-test was administered during the week following the final program session. Although as will be described below some families assigned to program conditions did not receive the program, all participants were treated in the analysis as they were randomized.

Participants

Participant Recruitment

Five organizations located in the Washington, DC metropolitan area were responsible for implementing the Strengthening Washington DC Families Project (SWFP). One of the five was a pre-release center and targeted families of incarcerated parents.⁵ The other four sites recruited families from their communities using a variety of strategies. Culturally matched site coordinators recruited at local

events (such as health fairs) and in shopping malls. They worked with schools, churches, and other social service providers to get referrals. They knocked on doors, called, and wrote letters to identify families. Brochures, advertisements, and Public Service Announcements (PSAs) were also used.

Although the program was designed to target youths identified as at-risk for conduct problems or substance use, in this study “at-risk” was defined primarily on the basis of participant youths’ residence in a high-risk neighborhood. Polizzi-Fox *et al.* (2004) described the implementing organizations and the communities from which families were recruited. One community was extremely impoverished, with a median household income of \$30,533 and a 54% unemployment rate, according to the 2000 Census. The unemployment rates in the others ranged from 25 to 33% and their median household incomes ranged from \$51,831 to \$68,074.

Description of Participants

Approximately 1400 families were approached, of whom 715 (51%) enrolled (e.g., registered and were pre-tested) in the program.⁶ These 715 enrolled families were the participants in this study. Family units were defined to include not only those with biological parents, but those with foster parents, step parents, boy or girlfriends living with the child, grandparents and other relatives who cared for the child. Eligible families had a child between the ages of 7 and 11, and they had to be able to read, speak, and understand English. Excluded from the sample were family members with mental disorders so severe that they could not function well in a group as well as those with salient needs for treatment (e.g., drug treatment) of a kind not offered by SWFP that would make participation difficult or inappropriate. All of the children in the family below the age of 12 were invited to attend the program, but children under the age of 7 received child care rather than programming.

Data were collected from all family members who participated in the program, but for the analyses reported in this paper, data from only the “target” child and the primary parent (usually the mother) attending the program are analyzed. During registration, parents identified which child (if more than one

⁵Over time, this site began to recruit families from the surrounding neighborhood as well as the center residents.

⁶Data were not collected on the families who did not enroll. It is therefore not possible to determine how the recruitment process influenced the characteristics of the sample.

child was participating) had the most problematic behavior. This child became the “target” child for purposes of the study.

The participating parents were predominantly African American (75%) and female (94%). Many were socio-economically disadvantaged, with 52% reporting a combined family annual income of less than \$20,000. The median number of children living in the home with the parent was two. Most of the children were between the ages of 7 and 11, as expected. Half of children included in the evaluation were 7 or 8 years of age, and half were between 9 and 11. None of the demographic characteristics examined differed significantly by experimental condition.

The sample of families can be compared with families in a recent replication of SFP with families of mothers receiving substance abuse treatment in a Philadelphia (Gutman *et al.*, 2004). The families in the Gutman *et al.* study were indisputably “at risk,” with 76% of mothers reporting having been unemployed in the past 3 years and 70% receiving TANF/welfare. According to parent reports at pre-test, 13% of children in the Philadelphia sample had health problems, compared with 11% in the SWFP sample. The percentage of children with learning disabilities was 16% and the percentage whose school progress was only fair or poor was 32% in both samples. The participants in the SWFP sample therefore appear “at risk” for subsequent problem behavior even though they were not selected on the basis of an *individual* diagnosis of risk.

Procedures

Implementing Organizations, Staffing, and Training

Five organizations were responsible for implementing SWFP. One was a community-based work release facility for male and female offenders who were within 6 months of their release dates. Residents in this facility were required to participate in a number of skill-building programs as part of their program contracts. SWFP was one of the several programs that fulfilled this program contract requirement. Another implementing organization was a resident-founded, non-profit corporation focused on increasing commerce within the community, providing resources to community residents to ensure their financial security, and helping families with high-risk children. The other three organizations were departments or divisions within local county or city gov-

ernments. These agencies offered SWFP as part of a broader array of community services. Family services were central to the mission of only one of these agencies. One provided education and substance abuse services as well as inpatient substance abuse treatment for community residents, and another focused on the health needs of its residents.

The organizational structure utilized in the SWFP involved one overarching governmental entity, which subcontracted funds to these five separate organizations, which in turn hired primarily contract part-time workers to implement SWFP. These program implementers generally had full-time jobs elsewhere and were not regular staff of the implementing organizations. The staffing at each site consisted of a half-time site coordinator, four trainers, child care workers, and a van driver, all of whom worked on an hourly basis. Most site coordinators had college degrees (67% graduate; 28% bachelors) and relevant prior experience as social service directors, teachers, counselors, or community specialists. Trainers were also well-educated: 48% had graduate degrees, 36% bachelors, 6% associates, and 6% had attended some college. The trainers generally had previous experience as teachers, counselors, school nurses, or in some other service provision role. All trainers and site coordinators received the standard SFP training, which was offered either by the program developer or another certified SFP trainer initially at the beginning of the project and at several points during the project for newly hired trainers. The length of the training sessions ranged from 14 to 20 hr. Trainers were supervised by the site coordinators, and the site coordinators were supervised both by their supervisors within their organizations and by the full-time supervisor for the project who worked for the overarching governmental organization. The amount and quality of supervision is not known.

Experimental Conditions

Just prior to beginning the pre-test, families were randomly assigned by graduate research assistants (GRAs) to the four study conditions as follows: 176 in the child skills training only (CT) condition, 177 in the parent skills training only (PT) condition, 188 in the parent and child skills training plus family skills training (FT) condition, and 174 in the minimal treatment (MT) control group. Families assigned to the FT condition received fourteen 3 hr skills training

sessions⁷ including: (a) 1 hr of pre-class activities including a family meal, (b) 1 hr of parent's skills training and a simultaneous children's skills training class, and (c) 1 hr of family skills training conducted in two multi-family groups with instruction and coaching by two group leaders in each group. All participating family members attended the pre-class activity. Children below the age of 7 received child care during the 2 hr of programming.

During the parents' training classes, parents met (without children) with two group leaders to learn to increase desired behaviors in children by increasing attention and rewards for positive behaviors, as well as about clear communication, effective discipline, substance use education, problem solving, and limit setting. During their separate skills training class—also led by two group leaders—the children received instruction in effective communication, behavioral principles of rewarding desired behaviors and ignoring undesired behaviors, understanding feelings, anger control, social skills, problem solving, resisting peer pressure, consequences of substance use, and compliance with parental rules. During the family skills training sessions, families engaged in family activities and practiced therapeutic child play, family meetings, communication skills, effective discipline, and reinforcing positive behaviors in each other. They also planned family activities together.

The PT condition involved the same 1-hr per week parent skills training session described earlier for the FT condition. The children of the parents receiving this training were offered child care during this hour. The CT condition involved the same 1-hr per week child skills training session described earlier for the FT condition. Parents of these children were not expected to receive services, but in 7 of the 10 CT groups, the site coordinator deemed it necessary to offer some “alternative services” to the parents of the children receiving child training while they waited for their children. These services consisted of either the same four-session health and wellness program offered to the MT families described below (four groups), or a combination of “unstructured discussions” with the site coordinator and recreational activities such as viewing entertainment videos or receiving manicures (three groups). Site coordina-

tors were instructed to avoid topics related to family management and child-rearing in their alternative services. Site coordinators reported that 57% of the parents assigned to the CT condition attended any sessions, and of those who attended any, the average number of sessions attended was 6.7.

The parents in the MT groups were to be offered a standard program consisting of four sessions on health and wellness, with instructional videos and discussions related to HIV/AIDS and gun control. In actuality, six of the eleven MT groups received no services, three received the standard MT program, one received unstructured discussions with the site coordinators and one received a stress management intervention. Site coordinators reported that approximately half of the participants assigned to the MT condition attended any sessions, and of those who attended any, the average number of sessions attended was 3.7. No significant differences were found across experimental conditions in the percentage reporting at the post-test that they had received any other family services, employment services, services to help get along with others, or services to help “set limits.”

The program ran in “cycles.” Each program site alternated between running a 14-session CT and PT intervention simultaneously or a 14-session FT and a four-session MT intervention simultaneously. During the CT/PT cycles, approximately 24 recruited families were randomly assigned to receive either the PT or CT condition. During the FT/MT cycles, approximately 24 recruited families were randomly assigned to receive either the FT or the MT condition. Two sites operated three of each type of cycle. Two sites operated four of each type of cycle. One site operated three CT/PT and four FT/MT cycles. The total number of each type of cycle was therefore 17 CT/PT, and 18 FT/MT.⁸

Assessment

Pre- and post-tests were administered to the parents and children of participating families before and directly after the intervention.⁹ The post-testing

⁷The program was designed to be delivered once per week over a 14-week period. In SWFP, 16 of the 35 classes were run using twice per week sessions over a 7-week period. This local adaptation was approved by the program developer.

⁸Each FT condition included a child and a parent training class as well as a family training class, so the total number of each parent and child training class delivered was 35 (17 plus 18).

⁹Two follow-up surveys were also conducted at 6 and 18 months following the completion of the intervention. Data from these surveys were examined using latent growth curve analyses to assess the extent to which any of the four experimental

session occurred the week following the final program session. Families assigned to all conditions, including those no longer attending sessions, were invited. Make up sessions were held, and individual test administrations either at the home or in an agreed upon public location were conducted following the group sessions until a 70% response rate was achieved. The data collection procedures therefore varied across individuals in the study in order to achieve a reasonable response rate, but not across study groups.

Tests were administered primarily in small groups. A graduate research assistant read the survey and response options out loud so that participants who chose not to read for themselves or had trouble reading could follow along. If it became clear that all or most of the participants were reading the survey independently, the GRA stopped reading to the entire group and instead worked individually with the smaller groups of less literate participants.

Retention Strategies

Attendance incentives were provided, including transportation, child care, and a family dinner at each session. Food vouchers ranging in price from \$25 to \$100 were given following each testing session. Additional \$25 vouchers were provided at the third and seventh training sessions and small incentives (usually worth less than \$1) were awarded for punctual attendance and the completion of homework at each session.

Outcome Measures

All outcome measures examined in this paper were taken from pre- and post-tests constructed from standardized scales used in prior research. When possible, complete scales developed in prior studies were used. In some cases, the original scales included too many items to be able to include them all in an instrument to be administered in one sitting. In these cases

conditions resulted in longer term changes in the same outcome measures examined in this report. The results of this exploration are reported in a separate report (Wilson, 2004). No long-term effects of any of the experimental conditions were found. Because Wilson's report uses a different method of analysis and a different sample (e.g., only those families include in the follow-up surveys), page limitations do not allow us to include the details of the analysis in this report. It will be submitted for publication separately.

(most notably, for parent reports of the target child's behavior and positive adjustment), items that seemed to capture the outcomes most directly targeted by the program were selected jointly by the program developer and the first author.

These items and scales were combined to create three different surveys: A survey consisting of 56 items was administered to younger children (those aged 7 and 8); a 138-item questionnaire was administered to older children (aged 9 through 11), and a 195-item survey was administered to parents. The parent survey asked the parent to report on the behavior of the target child, the parents' own behaviors and attitudes, and on the family in general. The child surveys also asked about the family and about risk and protective factors targeted by the program in addition to child problem behaviors. The older child survey included several items and scales intended to measure additional risk factors (e.g., peer drug models, drug availability, time spent with parents, impulsive behavior) and substance use that were not included in the younger child survey. Because these measures are available only for the older youths in the study, they are not reported here. A separate report (Gottfredson *et al.*, 2004) focusing on the older child sample found no significant differences across study groups on any of the measures collected only for older youths.

Items were combined to form scales using expectations based on prior research and factor analyses (FA) conducted in the study sample. That is, when intact scales had been taken from prior research, we scored the scales exactly as had been done in the prior research. When selected items were taken from pre-existing scales, we included all items thought to measure the same or related constructs in an exploratory factor analysis. Items that loaded highly (0.3 or greater) on the same factor were averaged (after recoding to ensure all items were scored in the same direction and had the same response categories) to create new scales. Fifteen different scales measuring child antisocial behavior (3), positive adjustment (3), intentions to use drugs (1), negative peer influence (1), and family functioning (7), were initially scored in this manner.¹⁰

¹⁰Items measuring parental substance use, child lifetime use, and parental depression were also included in the surveys. The variability in parent and child self-reports of substance use was low, and these measures did not converge with other scales that should have been related. In addition, several youths reported a decrease in lifetime drug use from pre to post, possibly indicating low

Because many of these original scales were correlated with one another, analysis using all of them as separate outcome measures would have inflated the type I error probability. To reduce the number of correlated outcome measures, a second exploratory FA was conducted on all of the original scales to identify higher order factors. This FA identified (based on both the scree plots and the difference in the eigenvalues) the seven underlying factors described below, all of which were directly targeted by the SWFP program (except for the negative peer associations scale, which was intended to capture a possible negative effect of grouping high-risk youths together). This FA guided the creation of the seven scales used in this analysis. The original scales that loaded highly on each of the seven factors were averaged to create the final scales used in the analysis.¹¹

For all scales, a high score indicates more problematic behavior or greater risk/lower protection. Table 1 shows the number of items in each scale, the scale range, and the alpha reliability. Information is provided in this table both for the original scales and for the higher order scales used in the analyses.¹²

validity in these measures. These scales and items were not analyzed for this report.

¹¹Several of the original scales were not normally distributed and their response formats differed from other scales with which they were to be combined. To resolve both problems, square root transformations were taken for the non-normally distributed scales prior to combining them with the other scales. Specifically, to create the family supervision and bonding factor, the scale representing closeness to parents (which was skewed and had a range of 1–4) was transformed by taking the square root in order to average it with the parental supervision scale (which had a range of 1–2). To create the parenting skills factor, confidence in parenting and consistency in discipline (both of which were measured on 1–5 scales and were skewed) were also transformed by taking the square root before combining them with the other parenting skills scales that had 1–2 ranges. For these two scales, a small number of cases (2–4) were also trimmed to pull the small number of square root transformed scores that were greater than 2 into the required 1–2 range. After these original scales were combined into higher order scales, the distributions of the resulting scales were again checked. The distributions of these higher order scales were closer to normal than were the original scales. Nevertheless, for the three higher order scales whose distributions contained more than a minimal amount of skew (rebellious behavior, family supervision and bonding, and parenting skills), analyses were conducted using both untransformed and transformed variables. These analyses produced essentially the same results, so the untransformed results are presented.

¹²A manual (Cowan & Gottfredson, 2000) containing the wording of all items, their response formats, their sources, and instructions

Table 1. Scale Ranges, Number of Items, and Reliabilities

Measures	Original scale range	No. of items	Alpha
Child problem behavior factors			
Parent reports	1–4	24	0.85
Hyperactivity/impulsivity	1–4	8	0.66
Antisocial behavior	1–4	16	0.83
Child reports	1–3	13	0.80
Rebellious behavior	1–3	13	0.80
Child risk and protective factors			
Intentions to use drugs (child reports)	1–2	3	0.87
Child positive adjustment (parent reports)	1–4	37	0.93
Social skills	1–4	23	0.91
School progress	1–4	3	0.52
Sociability	1–4	11	0.78
Negative peer associations (child reports)	1–2	8	0.51
Family factors			
Family supervision and bonding (child reports)	1–2	21	0.79
Closeness to parents ^a	1–4	8	0.83
Parental supervision	1–2	13	0.62
Parenting skills (parent reports)	1–2	37	0.81
Confidence in parenting ^a	1–5	5	0.60
Consistency in discipline ^a	1–5	8	0.65
Family organization	1–2	8	0.67
Family cohesion	1–2	7	0.65
Family conflict	1–2	9	0.68

^aSquare root transformation was made to reduce skew or kurtosis.

The final scales used in the analyses are as follows:

Child Problem Behavior

Parent Reports. This scale combines two parent scales measuring the target child's behavior: Hyperactivity/impulsivity and antisocial behavior. Items in the hyperactivity/impulsivity scale ask (for example) how often (in the past 3 weeks) the target child "interrupts or intrudes on others" and "is easily distracted." Items in the antisocial behavior scale ask (for example) how often (in the past 3 weeks) the target child fights and takes others' property. Items are taken from the Parent Observation of Children's Activities (POCA-R; Kellam, 1990) and the Social Skills Rating System (SSRS; Gresham & Elliott, 1990).

for recoding items to obtain a common metric among items in each original scale is available from the first author.

*Youth Reports.*¹³ This scale measures child self-reports of rebellious behaviors. Items ask how often the child (for example) fights with other students, cheats on tests, and breaks other people's things. The scale is taken in its entirety from the *What About You* survey (Gottfredson & Gottfredson, 1999).

Child Risk and Protective Factors

Intentions to Use Drugs (youth reports). This scale, taken in its entirety from the *What About You* survey (Gottfredson & Gottfredson, 1999), measures intentions to use alcohol, marijuana, and cigarettes. A sample item is: "I will never drink beer, wine, or hard liquor."

Child Positive Adjustment (parent reports). This scale combines three parent report scales measuring the target child's positive behavior: Social skills, school progress, and sociability. Items in the social skills scale ask (for example) how often (now and in the past month) the target child "tries hard to do good work on tasks" and "compromises in conflict situations with others." Items in the school progress scale ask about the child's school progress, liking for school, and school attendance. Items in the sociability scale ask (for example) how often (in the past 3 weeks) the target child socializes and interacts with other children and avoids other children. Items are taken from the Parent Observation of Children's Activities (Kellam, 1990) and the Social Skills Rating System (Gresham & Elliott, 1990).

Negative Peer Associations (youth reports). This scale includes a subset of items from the *What About You* survey (Gottfredson & Gottfredson, 1999) and reflects the level of delinquent peer influence. A sample item is: "My friends often try to get me to do things the teacher doesn't like."

¹³Child and parent reports of problem behavior did not load on a single factor, as anticipated. The correlations between child and parent reports of similar behaviors were modest. The correlation between child reports of rebellious behavior and parent reports of antisocial behavior, for example, was .25. An extensive analysis of the validity of the child and parent reports utilizing measures available from school records for a subset of the sample indicated that neither could be rejected as obviously invalid. Correlations of child reports of child characteristics generally had slightly higher correlations with measures of school performance taken from school records than parent reports of child characteristics, but both were correlated in the expected direction. Therefore, the two sets of measures of problem behavior were retained for analysis but were kept as separate factors.

Family Factors

Family Supervision and Bonding (youth reports). This scale combines a scale taken in its entirety from the *What About You* survey (Gottfredson & Gottfredson, 1999), parental supervision, and another, Closeness to parents, taken in its entirety from the Survey of Risk and Protective Factors Associated with Adolescent Alcohol, Tobacco, and Other Drug (ATOD) Use (Arthur *et al.*, 2002). Items in the former include "My parents usually know if I do something wrong" and in the latter include "I feel close to my mother."

Parenting Skills (parent reports). This scale comprises five subscales. Three (family cohesion, organization, and conflict) are shortened versions of scales from the Moos Family Environment Scale (FES; Moos & Moos, 1986). The cohesion subscale measures the degree to which the family unit works well together and feels close to one another. The conflict subscale measures the extent of conflict within the family, and the organization scale measures the family's organization as it relates to financial matters, household duties, and so on. A fourth subscale, based on measures used in Kumpfer's prior evaluations of SFP, is confidence in parenting which includes items such as "How much do you enjoy caring for this child" and "How would you rate your ability to parent this child." Finally, the consistency in discipline scale includes items such as "How often do you give in to the child's demands or excuses not to complete work?" and "How often does the punishment you give your child depend on your mood." This measure is modeled after questions from the Oregon Youth Study (cited in Gottfredson *et al.*, 1996).

Process Measures

Participation

Attendance and participation forms were completed weekly (or semi-weekly) at the end of each parent training session by program trainers. Attendance records were collected for 93% of the families included in this analysis. Two variables were created from the attendance records: Attended at least one session (yes/no) and the number of sessions attended. An average participation score was constructed from trainer ratings of each subject's participation on seven variables (attention, amount and appropriateness of sharing and disclosing, interest level,

motivation level, competence in concept). This score ranges from 1 (low participation) to 5 (high participation).

Fidelity

Observations were conducted by University of Maryland GRAs four times during the first program cycle offered by each site and twice during each subsequent program cycle. Observations were undertaken using a standardized fidelity checklist created by the program developer and tailored to the curriculum manuals. The fidelity checklist items included whether or not each major activity was covered as well as ratings of the trainers' program delivery and an overall rating of session quality. In all, 79 observations of each of the child training and parent training were conducted, and 44 observations of family training were conducted.¹⁴ Observations from all observers and classes in each cycle were averaged within each site, so that one score for each type of class in each cycle was available for analysis for each of the observation variables. These observations were used to compute the following measures of program quality: Session length, percent of items covered, trainers' average program delivery rating, and average overall session rating. For the latter two measures, observations were recorded using a three-point scale in which a rating of "1" was "below average," "2" was "average," and "3" was "above average."

Parent Comprehension

At the end of 11 of the 14 parent training sessions, participants completed a "session review" form. The purpose of the form was to assess how much information the participants retained from that night's session. All parents who attended the parent training sessions were expected to complete these forms. Among parents in attendance at the parent or family training sessions, 85% completed parent session review forms. These records were used to calculate the average percentage of items correct across all

sessions attended for each participant. On the night of the post-test, five or six parents whose families were randomly assigned to the CT or MT conditions were asked to complete three session review forms each. These session review forms were identical to the session review forms completed by participants in the FT and PT conditions.

Other Services

As a check on the extent to which potentially confounding services were provided to any study group, we included in the post-test a set of questions asking parents to report whether or not they had received any of a list of different services *other than the SFP program* in the past 3 months (e.g., concurrently with the SFP treatment).

Statistical Analysis

The statistical analysis included a descriptive analysis of process measures, an analysis of overall attrition and differential attrition, and analyses of the parent- and child-reported outcomes using analysis of covariance. A hierarchical approach to the data analysis was considered and rejected due to the small number of sites. Instead, site was included as a dummy variable for each family in all outcome analyses to adjust for any differences in the marginal effect of site. The pre-test measure corresponding to the outcome variable was also used as a covariate in each analysis.

RESULTS

Implementation

Attendance and Participation

Across all three program conditions, 69% of enrolled intervention families¹⁵ attended at least one training session. Seventy percent of CT, 61% of PT, and 76% of FT families attended at least one session ($F(2,499) = 4.20; p < .05$), with a significant difference observed between the FT and PT groups only

¹⁴The numbers of observations differ by type of session because each observation of a family skills training condition generated three distinct class observations—one for each of the child, parent, and family training sessions. The observations of classes in the CT and PT skills training conditions generated only one observation each.

¹⁵Each family was coded as having attended at least one session if any family member attended any of the 14 sessions. The total sessions attended for each family was the number of sessions attended by any family member.

($p < .05$). Among those 348 families who attended at least one training session, the average number of sessions attended was 8.4 out of the 14 delivered. Families assigned to the CT condition attended, on average, 8.5 sessions whereas families in the PT condition attended an average of 7.5 sessions. The average number of sessions attended by the FT families was 9.1. These differences across conditions were statistically significant ($F(2,345) = 3.76; p < .05$), with a significant difference observed between the FT and PT groups only ($p < .05$). The same pattern was observed for on-time arrival and completion of homework. The average participation level (for individual family members) across all program conditions was 4.2 on a five-point scale. The differences across conditions were statistically significant ($F(2,581) = 6.59; p < .01$), with the average participation level significantly higher for the PT than for the FT and CT groups (4.4 compared to 4.2 and 4.2, respectively).

Implementation Fidelity

Each type of session was designed to last 60 min. According to the observations, the child training classes lasted 54 min and the parent classes lasted 55 min, on average. The family training sessions (which were delivered directly following the parent and child sessions for families in the FT condition) were briefer, running an average of 36 min. Most of the program material was covered in the child (92%) and parent (90%) training sessions, but only 62% of the items to be covered in the family-training classes were covered. This is not surprising considering the sessions were delivered in a much shorter time than was intended by the program developers. The average delivery ratings from observations were significantly different across conditions ($F(2,85) = 4.58; p < .05$). Pairwise contrasts showed that the average delivery rating was significantly lower for FT than for the PT and CT groups (2.6 on a three-point scale compared to 2.8 and 2.8, respectively). The average leader ratings were as follows for the three types of training sessions: 2.8 (CT), 2.7 (PT), and 2.6 (FT), also on a three-point scale ($F(2,85) = 2.55, p < .10$). The post-hoc pairwise contrasts revealed again that the FT condition was marginally significantly lower than PT and CT groups.

Parent Comprehension

Parents who participated in and completed assessments in the PT and FT conditions answered

86% of the questions on the session review forms correctly. The average percent correct across reviews completed by parents in the CT and MT conditions (who had not received training) was 76%,¹⁶ and the percent correct was significantly different for parents who received the training than for those who did not ($F(1,343) = 53.84; p = .00$). These reviews provide evidence that the parents who attended the sessions understood the material covered, but the high percentage of correct responses for the parents in the conditions that did not receive any parent training suggests that much of the material covered in the parent sessions is common knowledge.

Staff Turnover

Over the course of the 5-year project, the program experienced a 420% turnover in the site coordinator position (that is, 21 different people filled the five positions) and 315% turnover in the program trainer position (63 different people filling 20 positions).¹⁷ In some cases, trainers were promoted to site coordinators, so the overall turnover rate across both positions was 288%. The overall project manager position was held by three different persons. Despite the high rate of turnover, all trainers and site coordinators were trained in SFP.

Attrition from the Research

The overall attrition rate from the pre-test to the post-test for this study was 30%,¹⁸ for a post-tested N of 502 families. This rate is moderately high, but in line with other implementations of SFP reported in earlier research (Kumpfer *et al.*, 2002). Nevertheless, it raises the possibility that (a) the results of the research does not generalize to the entire study sample and (b) differential patterns of attrition across the four study groups undermine confidence in the causal interpretations of the results.

¹⁶The percentage correct can be calculated either by assuming all unanswered questions are incorrect (as is reported above) or by calculating the percentage correct only on the basis of those items with valid responses. The percentage correct based on this latter method produces similar results: 89% versus 77% correct, $p < .01$.

¹⁷Trainer turnover generally occurred between classes so that minimal disruption of ongoing classes took place.

¹⁸That is, 70% of the parents who completed a pre-test also completed a post-test.

An analysis of attrition bias by condition and site showed the attrition rate did not vary by condition or site. To further explore the correlates of attrition, a logistic regression analysis was conducted. This analysis regressed the log odds of attrition on several parent-reported baseline measures¹⁹ to examine whether baseline characteristics of the parents and children were related to attrition. This analysis yielded only one significant association at the $p < .05$ level out of 24 tests conducted, or fewer than would be expected by chance. Parents' age was the only significant predictor, with older parents more likely to complete a post-test.

A series of two-way analyses of variance were also performed to examine the extent to which attrition from the study, condition, and their interaction were related to baseline measures. The interaction between attrition and condition tested whether baseline differences between those who left the study and those who did not differed across conditions. This analysis was performed on parent-reported baseline and demographic measures (see footnote 15). The results revealed no statistically significant interactions (all p -values were greater than .60). Based on these analyses, it seems unlikely that attrition from the study biases the results in any important way. Further, the similarity between those who left the study and those who did not suggests that the results can safely be generalized to the entire study sample.

Outcome Analysis

Table 2 shows the pre-test, post-test, and gain score mean, standard deviation, and sample sizes for each condition. This table shows that the general direction of change from pre- to post-test was in the positive direction, regardless of experimental condition. This change is plausibly attributable to a regression artifact.²⁰ ANCOVA was used to assess the ex-

Table 2. Pre-test, Post-test, and Change Score Means for Parent- and Child-Reported Outcome Measures

Variable by site and condition	Pre-test		Post-test		Change score		
	Mean	SD	Mean	SD	Mean	SD	N
Child problem behavior factors							
Parent reports							
Family	1.90	0.39	1.77	0.33	-0.13	0.33	127
Child only	1.93	0.45	1.77	0.43	-0.16	0.32	111
Parent only	1.88	0.41	1.78	0.35	-0.10	0.31	114
Minimal	1.88	0.39	1.79	0.40	-0.09	0.34	125
Child reports							
Family	1.30	0.26	1.33	0.34	0.03	0.32	126
Child only	1.43	0.35	1.36	0.33	-0.07	0.26	115
Parent only	1.37	0.31	1.36	0.37	-0.01	0.32	107
Minimal	1.44	0.35	1.39	0.35	-0.05	0.38	112
Child risk and protective factors							
Intentions to use drugs (child reports)							
Family	1.17	0.32	1.18	0.36	0.02	0.42	123
Child only	1.20	0.36	1.19	0.37	-0.01	0.46	112
Parent only	1.18	0.33	1.11	0.28	-0.07	0.37	104
Minimal	1.20	0.36	1.20	0.36	0.00	0.40	111
Child positive adjustment (parent reports)							
Family	2.22	0.44	2.06	0.41	-0.17	0.36	119
Child only	2.21	0.53	2.11	0.52	-0.1	0.33	108
Parent only	2.14	0.46	2.06	0.44	-0.08	0.35	108
Minimal	2.16	0.45	2.11	0.48	-0.05	0.3	121
Negative peer associations (child reports)							
Family	1.24	0.19	1.25	0.21	0.01	0.21	127
Child only	1.30	0.21	1.23	0.20	-0.08	0.23	114
Parent only	1.29	0.21	1.24	0.20	-0.05	0.20	106
Minimal	1.28	0.21	1.24	0.20	-0.04	0.19	111
Family factors							
Family supervision and bonding (child reports)							
Family	1.24	0.18	1.24	0.18	0.00	0.08	119
Child only	1.25	0.15	1.23	0.15	-0.02	0.08	107
Parent only	1.26	0.17	1.23	0.16	-0.03	0.08	102
Minimal	1.27	0.17	1.24	0.17	-0.02	0.09	109
Parenting skills (parent reports)							
Family	1.40	0.16	1.34	0.14	-0.06	0.13	120
Child only	1.40	0.19	1.35	0.18	-0.05	0.15	109
Parent only	1.38	0.16	1.32	0.14	-0.06	0.13	111
Minimal	1.36	0.15	1.33	0.14	-0.02	0.10	123

tent to which the post-test means on each of the seven outcome scales derived from parent and child surveys varied across condition after adjusting for pre-test scores and site differences. The results of this analysis are reported in Table 3. The group difference on negative peer associations was statistically significant ($p < .05$), and two others (parent reports of child positive adjustment²¹ and child reports of family su-

¹⁹These parent baseline measures included all five of the parenting skills scales used in this study; parent reports of their target child's antisocial behavior and school progress; parent reports of their own alcohol problems and illicit drug use; race; age; income; and number of children.

²⁰Regression would be expected given the high-risk nature of the population and participant self-selection for child problem behavior. As a further check, we examined gain scores for participants at different ranges of the pre-test distributions. This analysis revealed that families who tested at the more negative end of the distribution at the time of the pre-test improved the most between pre- and post-test. The association of gain score with pre-test was highly significant and did not differ by experimental condition.

²¹In analyses that examined each of the original scales separately, the F -test for parent reports of social skills reached statistical significance ($F = 5.62; p = .05$).

Table 3. Test of Difference Between Conditions at Post-Test

Measures	<i>F</i>	<i>p</i>	Pairwise contrasts
Measures			
Child problem behavior factors			
Parent reports	1.26	0.29	
Child reports	0.47	0.71	
Child risk and protective factors			
Intentions to use drugs (child reports)	1.81	0.14	
Child positive adjustment (parent reports)	2.22	0.09	F < M*; F < C ^a
Negative peer associations (child reports)	2.76	0.04	C < F**; P < F ^a
Family factors			
Family supervision and bonding (child reports)	2.17	0.09	P, M < F*; C < F ^a
Parenting skills (parent reports)	0.96	0.41	

Note. All scales are scored so that a higher score is less favorable. The analysis of covariance models included the pre-test score of the dependent variable, site, and condition. Degrees of freedom in the numerator of the *F*-statistic is always 3, and in the denominator ranges from 404 to 430.

^a*p* < .10.

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

A1

supervision and bonding) were marginally significant (*p* < .10).

Pairwise contrasts between conditions for the significant and marginally significant findings showed that children in the CT condition (*p* < .01) and the PT condition (*p* < .10) reported less negative peer associations than children in the FT condition. It was anticipated that grouping high-risk youths together might increase negative peer associations. Both the family and child conditions grouped the children for skills training, but a relative increase in negative peer associations was observed only for the FT condition.

The pairwise contrasts for the marginally significant effect on parent reports of child positive adjustment showed that the FT families outperformed the MT families (*p* < .05), and the CT families (*p* < .10), but not the PT families on this outcome. For the marginally significant effect on child reports of family supervision and bonding, the pairwise contrasts showed that the FT condition families fared significantly *worse* than the MT and PT families (*p* < .05), and marginally worse than for the CT condition (*p* < .10).

Finally, Table 4 shows the effects sizes (ESs) for the pairwise comparisons of greatest interest. These ESs were computed as covariate-adjusted mean dif-

ferences between conditions, adjusting for pre-test and site, divided by the pooled within-conditions and within-site standard deviation. These are standardized mean difference-type effect sizes, similar in metric to Cohen's *d* (Lipsey & Wilson, 2001). Presented for descriptive purposes only, these ESs are intended to call attention to trends in the data that, although not statistically significant, may suggest hypotheses for testing in subsequent research. Most of the effect sizes are small. Only three of the 35 effect sizes are as high as .3, which is generally considered the bottom range for a "moderate" sized effect, and one of these is in the unfavorable direction (child reports of family supervision and bonding contrasting the FT and PT conditions). A few general patterns are evident: Effects were larger and more positive for parent-reported than for child-reported outcomes (−.13 versus .08, on average). Effect sizes for the three experimental conditions relative to the MT showed that effects were smaller for the FT condition than for either the CT or the PT conditions (−.01 versus −.08 and −.07, on average). The effect sizes for child problem behavior were somewhat higher for the CT than for the other conditions, and the effects sizes for the parenting skills were somewhat higher for the PT than for the other conditions. The FT condition performed worse than the CT condition on all outcomes except parent reports of child positive adjustment and parent reports of parenting skills, and worse than the PT condition on all outcomes except parent reports of child positive adjustment and parent reports of child problem behavior. Although tentative due to the preponderance of non-significant differences, the pattern suggests that the CT and PT conditions outperformed the FT condition in part because the children in the FT condition were more exposed to negative peers and became less bonded to their parents than the children in the PT and CT families.

A final analysis examined the hypothesis that the weaker than anticipated effects observed for the FT in this study resulted from the provision of services to the MT groups. As noted earlier, six of the eleven MT cycles received no alternative services. Support for the hypothesis would be found if the effects for the FT were stronger and more positive (relative to the MT) in these six cycles. All ANCOVA analyses as described above were repeated for the MT and FT families, adding an indicator for whether or not alternative services were provided for the MT families and a treatment by MT services interaction term. These analyses provided no support for the hypothesis: FT families did not outperform MT

Table 4. Effect Sizes for Comparisons, Adjusting for Pre-test Differences

Measures	FT/MT	CT/MT	PT/MT	FT/CT	FT/PT
Child problem behavior factors					
Parent reports	-0.11	-0.21	-0.04	0.08	-0.08
Child reports	0.02	-0.09	0.04	0.17	0.02
Child risk and protective factors					
Intentions to use drugs (child reports)	0.00	-0.01	-0.07	0.00	0.06
Child positive adjustment (parent reports)	-0.31	-0.12	-0.11	<i>-0.30</i>	-0.18
Negative peer associations (child reports)	0.21	-0.11	-0.01	0.29	<i>0.21</i>
Family factors					
Family supervision and bonding (child reports)	0.27	0.04	-0.04	<i>0.24</i>	0.30
Parenting skills (parent reports)	-0.17	-0.09	-0.23	-0.10	0.07

Note. Negative effect sizes are desirable. Bold ESs are for pairwise comparisons that were statistically significant ($p < .05$) in the ANOVA analyses. Italicized ESs are for pairwise comparisons that were marginally statistically significant ($p < .10$) in the ANOVA analyses.

families, regardless of the provision of MT alternative services.

DISCUSSION

This study sought to replicate the SFP program and its separate components with a substantially different population (e.g., predominantly African American and urban) than in the original research and under more challenging conditions (e.g., with multiple implementing organizations serving predominantly socio-economically disadvantaged communities) than in any prior test. Results showed that only one of the study measures (negative peer associations) differed significantly across any of the study groups, and two others (family supervision and bonding, and child's positive adjustment) showed marginally significant group differences. Little support was found for the primary hypothesis tested in the research—that the FT condition produces more positive outcomes on the measures targeted by the program than the MT, PT, or CT conditions. No statistically significant positive effects of the full program were observed, although results for parents' reports of child's positive adjustment were marginally significant favoring the FT condition over the MT and CT conditions, but not the PT condition. Other significant and marginally significant findings suggested that the FT condition produced worse outcomes: Results indicated a statistically significant relative increase in negative peer associations for children of families assigned to the FT condi-

tion compared with the CT and PT conditions. A marginally significant effect also showed children in the FT condition reported lower family supervision and bonding than children in the other three conditions. Finally, none of the results suggested that the CT or the PT groups outperformed the MT group.

The results from this randomized trial are not as strong as anticipated from prior SFP research (Kumpfer *et al.*, 2002), but are similar to the results reported from a recent SFP replication conducted with a sample of mostly African American Philadelphia families (Gutman *et al.*, 2004). That study was conducted under conditions more similar to the original SFP research: The parents were receiving substance abuse treatment and were characterized by a number of factors that placed their children at elevated risk for problem behavior. The program was delivered with high fidelity by research staff, following the once per week session format as recommended by the developer. Immediate drop-out of participants was rare in that study (only 7.8% as opposed to 26.6% for FT families in SWFP), but the attendance patterns for participants who did attend were similar across studies (for example, 38.8 and 34.9% of families attended 11–14 sessions in the Philadelphia and SWFP FT samples). The Gutman study found positive effects of the full SFP program on parenting skills, but, similar to our study, no effects on child pro-social skills, problem behavior, or school progress.

Methodological issues, differences in the population, and implementation factors might explain the

weaker than anticipated effects found in this recent SFP research.

Methodological Issues

Previous studies of SFP relied solely on parent-reported outcomes. This study shows that the results differed somewhat depending on the source of the outcome data. Although both sources of data produced reasonably reliable measurement, parent and child reports of similar behaviors were not highly correlated, and parent reports generated more positive effects than child reports. Examination of the validity of both sets of measures (see footnote 9) failed to produce evidence that child reports were less valid than the parent reports. This finding underscores the importance of using multiple sources of measurement when assessing program effects. In this study, the child reports yielded important new information about potential negative effects of the family skills training condition that had not previously been considered.

The design of this study is also more rigorous than the designs used in most prior SFP research, which have been quasi-experimental. Even the original NIDA-supported research (Kumpfer & DeMarsh, 1985) did not fully randomize families to conditions. In that study, the three treatment conditions (CT, PT, and FT) were randomly assigned, but the control group was not. Fully randomized designs are more credible.

Appropriateness of the Population for the SFP Program

The fact that two different studies of SFP effects on predominantly African American, urban populations found minimal effects on child outcomes raises the possibility that the program is not effective for this population. The enormous difficulties related to recruitment and retention of families experienced in the SWFP make clear that the program was not perceived as a high priority by many clients recruited and enrolled in the program. Elsewhere (Polizzi-Fox *et al.*, 2004) we reported results from a qualitative analysis of interviews with program implementers in the SWFP. Although more than one implementer thought the content of the program (especially its emphasis on non-physical discipline strategies) was not as culturally relevant to the population

being served as it could have been, by and large the most frequently voiced concern was that the program was not well-suited for low SES families. Although *cultural* appropriateness of the program has received considerable attention in the prevention field (Kumpfer *et al.*, 2002), issues related to *social class and community disorganization* seem to have more bearing on the program effects than cultural variables, according to program implementers. For socio-economically disadvantaged families, day-to-day challenges of life made it difficult to plan for and carry out long-term programmatic changes in their lives.

Prinz and Miller (1996) discuss situational demands and constraints such as poverty, unemployment, and health that influence the engagement of parents in skills training. Two additional qualitative analyses of program attrition and low participation conducted during the course of the SWFP confirmed the importance of such factors in explaining program drop-out and nonattendance: Polizzi and Gottfredson (2003) conducted phone interviews with program non-completers. In this study, adult non-completers reported that, in addition to often being misinformed during recruitment about the content of the program, they often lacked accessible transportation—even though the program was supposed to provide it. They also cited family illness and scheduling conflicts as major problems preventing them from completing the program. Vilmenay (2002) conducted additional telephone interviews to identify reasons for low participation among subjects. Interviews were administered at two different points during the program class with families who missed two classes prior to the fifth training session or seven or more classes before the twelfth training session. Thirty-two percent of the poor attending families could not be contacted because their telephones were either disconnected, not in service, or the phone was never answered, indicating that the transient lifestyle of the clients may in part explain their low attendance. The interviews revealed that the most frequent reasons for non-participation included schedule conflicts, personal issues, and misinformation or no contact by the site coordinator.

These qualitative findings suggest that developing a client population for a program such as SWFP in socio-economically disadvantaged areas is likely to be a major undertaking. Previous reviews of family interventions (Sanders, 1996) have concluded that major alterations in program structure may be nec-

essary to reach the families who would most benefit from family interventions. Sanders *et al.* (2002), in their examination of the development and dissemination of their family intervention—Triple P—suggested that agency support, sufficient budgetary funds, and community outreach are essential to engage and retain families. They also recommended the use of focus groups to help shape the intervention when working with diverse populations. Many of the recommended strategies (e.g., transportation, child-care, locating the program within the community in close proximity to the clients) were tried with little success in the SWFP. We conclude that such auxiliary services may be necessary but are not likely to be sufficient to raise attendance and participation levels in truly challenged communities. In communities in which families find it difficult to provide meals and shelter for their children, in which people are fearful for their safety, and in which medical and mental health crises must be dealt with on a regular basis, more basic needs will have to be met before families will be able to engage in family strengthening sessions. Future efforts should plan for ample time and resources to explore the needs of the community prior to the start of the program and should anticipate the need to embed the program into a more comprehensive effort to meet community needs.

Implementation Fidelity

Our observations showed reasonably high fidelity to the program model for the PT and CT sessions, with 90 and 92% of the material in the trainers' manual being covered. The FT sessions, however, were often cut short and only 62% of the expected material was covered.

Although the retention rates were higher in the FT than the PT and CT conditions (76% versus 61% and 70% attended at least one session), the family training hour of the FT condition was clearly not well implemented. Many of the scheduled activities were not carried out, and observations showed that the quality of program delivery was lower in the FT than in the CT or PT sessions. This may explain the poor results for this condition relative to the CT and PT training. On the other hand, the recent Philadelphia replication of the full SFP program, in which the SFP trainers were hired and monitored by the research staff, achieved high fidelity but still did not produce positive outcomes for children in the participating families. The SWFP experience suggests that when

implemented incompletely as it was in this study, the full SFP condition has the potential to produce negative outcomes, perhaps by providing a social context in which youths are free to socialize in an unstructured environment. This type of unwanted negative effect is consistent with recent work by Dishion *et al.* (1999) who show that grouping youth—especially high-risk youths—for interventions may increase their vulnerability to “deviancy training,” or subtle reinforcement of deviant beliefs and behaviors through laughter, social attention, and interest. This study highlights the importance of *structure* in family interventions and suggests the need for additional studies of the FT condition to better understand the trade-offs involved between providing activities that increase family attendance levels (such as the family dinner) but that may allow undesirable effects to emerge. Because negative peer associations are the most potent predictor of subsequent problem behavior (Gottfredson, 2001), such research should take high priority and until this dynamic is better understood, intervention that group high-risk youths together in relatively unstructured settings for intervention should be attempted only with due caution.

Despite what appeared to be reasonably high fidelity of implementation in the PT and CT conditions, these program conditions did not achieve the expected results. This suggests that the success of these major program components hinges on something other than coverage of the material contained in the trainer's manual.²² While the trainers in SWFP succeeded at “covering the material,” they may not have possessed the therapist skills necessary to produce a meaningful change in the clients, despite their educational and work history credentials. As Prinz and Miller (1996) have suggested, special skills may be required to achieve client engagement. Therapists may have to focus on increasing therapist–client interaction, including building on existing family competencies, customizing the program to meet the family's values, needs, and routines, broadening the focus of treatment, and building social connections.

²²That the content of the program is less important than other aspects of the program, such as the opportunity it affords to make social connections, is underscored by the fact that the parents assigned to the MT and CT groups were able to answer 76% of the content items correctly without benefit of the parent sessions. Knowledge of the parenting skills covered in the training session is probably not the most important active ingredient. This finding corresponds to findings from drug prevention research indicating that increasing knowledge about substances and their effects does not reduce substance use (Botvin, 1990).

In SWFP, these skills were neither identified as qualifications necessary to join the project nor were they covered in a meaningful way in the trainer training. This is an important lesson for attempts to scale up model programs: The level of staff competencies available in trials in which the developer has more control over the selection and training of staff is not likely to be replicated in more natural settings. This may be especially true in more disorganized areas in which the challenges of recruiting more capable staff are confounded by their unwillingness to work in areas perceived to be dangerous. Site coordinators and trainers complained regularly that their jobs required far more work than they had expected, for not enough pay. As the extremely high turnover rate among trainers and site coordinators in the SWFP program attests, the rewards for working on this program were insufficient to outweigh the challenges. Clearly, more resources, both financial and infrastructure-enhancing, will be needed to attract and train skilled trainers and to administer the program effectively in future attempts to scale up the SFP program in areas such as those served in SWFP.

This report documents that efforts to disseminate evidence-based programs into less amenable settings than were used in the original research may not be as effective as the initial efforts. It joins the ranks of several other recent effectiveness trials (e.g., Alper, 2002; Cho *et al.*, 2004; Gutman *et al.*, 2004; Henggeler *et al.*, 1997) that show that enhancements to model programs may be required to boost program effectiveness in more challenging settings.

Implications for Future Directions

The work leads to two recommendations for future attempts to scale up model programs: Conduct pilot and “pipeline studies” and enhance training and technical assistance. Before launching full-scale dissemination in an area, it is advisable to conduct pilot trials of the program procedures—especially with respect to recruitment and retention—and pipeline studies to see if the anticipated population is actually available. These relatively low cost activities would be helpful in identifying major challenges early in the process. Such a preliminary study would have immediately surfaced the largest challenges faced in this study: Recruitment and retention. Of 1403 families recruited, only 1036 (74%) registered. Only 715 (69% of the registered families) showed up to complete a pre-test and only 69% of these pre-tested fam-

ilies attended at least one session. Second, the level of training and technical assistance required to achieve high fidelity implementation is likely to vary considerably from place to place. Some places may have a suitable infrastructure and highly skilled staff already in place. In these settings, the usual training with minimal technical assistance may suffice. In other settings, a more intensive process may be required to develop the necessary prevention infrastructure. Different training and technical assistance modules may be required to meet the needs of different locations. An initial assessment of the resources (personnel and other) available at the site might be used to determine what level of training and technical assistance is required. Also, minimum standards should be established for conditions that must be in place in the host setting prior to implementation.

Finally, the SWFP project results suggest that the prevention field should consider shifting the focus away from “installing” model programs into settings and towards creating a fit between existing knowledge about effective prevention practices and specific environments or markets. The emerging challenges are to build bridges between this knowledge and the host environments, helping to create the infrastructure necessary to support research-based practices as necessary. Also, a more thorough assessment of community needs may reveal more pressing needs than those addressed by the program. It is possible, for example, that the premise that greater parental knowledge and skill are what is needed by disadvantaged families in disadvantaged communities may miss the mark, and that more effective interventions might appropriately focus on social organization or family economic requirements.

The Society for Prevention Research has recently identified as a major priority promotion of research to further the understanding of how research-based policies and practices can be effectively applied on a broader scale (Botvin, 2003). This agenda for prevention researchers includes systematically testing the efficacy of various dissemination strategies in experimental studies. These studies should deliberately contrast different approaches to achieving high fidelity and adoption of effective strategies; measure the effects of these interventions on both the quality of implementation and on prevention outcomes; and eventually examine benefit in relation to cost. Moreover, as Rotheram-Borus and Duan (2003) noted, the prevention field should attend to the development of theories of adherence to participation and the incorporation of such theo-

ries into the programs. These theories of adherence should address accessibility of the program, provider attitudes, and consumers' perceptions of the prevention setting. As Rotheram-Borus (2004) suggested, it is time for an "extreme makeover" in prevention that focuses more on what each population needs and less on strict adherence to model programs.

ACKNOWLEDGMENTS

We wish to thank Suzie Johnson and Carol Small of the Metropolitan Washington Council of Government and the many directors, site coordinators, and trainers in the five implementation sites in the metropolitan area of Washington, DC for their assistance with this research. Allison Payne, Bianca McClenny, and Amanda Hall also helped with the research, and Charles Turner, Richard Catalano, Gary Gottfredson, Carol Metzler, and two anonymous reviewers provided useful feedback. Support for this research was provided by the National Institute on Drug Abuse through a grant to the University of Utah.

REFERENCES

- Aktan, G. (1995). Organizational framework for a substance use prevention program. *International Journal of Addictions, 30*, 185–201.
- Aktan, G., Kumpfer, K. L., & Turner, C. (1996). Effectiveness of a family skills training program for substance abuse prevention with inner city African-American families. *International Journal of the Addictions, 31*, 158–175.
- Alper, J. (2002). The nurse home visitation program. In S. L. Isaacs & J. R. Knickman (Eds.), *To improve health and health care. Volume V. The Robert Wood Johnson Anthology* (pp. 3–22). San Francisco: Jossey-Bass.
- Arthur, M. W., Hawkins, J. D., Pollard, J. A., Catalano, R. F., & Baglioni, A. J. (2002). Measuring risk and protective factors for substance use, delinquency, and other problem behaviors: The Communities That Care Youth Survey. *Evaluation Review, 26*, 576–601.
- Botvin, G. J. (1990). Substance abuse prevention: Theory, practice, and effectiveness. In M. Tonry & J. Q. Wilson (Eds.), *Drugs and crime* (pp. 461–520). Chicago: The University of Chicago Press.
- Botvin, G. J. (2003). *From research to policy: Advancing prevention science and practice*. Presidential Address presented at the 11th Annual Meeting of the Society for Prevention Research, Washington, DC.
- Brook, J. S., Brook, D. W., Gordon, A. N., Whiteman, M., & Cohen, P. (1990). The psychosocial etiology of adolescent drug use: A family interactional approach. *Genetic, Social and General Psychology Monographs, 116*, 111–267.
- Cho, H., Hallfors, D., Kim, H. M., Khatapoush, S., & Sanchez, V. (2004). *Findings from a randomized controlled trial of the effectiveness of an indicated prevention program*. Paper presented at the annual meeting of the Society for Prevention Research, Quebec City.
- Cowan, D., & Gottfredson, D. C. (2000). *The strengthening Washington, DC families program: Development of scales to measure SFP goals and objectives*. College Park, MD: Department of Criminology and Criminal Justice. Technical report available from authors.
- DeMarsh, J., & Kumpfer, K. L. (1986). Family-oriented interventions for the prevention of chemical dependency in children and adolescents. *Journal of Children in Contemporary Society, 18*, 117–151.
- Dembo, R., Williams, L., Wish, E., Dertke, M., Berry, E., Getreu, A., Washburn, M., & Schmeidler, J. (1988). The relationship between physical and sexual abuse and illicit drug use: A replication among a new sample of youths entering a juvenile detention center. *International Journal of the Addictions, 23*, 1101–1123.
- Dishion, T. J., McCord, J., & Poulon, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*(9), 755–764.
- Fox, D. M., & Gottfredson, D. C. (2002). Differentiating completers from non-completers of a family-based prevention program. *The Journal of Primary Prevention, 24*, 111–124.
- Gottfredson, D. C. (2001). *Schools and delinquency*. New York: Cambridge University Press.
- Gottfredson, D. C., Kumpfer, K., Fox, D., Beatty, P., & Wilson, D. (2004). *The Strengthening Washington D.C. Families Project: Relating implementation to program outcomes*. Paper presented at the annual meeting of the Society for Prevention Research, Quebec City.
- Gottfredson, G. D., & Gottfredson, D. C. (1999). *What about you?* Ellicott City, MD: Gottfredson Associates, Inc.
- Gottfredson, D. C., Harmon, M. A., Gottfredson, G. D., Jones, E. M., & Celestin, J. A. (1996). *Compendium of prevention program outcomes and instrument locator*. Ellicott City, MD: Gottfredson Associates, Inc.
- Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system manual*. Circle Pines, MN: American Guidance Service.
- Gutman, M. A., Foltz, C., Mittal, R., & Kaltenbach, K. (2004). *Outcomes of a family-based prevention model with women in substance abuse treatment and their children: The Philadelphia Strengthening Families Project*. Philadelphia: Treatment Research Institute (submitted for publication).
- Harrison, S., Proskauer, S., & Kumpfer, K. L. (1995). *Final evaluation report on Utah CSAP/CYAP project*. Submitted to the Utah State Division of Substance Abuse. University of Utah, Social Research Institute.
- Hawkins, J. D., Catalano, R. F., & Miller, J. L. (1992). Risk and protective factors for alcohol and other drug problems in early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112*, 64–105.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., & Schere, D. G. (1997). Multisystematic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65*, 821–833.
- Kameoka, V. A. (1996). *The effects of a family-focused intervention on reducing risk for substance abuse among Asian and Pacific-Island youths and families: Evaluation of the Strengthening Hawaii's Families Project*. Honolulu: University of Hawaii, Social Welfare Evaluation and Research Unit.
- Kandel, D. B., & Andrews, K. (1987). Processes of adolescent socialization by parents and peers. *International Journal of the Addictions, 22*, 319–342.
- Kellam, S. G. (1990). Developmental epidemiological framework for family research on depression and aggression. In G. R. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 11–48). Hillsdale, NJ: Lawrence Erlbaum.

- Kumpfer, K. L., & Alvarado, R. (2003). Family interventions for the prevention of drug abuse. In R. Weissberg & K. L. Kumpfer (Guest Eds.), *American Psychologist*, *58*, 457–465.
- Kumpfer, K. L., & DeMarsh, J. P. (1985). Prevention of chemical dependency in children of alcohol and drug abusers. *NIDA Notes*, *5*, 2–3.
- Kumpfer, K. L., DeMarsh, J. P., & Child, W. (1989). *Strengthening Families Program: Children's skill training curriculum manual (prevention services to children of substance-abusing parents)*. Salt Lake City, UT: Department of Health, Alcohol and Drug Research Center.
- Kumpfer, K. L., Alvarado, R., Smith, P., & Bellamy, N. (2002). Cultural sensitivity in universal family-based prevention interventions. In K. Kavanaugh, R. Spoth, & T. Dishion (Special Edition Eds.), *Prevention Science*, *3*, 241–244.
- Kumpfer, K. L., Alvarado, R., Tait, C., & Turner, C. (2002). Effectiveness of school-based family and children's skills training for substance abuse prevention among 6–8 year old rural children. *Psychology of Addictive Behaviors*, *16*(4S), 65–71.
- Kumpfer, K. L., Alvarado, R., Whiteside, H. O., & Tait, C. (2004). The strengthening families program (SFP): An evidence based, multi-cultural family skills training program. In J. Szapocznik, P. Tolan, & S. Sambrano (Eds.), *Preventing substance abuse* (pp. 3–14). Washington, DC: American Psychological Association.
- Lipsey, M. W., & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful intervention* (pp. 313–345). Thousand Oaks, CA: Sage.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Loeber, R., & Stouthamer-Loeber, M. (1986). Family factors as correlates and predictors of juvenile conduct problems and delinquency. In N. Morris & M. Tonry (Eds.), *Crime and justice: An annual review of research* (pp. 29–149). Chicago: University of Chicago Press.
- McMahon, R. J. (1999). Parent training. In S. Russ & T. H. Ollendick (Eds.), *Handbook of psychotherapies with children and families* (pp. 133–180). New York: Kluwer Academic/Plenum Publishers.
- Mooney, S. (1995). Parent training: A review of adlerian, parent effectiveness training, and behavioral research. *Family Journal*, *3*, 218–230.
- Moos, R. D., & Moos, B. S. (1986). *Family environment scales manual*. Palo Alto, CA: Consulting Psychologists Press, Inc.
- National Institute on Drug Abuse. (1997). *Preventing drug use among children and adolescents: A research-based guide* (DHHS, Publication No. ADM: NIH Publication No. 97-4212). Rockville, MD: U.S. Government Printing Office.
- Olds, D. L., Henderson, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., & Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior. *Journal of the American Medical Association*, *280*, 1238–1244.
- Polizzi, D. M., & Gottfredson, D. C. (2003). Differentiating completers from non-completers of a family-based prevention program. *Journal of Primary Prevention*, *24*(2), 111–124.
- Polizzi-Fox, D., Gottfredson, D. C., Kumpfer, K. K., & Beatty, P. D. (2004). Challenges in disseminating model programs: A qualitative analysis of the Strengthening Washington, DC Families Program. *Clinical Child and Family Psychology Review*, *7*(3), 165–176.
- Prinz, R. J., & Miller, G. E. (1996). Parental engagement in interventions for children at risk of conduct disorder. In R. D. Peters & R. J. McMahon (Eds.), *Preventing childhood disorders, substance abuse, and delinquency* (pp. 161–183). Thousand Oaks, CA: Sage.
- Rotheram-Borus, M. J., & Duan, N. (2003). Next generation of preventive interventions. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*, 518–526.
- Sanders, M. R. (1996). New directions in behavioral family interventions with children. In T. H. Ollendick & R. J. Prinz, et al. (Eds.), *Advances in clinical child psychology: Vol. 18* (pp. 283–330). New York, NY: Plenum Press.
- Sanders, M. R., Turner, K. M. T., & Markie-Dadds, C. (2002). The development and dissemination of the Triple P—Positive Parenting Program: A multi-level, evidence-based system of parenting and family support. *Prevention Science*, *3*, 173–198.
- Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy*, *27*, 171–186.
- Simcha-Fagan, O., Gersten, J. C., & Langner, T. S. (1986). Early precursors and concurrent correlates of patterns of illicit drug use in adolescence. *Journal of Drug Issues*, *16*, 7–28.
- Vicary, J. R., & Lerner, J. V. (1986). Parental attributes and adolescent drug use. *Journal of Adolescence*, *9*, 115–122.
- Vilmenay, M. (2002). Exploring barriers to at-risk family participation in the Washington, DC Strengthening Families Prevention Program. Unpublished Master's thesis. College Park, Maryland: University of Maryland.
- Webster-Stratton, C., & Taylor, T. (2001). Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0–8 years old). *Prevention Science*, *2*(3), 165–192.
- Whitbeck, L., & Smith, J. (2001). *Outcomes of a culturally-modified family program for Ojibway Indian families*. Paper presented at the annual meeting of the Society for Prevention Research, Washington, DC.
- Wilson, D. B. (2004). *Technical report: D.C. Strengthening Families Project: Growth curve analysis*. Technical report available from the author.

Query to Author

A1: Au: The alphabet H appearing in Table 3 has been changed to Superscript. a. Please check for correctness.