Published in final edited form as:

Eval Rev. 2011 April ; 35(2): 153–188. doi:10.1177/0193841X11403989.

Evaluation of Free to Grow: Head Start Partnerships to Promote Substance-Free Communities

Mark Wolfson¹, Heather Champion², Todd Rogers³, Rebecca H. Neiberg⁴, Dianne C. Barker⁵, Jennifer W. Talton⁴, Edward H. Ip⁴, Ralph B. D'Agostino Jr⁴, Maria T. Parries¹, and Doug Easterling¹

¹Department of Social Sciences and Health Policy, Wake Forest University School of Medicine, Winston-Salem, NC, USA

²Research, Innovation and Product Development, Center for Creative Leadership, Greensboro, NC, USA

³RTI International, San Francisco, CA, USA

⁴Department of Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA

⁵Barker Bi-Coastal Health Consultants, Inc., Calabasas, CA, USA

Abstract

Free to Grow: Head Start Partnerships to Promote Substance-free Communities (FTG) was a national initiative in which local Head Start (HS) agencies, in partnership with other community organizations, implemented a mix of evidence-based family-strengthening and community-strengthening strategies. The evaluation of FTG used a quasi-experimental design to compare 14 communities that participated in the FTG intervention with 14 matched comparison communities. Telephone surveys were conducted with two cohorts of the primary caregivers of children in HS at baseline and then annually for 2 years. The survey was also administered to repeated cross-sectional samples of primary caregivers of young children who were not enrolled in HS. No consistent evidence was found in changes in family functioning or neighborhood conditions when the 14 FTG sites were compared to 14 matched sites. However, caregivers of young children who were not in HS in three high-implementing FTG sites showed evidence of improvements in neighborhood organization, neighborhood norms against substance abuse, and child disciplinary practices. Results provide highly limited support for the concept that family and neighborhood conditions that are likely to affect child development and well-being can be changed through organized efforts implemented by local HS programs.

Reprints and permission: sagepub.com/journalsPermissions.nav

Corresponding Author: Mark Wolfson, Department of Social Sciences and Health Policy, Division of Public Health Sciences, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC27157, mwolfson@wfubmc.edu.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Keywords

community intervention; substance abuse; community partnerships; evaluation

Introduction

It has been well established that the contexts in which children are raised—including both the immediate family environment and the larger neighborhood environment—influence exposure to risk factors and the acquisition of protective factors (Hawkins, Catalano and Miller 1992; Hawkins, Van Horn and Arthur 2004). In recent years, the concept of strengthening families and communities has been incorporated into a variety of promising approaches to preventing substance abuse and child abuse and neglect. Research suggests that evidence-based parent education and skill-building programs can lead to improvements in parenting skills, parent/child communication, and the quality of the parent/child relationship (Kosterman et al. 2005; Houge et al. 2002, Spoth, Redmond and Shin 2001; Zubrick et al. 2005), all of which are important risk factors for substance use as children age. Most of these studies, however, focus on parents of pre-adolescents or adolescents; there is less evidence of the effectiveness of parenting programs aimed at the parents of pre-school children (an exception is Webster-Stratton 1998). A number of recent programs have used a combination of parenting education, skill building, and service linkages to strengthen families (Casey 2001; Harrell, Cavanagh, and Sridharans 1998). Although rigorous evaluation of community strengthening (CS) efforts to reduce individual- and community-level risk factors for substance abuse has a relatively short history (Sorensen et al. 1998) and demonstrated inconsistent findings (Hallfors, Cho and Livert 2002), some studies have shown that effectively mounted community policy, education, and enforcement initiatives can favorably affect rates of substance use and problems (Wagenaar et al. 2000; Holder et al. 1997; Hingson et al. 2005; Forster et al. 1998). Relatively little of this work has examined CS efforts focused on the environments of young children, especially young children in poor neighborhoods.

Free to Grow: Head Start Partnerships to Promote Substance-free Communities was a national demonstration initiative to implement a comprehensive, integrated approach to the prevention of substance abuse, child abuse and neglect, and other risk behaviors. Funded by the Robert Wood Johnson Foundation (RWJF) and the Doris Duke Charitable Foundation (DDCF), in partnership with the National Head Start Bureau, 15 Head Start (HS) programs collaborated with local schools, law-enforcement agencies, and other community partners to implement the Free to Grow (FTG) model. The FTG model emphasized the family and community context within which children live, rather than focusing solely on the immediate educational needs of the individual child. With support from the FTG National Program Office (NPO) at the Mailman School of Public Health at Columbia University, sites were expected to implement a mix of evidence-based family-strengthening and community-strengthening strategies and to infuse a set of FTG principles and practices within the HS agency and partner organizations.

As described elsewhere (Jones, Gutman, and Kaufman 1999; Wolfson et al. 2003; Wolfson and Easterling 2004; also see <http://www.freetogrow.org>), FTG was operationalized as a comprehensive set of locally tailored, integrated family- and community-strengthening strategies that address reduction of risk and enhancement of protective factors to serve as a buffer against substance use and abuse and other high-risk behaviors as children grow older. Sites were expected to implement a core set of family-strengthening strategies, including family assessment, case management, and substance abuse/mental health treatment. Additionally, sites were expected to implement a group-level intervention for supporting families in their development of skills and resources, such as family support and/or parent education.

With respect to community-strengthening strategies, the FTG model called for a leadership development approach that involved both HS and non-HS parents, as well as a multipronged strategy for improving environmental conditions and effecting policy change. The NPO recommended eight distinct strategies for environmental and policy change (community assessment, neighborhood-level awareness-raising, community-level awareness-raising and education, community action groups, community action, neighborhood revitalization, policy advocacy, policy change), from which sites could select a subset that suited local conditions and resources (for strategy definitions, see <http://www.freetogrow.org>).

Following a pilot phase in the 1990s that was used to develop the FTG model (Jones, Gutman and Kaufman 1999; Harrington et al. 2000), RWJF and the NPO solicited applications from all U.S. HS sites, resulting in 18 sites receiving funding for a 1-year capacity-building effort (June 2001–May 2002). On the basis of their capacity-building performance, 15 sites were selected to implement FTG over a 3-year period, starting June 2002.

Sites were allowed considerable latitude regarding strategy implementation, the specific curricula or protocols employed, and the sequencing of strategies. All sites were expected to integrate family and community strategies and to adopt FTG principles and practices regarding organizational development, strategic planning, and multisectoral collaboration with community partners. The FTG model directed funded HS agencies to develop a team of core community collaborators, including “mandated partners” representing local law enforcement and neighborhood schools. Moreover, the FTG model was designed to engage HS parents and other residents of low-income neighborhoods as change agents. A discussion of the implementation of the FTG model at three sites may be found in Wielawski 2006.

Throughout the initiative, the NPO offered a variety of technical assistance and training opportunities to guide and support sites in their planning and implementation of specific intervention strategies as well as general processes related to FTG principles and practices; for example, sites were provided with extensive training in family assessment, motivational interviewing, and neighborhood revitalization techniques, as well as assistance in collaborative leadership and sustainability planning. NPO support was provided through site visits, telephone meetings, e-mail exchanges, training workshops, and initiative-wide conferences.

In this article, we report the main program outcomes from the independent national evaluation of FTG, which was conducted by a team based at Wake Forest University School of Medicine.

Method

Overall Design

The evaluation of FTG employed a quasi-experimental design, in which 14 HS agencies and communities that participated in FTG (FTG sites) were compared on a variety of process and outcome measures with 14 matched HS agencies and communities (comparison sites). (A 15th FTG site, a Native American community in New Mexico, was treated primarily as a case study; outcomes are not reported in this article.)

In order to identify potential comparison sites, 2000 HS Program Information Report (PIR) data (United States Department of Health and Human Services [USDHHS] 2001) and 2000 Census data (US Census Bureau 2000) were used to generate a list of all HS agencies in the 14 states containing FTG sites. We used Mahalanobis metric matching to estimate multivariate distances between each FTG site and all potential comparison sites within each state as a function of several contextual variables that could be related to the outcomes to be measured. It has been shown that Mahalanobis metric matching is an effective technique for achieving balance between conditions at baseline in nonrandomized evaluations (Rubin 1976; Rosenbaum and Rubin 1985; D'Agostino 1998). Variables for matching were constructed using PIR and 2000 Census data and included population density of the city in which the HS was located, percentage of HS enrollment that used the center-based program option, whether the organization administering HS was a community action agency, total enrollment, race/ethnicity of enrolled children, presence/absence of an early HS program, and education level of the family and community partnership services expert. Additional Mahalanobis distances were calculated excluding population density and/or education level from the models since these two variables were missing for several potential comparison sites. There were no significant differences between the FTG and comparison sites on any of the matching variables.

Data Collection

The impact of FTG on families and communities was assessed through a telephone survey administered to children's primary caregivers (PCs). FTG's family- and community-strengthening efforts were aimed at families with children in HS, as well as other families with young children in the target neighborhood for each site. Consequently, the Primary Caregiver Survey (PCS) was administered to independent samples of HS and non-HS families in the FTG sites and matched comparison sites. Moreover, because FTG was not being fully implemented in all sites during the first year of implementation, a second cohort of HS families was recruited for participation in the PCS. Thus, three separate samples of PCs were used to assess the impact of FTG:

Cohort samples—To be eligible for the cohort samples, PCs had to have children enrolled in HS. In the FTG sites, the families had to live in the neighborhoods targeted for

FTG interventions, with boundaries provided by the FTG sites. Comparison sites were asked to provide comparable neighborhood boundaries surrounding their HS facility that most closely matched the FTG site based on size, demographics, and/or language. While neighborhoods were administratively identified by the HS agency in both FTG and comparison sites, they usually corresponded to “organic” neighborhoods that residents would describe as such. To participate in the PCS, families had to live in these neighborhoods. Additionally, in Cohort 2, the PC could *not* have a child enrolled in HS or early HS in the prior year. Lists of eligible participants were provided by the HS sites. Additional screening was done when the survey was conducted to confirm eligibility. All eligible participants were invited to participate in the survey.

Each cohort was interviewed annually for 3 years. Cohort 1 (C1) respondents were called for the first time between September 2002 and January 2003, and then during these same months in 2003–2004 and 2004–2005. Cohort 2 (C2) respondents were called for the first time between September 2003 and January 2004, and then during these same months in 2004–2005 and 2005–2006. In order to be eligible to participate at Time 2 or 3, a baseline survey had to have been completed. A PC who did not complete an interview at Time 2 was still eligible to be interviewed at Time 3. Participants who specifically requested not to be called again were removed from the sample.

The PCS was administered in English and Spanish by trained interviewers at the University of South Carolina (USC) Survey Research Laboratory. If 15% or more of the eligible participants at a given site spoke another language, the PCS was offered in that language. Additional subcontractors were used to conduct the survey in Haitian Creole and Hmong. The time required for completion of the surveys was 17–23 min (English), 16–55 min (Spanish), 27–66 min (Hmong), and 33–57 min (Haitian Creole). Due to quality control issues in data collection for the 2003 and 2004 Haitian Creole surveys, interviews conducted in this language were not included in the analyses presented in this article.

A number of strategies were employed to maximize response rates and retention in the two cohorts in this low-income, mobile sample of HS PCs (Champion et al. 2005). Gift card incentives were provided to individuals who completed the PCS (\$10 in 2002 and 2003; \$15 in 2004 and 2005). Additionally, multiple interim contacts with participants were made, including a brief 6-month follow-up interview and prenotification letters prior to both the 6-month interview and the PCS. Stickers, tattoos, and a refrigerator magnet were also sent to families, who could use a toll-free number to update their contact information. Additional strategies used to maximize retention included directory assistance and Internet searches to locate participants, and incentivizing HS staff and PCS participants for updating contact information.

The response rates for Cohort 1 were 66.86% at Time 1, 64.95% at Time 2, and 70.97% at Time 3; the overall retention rate was 53.85%. The response rates for Cohort 2 were 67.68% at Time 1, 70.59% at Time 2, and 70.18% at Time 3; the overall retention rate was 54.13% (see Table 1 for details on how response rates were calculated).

Our analysis of the impact of attrition revealed that participants who completed only one or two waves of the PCS did not show significant differences on baseline risk levels compared to participants who completed all three waves (data not shown; construction of the Risk Index is described below). Thus, between-group comparisons should be valid.

Community sample—In the community sample, PCs had to live in the neighborhoods targeted by the FTG sites or in comparable neighborhoods in the comparison sample. PCs also had to have children ages 3–8 years old (inclusive) who were not enrolled in HS. Using the defined boundaries, a random sample was selected from listed residential telephone numbers located within the target area (the sampling frame of age-targeted telephone numbers in each area was provided by Survey Sampling, Inc., of Westport, Connecticut). Calls were made by the USC survey center to sampled numbers to determine eligibility; once eligibility was established the interviewer proceeded with the interview protocol. The target sample size was 50 participants per community. Community sample PCs surveyed in January–June 2003 served as a baseline group. A second cross-sectional sample of PCs was surveyed in January–June 2006 to form a follow-up group.

The baseline PCS community sample consisted of 1,110 PCs, and the follow-up PCS community sample consisted of 1,093 PCs (no response rate is calculated because the size of the underlying eligible population is not known).

Measures

The PCS was used to measure demographic characteristics of the PC and the PC's family, as well as a series of neighborhood constructs, family constructs, and measures of substance use by the PC and the PC's family. Secondary data, including the HS PIR and Census data, were used to construct several site-level covariates used in the analyses. A measure of level of implementation of FTG, which was used in a site-level dose analysis, is described below.

Neighborhood constructs—Seven constructs were used to assess the impact of FTG on neighborhoods: Neighborhood involvement, school involvement, social connectedness, neighborhood organization, neighborhood availability of alcohol and illegal drugs, neighborhood norms against substance use, and satisfaction with police enforcement. All constructs were coded so that higher scores indicate better outcomes. Details on the construction of these constructs are provided in Table 2.

Family strengthening (FS) constructs—Six constructs were used to measure the impact of FTG on families: Family conflict, parenting frustration, physical discipline: hitting, physical discipline: spanking, use of rules, and time out (see Table 2).

Family substance use constructs—Four constructs were used to measure the impact of FTG on substance use attitudes and behavior within the PC's household: PC alcohol use, PC & household substance use, family norms—drinking, and family norms—getting drunk (see Table 2).

Site-level covariates—Site-level variables used as covariates in the analyses were derived from 2000 HS PIR and 2000 U.S. Census data. Information on HS enrollment size,

the education level of the family service worker in the HS, whether the HS was part of a community action agency, and the presence of early HS was obtained from the 2001 PIR data set. HS enrollment was included as a continuous variable. Education level of the family service worker was measured on a 5-point scale (<high school; high school graduate; associate degree/2 years college; college graduate; graduate degree). Population density, per capita income, percentage below poverty level, percentage less than high school education, percentage employed, and percentage owner occupied housing for the target neighborhood of each site were obtained from 2000 Census data. Z scores of percentage below poverty level, percentage less than high school education, and percentage employed were summed to comprise a socioeconomic status indicator.

Individual-level covariates—Individual-level variables used as covariates included child's gender; Risk Index score for the family (described below); age, race (Black, White, and other), and Hispanic ethnicity (yes/no) of the PC; number of children in the home (1, 2, 3, 4+); presence of an infant (<24 months) in the household (yes/no); family structure (single-parent home vs. more than one adult relative in home); and relationship of the PC to the HS child (mother vs. other). Age and number of children in the home were measured as continuous variables. Education level of the PC and household income were not included as individual-level covariates because they are reflected in the Risk Index. Individual-level variables that were allowed to vary with time in order to reduce missing values at baseline were age, race, Hispanic ethnicity, number of children in the home, presence of an infant, family structure, and relationship of the PC to the HS child.

The Risk Index was used as an individual-level covariate and as a stratification factor in the analyses presented below. The index was created using items from the PCS and was based on the Audubon Area Head Start Family Assessment Scoring tool, which was promoted by the NPO to be used or adapted by FTG sites in implementing the family assessment strategy (Boden and Levine 2000). The index was constructed using items on the PCS that were the same as items on the Audubon tool, or a close proxy for them, including education, employment, housing, services and resources, special needs/family support, family wellness/substance use, family finances, child care, and parenting. (One construct on the Audubon tool, transportation, did not have comparable items on the PCS.) To be consistent with the Audubon assessment, weights were used so that the individual construct scores could range from 1 to 5 and the total cumulative score could range from 10 to 50, with higher scores indicating higher risk.

Comparisons of FTG and comparison group demographic characteristics at baseline for C1 and C2, and the community sample, are presented in Table 3. For C1 and C2, there were significant differences ($p < .05$) between the FTG and comparison samples on race, Hispanic ethnicity, and Risk Index score. Participants in the FTG sites were more likely to be White, Hispanic, and score higher on the Risk Index. While the magnitude of the differences by condition for the Risk Index was modest, the differences by race and Hispanic ethnicity were considerably more substantial. We adjust for these differences in model estimation in order to minimize any bias due to baseline differences. There were no statistically significant differences between the FTG and comparison groups for any of these variables in

the community sample. At baseline, there were no significant differences ($p < .05$) between FTG and comparison sites on any of the site-level variables.

Analysis

Three approaches were used in the analyses, each addressing a somewhat different question. The main effects analysis examines whether PCs in the FTG sample, as a group, reported better outcomes over time compared to PCs in the comparison sample.

In addition to the main effects analysis, we examined whether families with different risk profile scores (measured at baseline) responded differently to the FTG intervention. Previous studies have found that the response to a preventive intervention may vary dramatically by baseline risk level (Olds et al. 1997; Wilson, Lipsey and Derzon 2003). The FTG approach employs a range of family- and community-strengthening strategies; an individual HS parent's exposure to any particular strategy is likely to vary by risk level, because of either self-selection or the way in which the strategy is targeted by an implementing site. For example, people at the highest risk at baseline may be most likely to participate in intensive case management, while individuals at the lowest risk at baseline may be most likely to participate in leadership training.

The final set of analyses examined the possibility that the main effects analysis might understate the effectiveness of the intervention because of uneven implementation across the FTG sites. We evaluated whether there was greater improvement in the outcome variables within those sites that fully implemented the FTG model, compared to the change that occurred within both the comparison sites and the low-implementing FTG sites. Essentially, we assessed whether site-level dose (SLD) had a moderating effect on the outcome variables (with higher levels of SLD associated with greater change in the outcomes; see Weitzman et al. 2004).

Each of these analytic approaches was applied to the three samples of caregivers described previously: Cohort 1, Cohort 2, and the community sample. In the analysis of the Cohort 1 and Cohort 2 samples, weights were applied to adjust for unequal sampling probabilities at each site (the weight used was the inverse of the probability that a PC at the site would be selected, with the probability calculated as $P(\text{selection}) = \text{number of PCs sampled} / \text{number of PCs with a child enrolled in the site's HS program}$). Weighting was not feasible in the analysis of the community sample, as we did not have information about the number of household with young children in each community. Two-tailed tests of significance were used. We use a conservative α level of $p < .001$, based on a Bonferroni adjustment, for statistical significance.

Main effects analysis—We used a stratified, nested, repeated measures approach to assess the effects of FTG on PCs of children in HS. Because measurements on parents in the same neighborhoods are associated with the same HS and sites within the same state are likely to be correlated, it was appropriate to use a model that includes fixed effects and time as well as features for handling clustering effects for state, community, and repeated measures on a PC (Murray 1998).

For all outcomes, we used a population-averaged approach that implements Generalized Estimating Equations (GEE) models to analyze the data (Liang and Zeger 1986). All analyses were performed using SAS callable SUDAAN (2006; SAS Institute, Cary, North Carolina).

Because some observations contained partially missing values for the multiple items used to construct the Risk Index, a complete-case approach would have discarded all partially missing cases. To avoid a wasteful approach, a multiple imputation procedure was used in conjunction with the GEE models (Rubin 1987). Observations with more than five missing items for calculating the Risk Index were removed. For observations with at least one but no more than five missing items for scoring the Risk Index, values were imputed through a predictive model based on a missing at random (MAR) assumption.

The models included all of the site- and individual-level covariates described in the previous section as fixed effects: number of children in the home; presence of an infant in the home; family structure; child's gender; Risk Index score; and PC age, race, Hispanic ethnicity, and relationship to HS child. Site/neighborhood covariates included whether the HS was part of a community action agency, education level of the family service worker, HS enrollment size, socioeconomic status (SES) factor score, percentage of housing occupied by owner, presence of an early HS program at the site, and population density and per capita income of the FTG or comparison neighborhood.

We use a stratified, repeated measures approach to assess the intervention effects of FTG in the community sample. Because measurements on parents in the same neighborhoods are associated with the same HS and sites within the same state are likely to be correlated, it is appropriate to use a model that handles the clustering effects between observations within the same neighborhood, which are nested within the same HS and state.

As in the cohort analyses, GEE models with multiply imputed missing values for the Risk Index were employed where the Condition by Time interaction is of interest. The same individual- and site-level covariates were used as in the Cohort analyses.

Analysis of the impact of FTG stratified by risk score—For the C1 and C2 samples, we used the Risk Index to stratify the sample into low-risk (i.e., those with a Risk Index score of 10–15) and high-risk groups (i.e., those with a Risk Index score 21 and above). Separate analyses were conducted for each of these two groups, using the same statistical methods employed in the main effects analyses. The models included all of the site- and individual-level covariates (with the exception of the Risk Index) included in the main effects analyses, and the same weighting scheme was used (described above).

Similar to the nonstratified analysis, we used a stratified, repeated measures approach to assess the intervention effects of FTG in the community sample.

Site-level dose analysis—Following the approach used by Weitzman and colleagues (Weitzman et al. 2004), we divided the sample of 28 sites into 3 treatment groups: (a) high-implementing FTG sites, (b) low-implementing FTG sites, and (c) comparison sites. Using SUDAAN, we tested the significance of the Treatment by Time interaction, controlling for

individual-level and site-level covariates. Separate analyses were conducted for each of the two cohort samples and for the cross-sectional community sample.

The 14 FTG sites were classified into high versus low implementers based on an expert-rating approach (Easterling et al. 2007). Ratings were developed for both the second and third years of implementation of FTG. Members of the evaluation team who had conducted site visits and who were familiar with the sites rated the degree to which each site had implemented the 14 different strategies specified within the FTG model (i.e., 14 ratings per site). Ratings were based on a 1–5 scale, with 1 = *no more than would be expected of a typical Head Start*, and 5 = *exceptional example of the FTG model*. Raters had access to a comprehensive data profile which summarized each site's standing on key measures using a variety of process evaluation protocols (Rogers, Easterling, and Barker 2006). Where there was notable disagreement between raters (standard deviation greater than 1), the raters entered into a “consensus conference” where they explained their ratings to one another, and then assigned revised ratings.

For the purposes of classifying sites by level of implementation, ratings were averaged across raters to generate 14 strategy-specific scores for each site. The 14 strategy-specific scores were summarized into 2 overall rating scores for each site, ranging from 1 to 5: (a) strength of the FS approach and (b) strength of the CS approach. The FS and CS scores derived from the evaluation team's ratings were compared against ratings that were independently generated by the NPO (again using a 1–5 scale). The correlation was .54 for the FS scores and .76 for the CS scores. Recognizing that the NPO and the evaluation team each brought unique knowledge to these ratings, the two sets of ratings were averaged, producing an overall FS score and an overall CS score for each FTG site (each ranging from 1 to 5). A site was classified as a “high implementer” if both its FS score and its CS score were at least 3.5. Based on this rather stringent criterion for evaluating implementation, 3 of the 14 FTG sites were classified as high implementers during Implementation Year 2. The same three sites came out as “high” based on the Year 3 ratings.

Results

Main Effects Analysis

Using the criterion of $p < .001$, no differences were detected between the FTG and the comparison samples from baseline to follow-up for any of the neighborhood, family, or substance-abuse measures (see Table 4).

Analysis of the Impact of FTG Stratified by Risk Index Score

No Time by Condition differences were observed among C1 and C2 high-risk participants on the 17 neighborhood and family and substance use constructs (see Table 5). Among the community sample high-risk participants, the neighborhood norms against substance use variable was significant in the Time by Condition analyses ($p = .0001$). Reported norms improved over time among participants in the FTG communities and worsened among participants in the comparison sites.

No Time by Condition differences were observed among low-risk participants in the C1, C2, and community samples (see Table 6).

Site-Level Dose Analysis

The analyses summarized in Table 7 test whether the three high-implementation sites showed greater positive change in PCS scores than what occurred in either the 11 low-implementing FTG sites or the 14 comparison sites (i.e., whether there was a significant interaction between Time and a Treatment Factor that classifies sites into FTG high, FTG low, and comparison conditions).

Within the two cohorts of HS parents and caregivers, there were no significant Time by Treatment interactions. In contrast, analysis of the cross-sectional community sample data revealed that FTG-high sites improved on some outcomes between 2003 and 2006. As shown in the last set of columns in Table 7, respondents in the FTG-high sites reported higher levels of neighborhood organization in 2006 than in 2003 (adjusted mean increased from 16.79 to 17.32), whereas values on this scale decreased over time in the FTG-low (18.54 to 17.95) and comparison sites (19.59 to 19.00; Time $\times$ Treatment interaction $p < .0001$). A similar pattern was found for neighborhood norms against substance use ($p < .0001$). These findings suggest that the CS strategies carried out in the three FTG-high sites had a measurable effect on neighborhood conditions as perceived by a representative community sample of parents of young children.

This sample of parents in the broader community also reported improvements in a family outcome: Physical discipline—spanking. The adjusted proportion of caregivers who reported that they spanked their child decreased from 60% to 47% in the FTG-high sites, compared to minimal changes in the FTG-low (56% to 50%) and comparison sites (63% to 60%; interaction $p = .0001$).

Discussion

FTG was an ambitious national demonstration program aimed at strengthening the families and neighborhoods of children in HS. In the national evaluation of FTG, we found no evidence that the initiative produced statistically significant changes in family functioning or neighborhood conditions when the 14 FTG sites were compared to 14 matched sites, as measured by a survey of PCs. Similarly, our analyses of subgroups of families, stratified by risk profiles at baseline, did not produce consistent evidence of impact—with improvements in only one of the outcome measures—neighborhood norms against substance abuse—and this limited to the community sample.

Our site-level dose analysis did, however, find some evidence of impact within the three sites that were most successful in implementing the FTG model. This evidence emerged from two cross-sectional surveys of caregivers in the HS communities but was not replicated within the longitudinal surveys of HS caregivers in the two cohorts. Within the community sample, the high-implementing FTG sites showed significantly greater improvements than the lower-implementing FTG sites or the comparison sites on two neighborhood conditions (neighborhood organization and neighborhood norms against substance abuse) and one

family-functioning measure (spanking). The changes in values on the scales correspond to a 25%–39% change in standard deviation units for these items. Moreover, the 7%–10% greater decline in reports of spanking in the high-implementing FTG sites, compared (respectively) to the low-implementing FTG and comparison sites, appears to be a meaningful level of change.

Is it possible to generate a compelling explanation for the pattern of results from the site-level dose analysis: no effects in either of the HS caregiver samples, but positive effects in the community sample? One possible explanation for the observed results—particularly for the neighborhood outcomes—is that the HS cohorts continued to be surveyed, even when the caregiver moved out of the neighborhood. In contrast, both waves of the community sample were drawn from the target neighborhoods and thus may have provided a more accurate assessment of how the target neighborhood had changed over the 4 years of the initiative.

In interpreting the community sample and cohort results from the site-level dose analyses, it is also important to recognize how these two populations of caregivers differ from one another. Participants in the community sample, as a group, had lower family risk levels than did participants in the HS sample (see Table 3). Caregivers in the community sample may have been more receptive to several of the community-strengthening strategies in the FTG model, including leadership development, community action, and neighborhood revitalization. This would suggest that implementation of FTG CS strategies, not delivery of programs that were focused primarily, if not exclusively, on HS families, may ultimately have been what produced any neighborhood-level impact in the high-implementing sites.

Evidence that the FTG program's impact was limited to the subset of highest performing sites suggests that incomplete, delayed, or inadequate implementation may have taken place in the other FTG sites (Lipsey and Cordray 2000). Moreover, this pattern may suggest fundamental flaws in the program model, or the way it was communicated to, and internalized, by the sites. The FTG model was quite complex, involving elements of organizational change, formation of genuine partnerships with diverse community organizations, and planning and delivery of integrated family- and community-strengthening strategies. Complex models, such as FTG, are difficult to define and communicate clearly to site leaders and staff, and may be inherently too challenging for many organizations to internalize and execute effectively (Szulanski 2003). This is especially true when the host organization is rooted in a culture with minimal connection to aspects of the model's principals and practices that are derived from a different culture. In FTG, HS agencies' commitment to classroom-based, early childhood education may have made it challenging for some agencies to engage fully and effectively in community-oriented efforts for wide-scale social and environmental change (Wielawski 2006).

Our evaluation of FTG had a number of limitations. First, we used a quasi-experimental design. Although the use of Mahalanobis metric matching enabled us to select a group of comparison sites similar to the FTG sites on most dimensions, there were still a number of significant differences at baseline between the FTG and comparison samples of PCs. However, we adjusted for these differences in model estimation, so baseline differences

should not bias the assessment of intervention effects. Second, we relied on self-reports of caregivers to assess changes in their families and in the surrounding neighborhood. While objective indicators of changes in neighborhood conditions are often desirable for the assessment of efforts aimed in whole or in part at community change, the varied nature of the FTG intervention across multiple sites—and corresponding variance in the neighborhood outcomes that it might produce—hampered our ability to design, preintervention, a comprehensive set of neighborhood measures. Additionally, community outcomes pursued by FTG sites were targeted at geopolitical levels lower than those readily detected by standard surveillance methods. Finally, the evaluation lacked adequate measures of individual-level exposure to specific FTG family- or community-strengthening interventions, precluding individual-level dose analyses which may have detected additional FTG effects.

FTG is one of several recent comprehensive community initiatives with broad and diffuse program strategies and goals (Wagner et al. 2000; O'Connor 1995; Kreuter, Lezin and Young 2000; Billing, 2004). The results of the national evaluation of FTG have implications for the developers, funders, implementers, and evaluators of such initiatives. First and foremost, it is important to consider our finding—and one that is likely to obtain in similar initiatives—that only a subset of sites can implement a program model with enough intensity, comprehensiveness, and fidelity to show measurable changes in families and communities (Hingson et al. 2005; Weitzman et al. 2004). In concept, there are a number of ways in which this shortcoming could be addressed: (a) the program model could be made easier to implement (e.g., by making it more prescriptive); (b) more central guidance and support could be provided; or (c) more selectivity could be exercised at the time of site selection, provided that factors associated with the capacity to be a high implementer can be identified and accurately assessed in prospective sites. Each of these approaches, however, has its own set of challenges.

In summary, the results presented here provide highly limited support for the concept that family and neighborhood conditions that are likely to affect child development and well-being can be attained through organized change efforts implemented by local HS programs.

Acknowledgments

We would like to offer our sincere thanks and appreciation to Ms. Judith Jones, Ms. Lori Levine, and other staff of the former Free to Grow National Program Office at the Mailman School of Public Health at Columbia University. We also offer our thanks to Dr. Richard Catalano and the other members of the Free to Grow Evaluation Advisory Committee, Francie Zimmerman of the Doris Duke Charitable Foundation, and our funders at the Robert Wood Johnson Foundation and the Office of Juvenile Justice and Delinquency Prevention. Tracy Orleans, our project officer at RWJF, deserves a special note of thanks for her support and guidance. We also appreciate the efforts of Dr. Robert Oldendick, director of the Survey Research Laboratory at the University of South Carolina. Finally, we offer our thanks to the staffs of the Head Start agencies at the participating Free to Grow and comparison sites, whose active support and participation made the evaluation possible.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Support for this research was provided by the Robert Wood Johnson Foundation (Grant No. 044123) and the Office of Juvenile Justice and Delinquency Prevention (Grant No. 2002-DR-FX-0001).

References

- Billing, D. Appalachian poverty. In: Mink, G.; O'Conner, A., editors. *Poverty and Social Welfare in America: An Encyclopedia*. Santa Barbara, CA: ABC-CLIO; 2004.
- Boden, L.; Levine, L. [accessed September 14, 2009] Audubon area HeadStart: Appendix G. The Kentucky Free to Grow Model 2006. 2000. http://www.freetogrow.org/usr_doc/appendixG1.pdf
- Casey Family Programs and the U.S. Department of Health and Human Services. *Starting Early Starting Smart: Summary of Early Findings*. Washington, DC: Casey Family Programs and the U.S. Department of Health and Human Services, Substance Abuse, and Mental Health Services Administration; 2001.
- Champion, H.; Parries, M.; Shrestha, A.; Kidd, K.; Martin, W. Retaining High-Risk, Mobile and Multi-Lingual Head Start Survey Participants in the National Evaluation of Free to Grow: Head Start Partnerships to Promote Substance-Free Communities. Paper presented at the Annual Meeting of the American Evaluation Association; Toronto, Canada. 2005.
- D'Agostino RB Jr. Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group. *Statistics in Medicine*. 1998; 17:2265–81. [PubMed: 9802183]
- Easterling, D.; Rogers, T.; Neiberg, RH.; D'Agostino, RB, Jr. Balancing Validity and Efficiency in the Assessment of Site-Level Dose. Paper presented at the Annual Meeting of the American Evaluation Association; Toronto, Canada. 2007.
- Forster JL, Murray DM, Wolfson M, Blaine TM, Wagenaar AC, Hennrikus DJ. The Effects of Community Policies to Reduce Youth Access to Tobacco. *American Journal of Public Health*. 1998; 88:1193–8. [PubMed: 9702146]
- Hallfors D, Cho H, Livert D. Fighting Back Against Substance Abuse: Are Community Coalitions Winning? *American Journal of Preventative Medicine*. 2002; 23:237–45.
- Harrell AV, Cavanagh SE, Sridharans S. Impact of the Children at Risk Program Final Report II 12-10-2007. 1998
- Harrington, Mary; Perez-Johnson, Irma; Meckstroth, Alicia; Bellotti, Anne; Love, John M. Final Report. Mathematica Policy Research, Inc; 2000. *Protecting Children From Substance Abuse: Lessons From Free to Grow Partnerships*.
- Hawkins JD, Catalano RF, Miller JY. Risk and Protective Factors for Alcohol and Other Drug Problems in Adolescence and Early Adulthood: Implications for Substance Abuse Prevention. *Psychology Bulletin*. 1992; 112:64–105.
- Hawkins JD, Van Horn ML, Arthur MW. Community Variation in Risk and Protective Factors and Substance Use Outcomes. *Prevention Science*. 2004; 5:213–20. [PubMed: 15566047]
- Hingson RW, Zakocs RC, Heeren T, Winter MR, Rosenbloom D, DeJong W. Effects on Alcohol Related Fatal Crashes of a Community Based Initiative to Increase Substance Abuse Treatment and Reduce Alcohol Availability. *Injury Prevention*. 2005; 11:84–90. [PubMed: 15805436]
- Holder HD, Saltz RF, Grube JW, Treno AJ, Reynolds RI, Voas RB, Gruenewald PJ. Summing Up: Lessons From a Comprehensive Community Prevention Trial. *Addiction*. 1997; 92:S293–301. [PubMed: 9231452]
- Houge A, Liddle HA, Becker D, Johnson-Leckrone J. Family-Based Prevention Counseling for High-Risk Young Adolescents: Immediate Outcomes. *Journal of Community Psychology*. 2002; 30:1–22.
- Jones, Judith E.; Gutman, Marjorie A.; Kaufman, Nancy J. Free to Grow: Translating Substance Abuse Research and Theory into Preventive Practice in a National Head Start Initiative. *The Journal of Primary Prevention*. 1999; 19:279–96.
- Kosterman, Rick; Hawkins, JD.; Spoth, Richard; Haggerty, Kevin; Zhu, Kangmin. Effects of a Preventive Parent-Training Intervention on Observed Family Interactions: Proximal Outcomes From Preparing for the Drug Free Years. *Journal of Community Psychology*. 2005; 25:337–52.
- Kreuter MW, Lezin NA, Young LA. Evaluating Community-Based Collaborative Mechanisms: Implications for Practitioner. *Health Promotion Practice*. 2000; 1:47–61.
- Liang KY, Zeger S. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*. 1986; 73:13–22.

- Lipsey MW, Cordray DS. Evaluation Methods for Social Intervention. *Annual Review of Psychology*. 2000; 51:345–75.
- Murray, DM. Design and Analysis of Group-Randomized Trials. New York, NY: Oxford University Press; 1998.
- O'Connor, A. Evaluating comprehensive community initiatives: A view from history. In: Connell, JP.; Kubisch, AC.; Schorr, LB.; Weiss, CH., editors. *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*. Washington, DC: The Aspen Institute; 1995.
- Olds DL, Eckenrode J, Henderson CR Jr, Kitzman H, Powers J, Cole R, Sidora K, Morris P, Pettitt LM, Luckey D. Long-Term Effects of Home Visitation on Maternal Life Course and Child Abuse and Neglect. Fifteen-Year Follow-Up of a Randomized Trial. *Journal of the American Medical Association*. 1997; 278:637–43. [PubMed: 9272895]
- Rogers, T.; Easterling, D.; Barker, D. Assessing Program Implementation in the National Evaluation of “Free to Grow”. Paper presented at the Annual Meeting of the American Evaluation Association; Portland, Oregon. 2006.
- Rosenbaum PR, Rubin DB. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *The American Statistician*. 1985; 39:33–8.
- Rubin, D. Multiple Imputation for Nonresponse in Surveys. New York, NY: Wiley; 1987.
- Rubin DB. Multivariate Matching Methods That are Equal Percent Bias Reducing, I: Some Examples. *Biometrics*. 1976; 32:109.
- Sorensen G, Emmons K, Hunt MK, Johnston D. Implications of the Results of Community Intervention Trials. *Annual Review Public Health*. 1998; 19:379–416.
- Spoth RL, Redmond C, Shin C. Randomized Trial of Brief Family Interventions for General Populations: Adolescent Substance Use Outcomes 4 Years Following Baseline. *Journal of Consulting and Clinical Psychology*. 2001; 69:627–42. [PubMed: 11550729]
- Szulanski, G. *Sticky Knowledge*. London, UK: SAGE; 2003.
- U.S. Census Bureau. [accessed September 14, 2009] 2000 Census of Population and Housing, Summary Tape File 3A. 2002. 2006<http://factfinder.census.gov/servlet/DatasetMainPageServlet>
- USDHHS. Project Head Start Program Information Report. Washington, DC: Administration on Children, Youth and Families; 2001.
- Wagenaar AC, Murray DM, Gehan JP, Wolfson M, Forster JL, Toomey TL, Perry CL, Jones-Webb R. Communities Mobilizing for Change on Alcohol: Outcomes From a Randomized Community Trial. *Journal of Studies on Alcohol*. 2000; 61:85–94. [PubMed: 10627101]
- Wagner EH, Wickizer TM, Cheadle A, Psaty BM, Koepsell TD, Diehr P, Curry SJ, Von Korff M, Anderman C, Beery WL, et al. The Kaiser Family Foundation Community Health Promotion Grants Program: Findings From an Outcome Evaluation. *Health Services Research*. 2000; 35:561–89. [PubMed: 10966086]
- Webster-Stratton C. Preventing Conduct Problems in Head Start Children: Strengthening Parenting Competencies. *Journal of Consulting and Clinical Psychology*. 1998; 66:715–30. [PubMed: 9803690]
- Weitzman ER, Nelson TF, Lee H, Wechsler H. Reducing Drinking and Related Harms in College: Evaluation of the “A Matter of Degree” Program. *American Journal of Preventive Medicine*. 2004; 27:187–96. [PubMed: 15450630]
- Wielawski, I. Free to Grow. In: Isaacs, S.; Knickman, J., editors. *To Improve Health and Health Care*. Vol. IX. Princeton, NJ: The Robert Wood Johnson Foundation; 2006. p. 3-27.
- Wilson SJ, Lipsey MW, Derzon JH. The Effects of School-Based Intervention Programs on Aggressive Behavior: A Meta-Analysis. *Journal of Consulting and Clinical Psychology*. 2003; 71:136–49. [PubMed: 12602434]
- Wolfson, M.; Foley, K.; Altman, D.; Barker, D.; Brittain, S.; Champion, H.; D’Agostino, RB., Jr; DuRant, R.; Rogers, T.; Williams, A. Free to Grow 2003 Annual Report. Winston-Salem, NC: Wake Forest University School of Medicine; 2003.
- Wolfson, M.; Easterling, D. Challenges in Evaluation of a National Demonstration Program to Promote Substance-Free Communities in Partnership With Head Start. Paper presented at Head Start’s 7th National Research Conference; Washington, DC. 2004.

Zubrick SR, Ward KA, Silburn SR, Lawrence D, Williams AA, Blair E, Robertson D, Sanders MR. Prevention of Child Behavior Problems Through Universal Implementation of a Group Behavioral Family Intervention. *Prevention Science*. 2005; 6:287–304. [PubMed: 16160760]

Biographies

Mark Wolfson, PhD, is Professor in the Department of Social Sciences and Health Policy at Wake Forest University School of Medicine, Winston-Salem, NC.

Heather Champion, PhD, is an Enterprise Associate at the Center for Creative Leadership in Greensboro, NC.

Todd Rogers, PhD, is Senior Scientist with RTI International in San Francisco, CA.

Rebecca H. Neiberg, MS, is a Biostatistician in the Department of Biostatistical Sciences at Wake Forest University School of Medicine, Winston-Salem, NC.

Dianne C. Barker, MHS, founded and runs Barker Bi-Coastal Health Consultants in Calabasas, California.

Jennifer W. Talton, MS, is a Biostatistician in the Department of Biostatistical Sciences at Wake Forest University School of Medicine, Winston-Salem, NC.

Edward H. Ip, PhD, is Professor in the Department of Biostatistical Sciences at Wake Forest University School of Medicine, Winston-Salem, NC.

Ralph B. D'Agostino, Jr., PhD, is Professor in the Department of Biostatistical Sciences at Wake Forest University School of Medicine. He is also the Director of the Biostatistics Core of the Comprehensive Cancer Center at WFUSM.

Maria T. Parries, is an Associate Project Manager in the Department of Social Sciences and Health Policy at Wake Forest University School of Medicine, Winston-Salem, NC.

Doug Easterling, PhD, is Associate Professor and Chair of the Department of Social Sciences and Health Policy at Wake Forest University School of Medicine, Winston-Salem, NC.

Table 1

Cohort 1 and Cohort 2 Primary Caregiver Survey (PCS) Response Rates^a

	Time 1			Time 2			Time 3			Time 1-3 Retention Rate	
	Sample Size	# Completions	Response Rate	Sample Size	# Completions	Response Rate	Sample Size	# Completions	Response Rate		
Cohort 1	2,870	1,919	66.86	1,900	1,234	64.95	1,223	868	70.97	53.85	
Cohort 2	1,411	955	67.68	945	667	70.59	654	459	70.18	54.13	

^a Response rates at Time 1 were calculated by dividing the number of completions by the total number of eligible participants received from enrollment lists from the participating Head Start sites. Time 2 and Time 3 response rates were calculated by dividing the total number of completions by the number of completions from the previous year after subtracting the people who indicated they did not want to be called again. The overall retention rate is calculated by the total number of completions at Time 3 for the specific cohort divided by the original sample size minus any "do-not-call" participants.

Table 2

Outcome Variables^a

Variable	Items	Response Options and Range
Neighborhood constructs		
Neighborhood involvement $\alpha = .82$	Participated in (past 12 months): neighborhood watch, block/tenant association, community council/board, community action group, neighborhood meeting	Response options: never (1), 1–2 times, 3–5 times, 6–9 times, 10 or more times (5); possible range: 5–25
School involvement $\alpha = .79$	Participated in (past 12 months): local PTA/PTO, school planning team, classroom volunteer	Response options: never (1), 1–2 times, 3–5 times, 6–9 times, 10 or more times (5); possible range: 3–15
Social connectedness $\alpha = .65$	# of friends in neighborhood (0, 1–2, 3–5, 6–9, 10+). Able to find help with a favor (never (1) to all of the time (5)). People help neighbors, close-knit neighborhood, neighbors do not get along, ^b neighbors can be trusted, neighbors do not share same values. ^b	Response options: strongly disagree (1), disagree, neither, agree, strongly agree (5); possible range: 7–35
Neighborhood organization $\alpha = .74$	A lot of crime in neighborhood ^b , neighbors take good care of houses/apartments, a lot of abandoned buildings in neighborhood ^b , too much drug use in neighborhood ^b , neighbors are afraid to go out at night ^b	Response options: strongly disagree (1), disagree, neither, agree, strongly agree (5); possible range: 5–25
Neighborhood availability of alcohol and drugs $\alpha = .83$	How easy for neighbors: Over age of 21 to buy alcohol; under age of 21 to get alcohol from friends/parents/neighbors; to buy marijuana; teenagers to buy illegal drugs	Response options: very easy (1), easy, difficult, very difficult (4); possible range: 4–16
Neighborhood norms against substance use $\alpha = .84$	How many adults in the neighborhood approved of: teenagers drinking alcohol, people using marijuana, people using illegal drugs (cocaine, LSD, PCP, heroin, speed, "meth," or ecstasy)	Response options: all (1), most, some, very few, none (5); possible range: 3–15
Satisfaction with police enforcement $\alpha = .82$	Police are doing a good job in the neighborhood, police enforce underage drinking laws, police enforce public drunkenness laws, police enforce laws against illegal drug sales in neighborhood	Response options: strongly disagree (1), disagree, neither, agree, strongly agree (5); possible range: 4–20
Family and substance abuse constructs ^c		
Family conflict	Fight a lot in your family ^b , family members hit each other ^b , family members hardly ever lose their tempers, family members throw things ^b , family members rarely criticize	Response options: strongly disagree (0), disagree (0), neither (0), agree (1), strongly agree (1); possible range: 0–5; 0 = strongly disagree, disagree, neither to all 5 items, 5 = strongly agree, agree to all 5 items
Parenting frustration	Finds it hard to understand child's concern, gives in to child or let child have his or her way, not worth getting child to obey	Response options: very often (0), often (0), occasionally (0), seldom (1), never (1); possible range: 0–3; 0 = very often, often, or occasionally to all 3 items, 3 = seldom or never to all 3 items
Physical discipline—spansks child	Spansks child	Response options: very often (1), often (1), occasionally (1), seldom (1), or never (0); possible range: 0–1
Physical discipline—hits child	Hits child with hard object (belt, hairbrush, stick, etc.)	Response options: very often (1), often (1), occasionally (1), seldom (1), or never (0); possible range: 0–1
Use of rules	Tells child reason for making the rules, uses rules to manage child's behavior, makes sure rules are followed	Response options: never (0), seldom (0), occasionally (0), often (1), very often (1); possible range: 0–3; 0 = never, seldom, or occasionally to all 3 items, 3 = very often or often to all 3 items

Variable	Items	Response Options and Range
Time-out	Uses time-out/sends child to his or her room	Response options: never(0), seldom(0), occasionally(1), often(1), very often(1); possible range: 0 to 1
PC alcohol use $\alpha = .77$	Number of drinking days in past 30 days, number of drinks on a drinking day, number of days having 5 or more drinks in a row	Response options: past 30 days use: 0, 1–2, 3–5, 6–9, 10–19, 20–29, all 30. # of drinks: (0, 1, 2, 3, 4, 5 or more), 5+ drinks: 0, 1, 2, 3–5, 6–9, 10–19, 20+; possible range: 0–17
PC and household substance use	PC (past 30 days): 5+ drinks in a row (1 or more days), should cut down drinking (yes), people annoyed by drinking (yes), felt bad/guilty about drinking (yes), drank first thing in morning (yes), marijuana use (1 or more times), any other illegal drugs (1 or more times), household: high or drunk, marijuana use, any other illegal drugs, illegal use of pain medicines	Response options: PC (past 30 days): as indicated by item, household use (all items): never (0), seldom, occasionally, often, very often (1); possible range: 0 to 1, 0 = no to all items, 1 = yes to any item
Family norms: drinking	How acceptable is it for household members to drink	Response options: very often (1), often, occasionally, seldom, or never (5); possible range: 1–5
Family norms: getting drunk	How acceptable is it for household members to drink enough to feel drunk or high	Response options: very often (1), often, occasionally, seldom, or never (5); possible range: 1–5

^a All variables are coded so that higher scores indicate better outcomes. The only exceptions are for the PC and Household Substance Abuse index (where a score of 1 indicates the presence of some substance use by the PC or someone in the household and a score of 0 indicates no substance use by the PC or anyone in the household), the PC Alcohol Use scale (where higher scores indicate more alcohol use by the PC), and the Physical Discipline: Spanking items (where a score of 1 indicates ever using hitting or spanking as a means of discipline).

^b These items have been reverse coded to be consistent with other items in the scale so that higher scores indicate better outcomes.

^c Due to the social desirability associated with answering personal questions about one's own parenting practices and family behaviors, the individual items for the Family Conflict and Parenting Frustration scales were dichotomized and then summed for the scale, as opposed to summing the multicategory responses to each item. For the same reason, the single-item constructs, physical discipline: spanking, hitting, and time-out, were dichotomized.

Table 3

Demographic Characteristics at Baseline^a

	Cohort 1			Cohort 2			Community		
	FTG % or Mean (std) ^a	Comp % or Mean (std) ^a	p Value	FTG % or Mean (std) ^a	Comp % or Mean (std) ^a	p Value	FTG % or Mean (std) ^a	Comp % or Mean (std) ^a	p Value
Race/ethnicity	n = 826	n = 1025	.0010	n = 410	n = 513	.0039	n = 525	n = 517	.9318
White	65.3	61		74.2	57.9		76.2	79.1	
Black	23.2	28.3		19.5	31.2		13.0	12.4	
Other	11.5	10.7		6.3	10.9		10.8	8.5	
Hispanic	n = 841	n = 1034	<.0001	n = 412	n = 517	<.0001	n = 524	n = 518	.8394
45.8		31.7		58.3	29.6		24.1	27.4	
Family structure	n = 863	n = 1052	.3694	n = 423	n = 530	.1976	n = 557	n = 553	.6161
Single Parent	26.2	29.1		22.7	31.5		15.3	17.7	
Relationship to child	n = 866	n = 1053	.9853	n = 423	n = 530	.6597	n = 556	n = 553	.5616
Mother	88.3	87.6		92.0	90.8		79.0	81.0	
Gender of child	n = 866	n = 1053	.1730	n = 423	n = 532	.1195	n = 557	n = 553	.2502
Female	49.1	45.9		47.5	49.8		38.06	43.04	
Infant in household	n = 863	n = 1052	.2484	n = 423	n = 530	.5085	n = 557	n = 553	.6113
Yes	26.9	25.3		29.8	27.9		18.3	19.7	
Age of primary caregiver	n = 842	n = 1035	.0882	n = 416	n = 518	.7791	n = 557	n = 553	.8251
29.7(7.3)		30.5(7.9)		30.0(7.2)	30.4(7.6)		34.3(8.6)	34.4(8.9)	
# of Children in household	n = 863	n = 1052	.8204	n = 423	n = 530	.7652	n = 557	n = 553	.6991
1.8(1.6)		1.8(1.6)		1.6(1.6)	1.6(1.5)		1.5(1.6)	1.5(1.6)	
Risk index score	n = 866	n = 1053	<.0001	n = 423	n = 532	<.0001	n = 557	n = 553	.8836
20.7(4.8)		19.0(4.3)		21.3(4.7)	19.0(4.5)		17.0(4.6)	17.2(4.7)	

^a Means and frequencies reported are unadjusted. *p* values are adjusted for the nesting structure.

Table 4

Main Effects Analysis: Cohorts 1 and 2 and Community Sample

Neighborhood constructs	Cohort 1				Cohort 2				Community Sample			
	Adjusted Means ^a		p Value for Tx × Time ^b		Adjusted Means ^a		p Value for Tx × Time ^b		Adjusted Means ^a		p Value for Tx × Time ^b	
	T1	T2	T3		T1	T2	T3		T1	T2		
Neighborhood involvement	FTG	6.21	6.07	6.24	ns	5.61	5.93	6.12	ns	5.72	5.87	ns
	Comp	5.61	5.75	6.16		5.50	5.86	6.25		5.76	5.62	
School involvement	FTG	4.41	4.68	4.83	ns	3.80	4.11	4.51	ns	4.99	4.73	ns
	Comp	4.01	4.46	4.74		3.73	4.39	4.86		5.31	4.85	
Social connectedness	FTG	23.47	23.90	24.21	ns	23.15	23.43	23.73	ns	24.27	23.59	ns
	Comp	24.26	24.74	24.98		24.04	24.66	24.98		25.78	25.27	
Neighborhood organization	FTG	18.16	18.32	18.88	ns	17.88	18.89	18.99	ns	18.31	18.01	ns
	Comp	19.10	19.38	19.65		19.04	19.69	19.79		19.53	19.04	
Neighborhood availability of alcohol and drugs	FTG	10.49	10.86	10.69	ns	10.62	10.50	10.88	ns	10.19	10.29	ns
	Comp	10.60	10.88	10.97		10.67	10.92	10.92		10.62	10.56	
Satisfaction with police enforcement	FTG	14.72	14.87	15.15	ns	14.40	14.75	15.04	ns	12.95	12.91	ns
	Comp	14.96	15.14	15.39		14.88	15.23	15.37		13.47	13.20	
Neighborhood norms against substance use	FTG	13.15	13.49	13.35	ns	13.32	13.31	13.62	ns	14.32	14.16	ns
	Comp	13.47	13.69	13.76		13.69	13.66	13.85		14.91	14.68	
Family and substance abuse constructs												
Use of rules	FTG	2.51	2.42	2.38	ns	2.27	2.34	2.24	ns	2.52	2.44	ns
	Comp	2.54	2.39	2.37		2.40	2.40	2.28		2.56	2.43	
Family conflict	FTG	3.41	3.58	3.69	ns	3.54	3.61	3.77	ns	3.43	3.22	ns
	Comp	3.48	3.61	3.58		3.53	3.61	3.83		3.48	3.39	
Parent frustration	FTG	1.91	2.13	2.2	ns	2.02	2.04	2.27	ns	1.90	1.98	ns
	Comp	1.77	2.02	2.13		1.94	2.14	2.14		1.99	2.09	
Physical discipline—spanking ^{c,d}	FTG	0.62	0.53	0.55	ns	0.57	0.59	0.56	ns	0.56	0.49	ns
	Comp	0.68	0.62	0.59		0.63	0.62	0.55		0.62	0.60	
Physical discipline—hitting ^{c,d}	FTG	0.10	0.12	0.11	ns	0.07	0.10	0.11	ns	0.09	0.06	ns
	Comp	0.13	0.11	0.12		0.10	0.11	0.12		0.09	0.08	
Time-out ^c	FTG	0.59	0.53	0.52	ns	0.42	0.42	0.46	ns	0.55	0.55	ns

Neighborhood constructs	Cohort 1						Cohort 2						Community Sample			
	Adjusted Means ^a			p Value for Tx × Time ^b			Adjusted Means ^a			p Value for Tx × Time ^b			Adjusted Means ^a		p Value for Tx × Time ^b	
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T1	T2
PC alcohol use ^d	Comp	0.57	0.49	0.48			0.58	0.57	0.44				0.57	0.52		
	FTG	1.07	1.04	1.21	ns		1.23	1.14	1.14	ns			1.46	1.70	ns	
	Comp	0.77	0.85	0.86			0.69	0.68	0.75				1.41	1.66		
PC and household substance use ^{c,d}	FTG	0.13	0.13	0.19	ns		0.14	0.17	0.14	ns			0.16	0.20	ns	
	Comp	0.08	0.09	0.09			0.09	0.08	0.09				0.13	0.18		
Family norms: drinking	FTG	2.93	3.02	2.92	ns		2.93	2.94	3.08	ns			2.55	2.64	ns	
	Comp	3.09	3.09	3.09			3.11	3.24	3.15				2.75	2.73		
Family norms: getting drunk	FTG	3.63	3.68	3.59	ns		3.57	3.59	3.66	ns			3.49	3.50	ns	
	Comp	3.72	3.72	3.69			3.77	3.79	3.77				3.59	3.58		
<i>n</i>		1,919	1,234	1,041			955	667	513				1,110	1,093		

^a Estimates for means adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Except where otherwise noted, higher scores indicate better outcomes.

^b Tests of interaction term adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Based on a Bonferroni adjustment, *p* values < .001 are treated as statistically significant and reported in the tables.

^c Dichotomous variable estimates reported as prevalences instead of means.

^d Lower scores indicate better outcomes.

Table 5
Analysis of the High Risk Sample (Based on Risk Index Score > 20): Cohorts 1 and 2 and Community Sample

Neighborhood constructs	Cohort 1					Cohort 2					Community Sample		
	Adjusted Means ^a			p Value for Tx × Time ^b		Adjusted Means ^a			p Value for Tx × Time ^b		Adjusted Means ^a		p Value for Tx × Time ^b
	T1	T2	T3			T1	T2	T3			T1	T2	
Neighborhood involvement	FTG	6.16	6.15	6.33	ns	5.36	5.64	5.94	ns		5.46	5.19	ns
	Comp	5.59	5.53	5.78		5.50	5.56	6.18			5.40	5.44	
School involvement	FTG	4.52	4.77	4.65	ns	3.60	4.19	4.39	ns		4.35	3.74	ns
	Comp	3.83	4.22	4.30		3.39	4.14	3.95			3.75	4.08	
Social connectedness	FTG	22.12	23.09	23.15	ns	22.49	23.09	23.57	ns		21.21	20.56	ns
	Comp	22.91	23.96	24.25		23.08	23.67	24.41			22.73	22.07	
Neighborhood organization	FTG	17.78	18.05	18.86	ns	17.35	18.51	18.46	ns		16.96	16.79	ns
	Comp	18.78	19.15	19.40		18.74	19.42	19.45			18.14	16.84	
Neighborhood availability of alcohol and drugs	FTG	10.54	10.84	10.79	ns	10.67	10.51	10.77	ns		9.95	10.26	ns
	Comp	10.78	11.19	11.38		10.77	10.83	11.05			10.71	10.52	
Satisfaction with police enforcement	FTG	14.54	14.91	15.15	ns	14.32	14.88	15.29	ns		14.00	14.43	ns
	Comp	15.01	15.42	15.64		14.96	15.20	15.80			14.71	13.90	
Neighborhood norms against substance use	FTG	12.84	13.37	13.17	ns	13.61	13.40	13.69	ns		12.61	12.93	.0001
	Comp	13.53	13.87	14.15		13.55	13.53	13.72			13.65	13.11	
Family and substance abuse constructs													
Use of rules	FTG	2.40	2.30	2.22	ns	2.17	2.23	2.08	ns		2.19	1.94	ns
	Comp	2.37	2.31	2.25		2.19	2.31	2.06			2.18	2.06	
Family conflict	FTG	3.37	3.49	3.72	ns	3.55	3.50	3.86	ns		3.21	2.81	ns
	Comp	3.42	3.55	3.58		3.45	3.60	3.75			3.41	3.17	
Parent frustration	FTG	1.74	2.02	2.06	ns	21.99	21.65	22.30	ns		1.68	1.85	ns
	Comp	1.58	1.86	2.01		21.24	21.87	21.97			1.85	1.85	
Physical discipline—spanking ^{c,d}	FTG	0.68	0.55	0.61	ns	0.62	0.62	0.56	ns		0.60	0.45	ns
	Comp	0.71	0.58	0.54		0.57	0.52	0.43			0.62	0.59	
Physical discipline—hitting ^{c,d}	FTG	0.14	0.13	0.12	ns	0.08	0.08	0.09	ns		0.12	0.09	ns
	Comp	0.17	0.13	0.14		0.09	0.09	0.13			0.08	0.17	
Time-out ^c	FTG	0.57	0.49	0.54	ns	0.35	0.44	0.41	ns		0.51	0.44	ns

Neighborhood constructs	Cohort 1						Cohort 2						Community Sample				
	Adjusted Means ^a			p Value for Tx × Time ^b			Adjusted Means ^a			p Value for Tx × Time ^b			Adjusted Means ^a		p Value for Tx × Time ^b		
	T1	T2	T3				T1	T2	T3				T1	T2	T1	T2	
PC alcohol use ^d	Comp	0.49	0.49	0.41			0.51	0.51	0.38				0.45	0.40			
	FTG	0.87	0.61	0.86		ns	1.12	0.91	0.65		ns		1.02	1.06			ns
	Comp	0.76	0.64	0.85			0.27	0.29	0.14				0.57	0.88			
PC and household substance use ^{c,d}	FTG	0.15	0.15	0.20		ns	0.15	0.20	0.14		ns		0.16	0.14			ns
	Comp	0.08	0.06	0.06			0.08	0.08	0.07				0.15	0.22			
	FTG	3.18	3.23	3.06		ns	3.07	3.01	3.33		ns		3.10	3.21			ns
Family norms: drinking	Comp	3.11	3.24	3.09			3.34	3.45	3.52				3.43	3.31			
	FTG	3.65	3.71	3.58		ns	3.61	3.59	3.75		ns		3.57	3.69			ns
	Comp	3.71	3.77	3.69			3.82	3.86	3.91				3.73	3.72			
<i>n</i>		878	556	467			476	330	254			278	282				

^a Estimates for means adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Except where otherwise noted, higher scores indicate better outcomes.

^b Tests of interaction term adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Based on a Bonferroni adjustment, *p* values < .001 are treated as statistically significant and reported in the tables.

^c Dichotomous variable estimates reported as prevalences instead of means.

^d Lower scores indicate better outcomes.

Table 6
Analysis of the Low Risk Sample (Based on Risk Index Score < 16): Cohorts 1 and 2 and Community Sample

Neighborhood constructs	Cohort 1					Cohort 2					Community Sample		
	Adjusted Means ^a			p Value for Tx × Time ^b		Adjusted Means ^a			p Value for Tx × Time ^b		Adjusted Means ^a		p Value for Tx × Time ^b
	T1	T2	T3			T1	T2	T3			T1	T2	
Neighborhood involvement	FTG	6.35	6.19	6.33	ns	6.28	6.74	6.91	ns	ns	5.81	6.11	ns
	Comp	5.76	5.90	6.66		5.63	5.92	6.19			6.10	5.82	
School involvement	FTG	4.39	4.91	5.53	ns	3.82	3.67	4.22	ns	ns	5.23	5.19	ns
	Comp	4.18	4.65	4.94		3.85	4.13	4.69			6.06	5.31	
Social connectedness	FTG	25.45	25.22	25.55	ns	24.00	24.29	24.58	ns	ns	25.61	25.02	ns
	Comp	26.24	26.15	26.65		25.89	26.26	26.57			27.60	27.15	
Neighborhood organization	FTG	19.15	19.19	18.96	ns	18.65	19.81	20.22	ns	ns	18.86	18.45	ns
	Comp	19.69	19.65	19.75		19.43	19.43	20.33			20.47	20.71	
Neighborhood availability of alcohol & drugs	FTG	10.38	10.84	10.61	ns	10.21	10.86	11.22	ns	ns	10.18	10.26	ns
	Comp	10.50	10.67	10.76		10.69	10.83	10.66			10.77	10.75	
Satisfaction with police enforcement	FTG	14.76	14.58	14.58	ns	14.39	14.33	14.53	ns	ns	14.30	14.12	ns
	Comp	15.11	15.13	15.49		15.02	14.73	15.06			15.03	15.06	
Neighborhood norm against substance use	FTG	13.52	13.47	13.36	ns	12.87	13.58	13.92	ns	ns	12.94	12.84	ns
	Comp	13.59	13.79	13.86		13.63	13.54	13.76			13.54	13.43	
Family and substance abuse constructs													
Use of rules	FTG	2.73	2.57	2.62	ns	2.45	2.45	2.43	ns	ns	2.65	2.70	ns
	Comp	2.73	2.60	2.61		2.47	2.53	2.40			2.70	2.60	
Family conflict	FTG	3.59	3.74	3.68	ns	3.40	3.63	3.44	ns	ns	3.47	3.42	ns
	Comp	3.56	3.76	3.67		3.83	3.77	4.11			3.54	3.61	
Parent frustration	FTG	2.13	2.26	2.36	ns	2.05	2.12	2.36	ns	ns	2.05	2.06	ns
	Comp	2.01	2.26	2.35		2.10	2.32	2.25			2.11	2.28	
Physical discipline—spanking ^{c,d}	FTG	0.52	0.51	0.45	ns	0.49	0.55	0.64	ns	ns	0.53	0.50	ns
	Comp	0.69	0.67	0.61		0.65	0.69	0.57			0.63	0.58	
Physical discipline—hitting ^{c,d}	FTG	0.07	0.09	0.10	ns	0.03	0.07	0.10	ns	ns	0.08	0.05	ns
	Comp	0.11	0.10	0.10		0.11	0.20	0.09			0.07	0.08	
Time-out ^c	FTG	0.59	0.62	0.47	ns	0.57	0.46	0.55	ns	ns	0.58	0.62	ns

Neighborhood constructs	Cohort 1					Cohort 2					Community Sample		
	Adjusted Means ^a		p Value for Tx × Time ^b			Adjusted Means ^a		p Value for Tx × Time ^b			Adjusted Means ^a		p Value for Tx × Time ^b
	T1	T2	T3			T1	T2	T3			T1	T2	
PC alcohol use ^d	Comp	0.65	0.43	0.54		0.56	0.59	0.47			0.62	0.55	
	FTG	1.13	1.29	1.60	ns	1.79	1.94	2.16	ns		1.77	2.00	ns
PC and household substance use ^{c,d}	Comp	1.04	1.45	1.22		0.84	0.97	1.30			1.77	2.10	
	FTG	0.10	0.05	0.20	ns	0.19	0.18	0.21	ns		0.16	0.22	ns
Family norms: drinking	Comp	0.08	0.12	0.09		0.04	0.04	0.07			0.11	0.15	
	FTG	2.75	2.80	2.78	ns	2.78	2.91	2.68	ns		2.30	2.39	ns
Family norms: getting drunk	Comp	2.89	2.80	2.94		2.82	2.89	2.60			2.38	2.39	
	FTG	3.62	3.58	3.53	ns	3.55	3.55	3.45	ns		3.50	3.47	ns
<i>n</i>	Comp	3.75	3.69	3.66		3.74	3.67	3.67			3.48	3.50	
		449	296	252		197	141	109			531	545	

^a Estimates for means adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Except where otherwise noted, higher scores indicate better outcomes.

^b Tests of interaction term adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Based on a Bonferroni adjustment, *p* values < .001 are treated as statistically significant and reported in the tables.

^c Dichotomous variable estimates reported as prevalences instead of means.

^d Lower scores indicate better outcomes.

Table 7
Site-Level Dose Analysis: High vs. Low Dose (FTG) vs. Comparison Sites—Cohorts 1 and 2 and Community Sample

Neighborhood Constructs		Cohort 1 ^a					Cohort 2 ^a					Community Sample ^d				
		Adjusted Means ^b			p Value for Tx × Time ^c		Adjusted Means ^b			p Value for Tx × Time ^c		Adjusted Means ^b			p Value for Tx × Time ^c	
		T1	T2	T3	T1	T3	T1	T2	T3	T1	T3	T1	T2	T3	T1	T3
Neighborhood Involvement	FTG High	6.20	6.13	5.82	ns		5.65	5.76	5.86	ns		6.04	5.59		ns	
	FTG Low	6.20	6.03	6.41			5.58	5.98	6.23			5.59	5.92			
	Comparison	5.62	5.75	6.16			5.51	5.87	6.26			5.77	5.64			
School involvement	FTG High	4.47	4.45	4.63	ns		3.58	3.71	4.11	ns		4.51	4.78		ns	
	FTG Low	4.38	4.75	4.90			3.84	4.23	4.64			4.92	4.60			
	Comparison	4.02	4.46	4.75			3.76	4.42	4.89			5.21	4.87			
Social connectedness	FTG High	22.88	23.26	23.80	ns		23.35	23.86	23.69	ns		23.91	22.94		ns	
	FTG Low	23.64	24.08	24.31			23.10	23.28	23.79			24.33	23.64			
	Comparison	24.29	24.77	25.01			24.02	24.64	24.96			25.74	25.22			
Neighborhood organization	FTG High	17.51	17.46	18.57	ns		17.43	18.79	18.21	ns		16.79	17.32		<.0001	
	FTG Low	18.34	18.59	18.94			18.00	18.88	19.30			18.54	17.95			
	Comparison	19.13	19.41	19.68			19.07	19.72	19.84			19.59	19.00			
Neighborhood availability of alcohol and drugs	FTG High	10.69	10.89	10.76	ns		10.94	10.55	10.83	ns		9.85	9.91		ns	
	FTG Low	10.42	10.86	10.68			10.52	10.50	10.93			10.16	10.28			
	Comparison	10.59	10.87	10.96			10.66	10.91	10.91			10.64	10.54			
Satisfaction with police enforcement	FTG High	14.42	14.79	14.82	ns		14.12	14.93	14.68	ns		14.86	14.30		ns	
	FTG Low	14.81	14.88	15.26			14.49	14.66	15.18			14.17	14.20			
	Comparison	14.98	15.15	15.40			14.89	15.24	15.38			14.88	14.66			
Neighborhood norm against substance use	FTG High	13.04	13.46	13.25	ns		13.63	13.75	13.61	ns		12.50	12.79		<.0001	
	FTG Low	13.19	13.49	13.39			13.23	13.17	13.68			12.99	12.84			
	Comparison	13.47	13.69	13.77			13.67	13.63	13.83			13.47	13.20			
Family and substance abuse constructs	FTG High	2.51	2.46	2.48	ns		2.25	2.33	2.23	ns		2.59	2.44		ns	
	FTG Low	2.52	2.41	2.34			2.27	2.35	2.24			2.50	2.44			
	Comparison	2.53	2.39	2.36			2.40	2.40	2.28			2.55	2.43			
Family conflict	FTG High	3.34	3.65	3.74	ns		3.73	3.81	3.99	ns		3.26	3.48		ns	

Neighborhood Constructs	Cohort 1 ^a				Cohort 2 ^a				Community Sample ^a			
	Adjusted Means ^b		p Value for Tx × Time ^c		Adjusted Means ^b		p Value for Tx × Time ^c		Adjusted Means ^b		p Value for Tx × Time ^c	
	T1	T2	T3		T1	T2	T3		T1	T2		
Parent frustration	FTG Low	3.44	3.55	3.67	3.50	3.55	3.71		3.46	3.10		
	Comparison	3.48	3.61	3.58	3.51	3.60	3.82		3.49	3.38		
	FTG High	1.96	2.22	2.32	1.96	1.95	2.16	ns	2.00	2.17	ns	
	FTG Low	1.90	2.10	2.15	2.03	2.07	2.30		1.89	1.92		
Physical discipline—spanking ^{d,e}	Comparison	1.76	2.02	2.12	1.95	2.14	2.15		1.97	2.05		
	FTG High	0.70	0.62	0.72	0.59	0.62	0.51	ns	0.60	0.47	.0001	
	FTG Low	0.60	0.50	0.48	0.57	0.58	0.59		0.56	0.50		
	Comparison	0.67	0.61	0.58	0.63	0.62	0.55		0.63	0.60		
Physical discipline—hitting ^{d,e}	FTG High	0.16	0.14	0.18	0.06	0.04	0.13	ns	0.17	0.03	ns	
	FTG Low	0.09	0.12	0.09	0.07	0.12	0.09		0.07	0.06		
	Comparison	0.13	0.11	0.12	0.10	0.11	0.12		0.09	0.08		
	FTG High	0.60	0.52	0.56	0.31	0.41	0.43	ns	0.56	0.51	ns	
Time-out ^d	FTG Low	0.58	0.54	0.50	0.46	0.42	0.47		0.56	0.56		
	Comparison	0.57	0.49	0.48	0.59	0.58	0.45		0.57	0.53		
	FTG High	0.73	0.91	1.05	1.40	1.60	1.46	ns	1.15	1.97	ns	
	FTG Low	1.17	1.06	1.25	1.20	1.00	1.03		1.58	1.64		
PC alcohol use ^e	Comparison	0.79	0.86	0.87	0.67	0.66	0.72		1.42	1.67		
	FTG High	0.11	0.12	0.15	0.13	0.28	0.18	ns	0.14	0.18	ns	
	FTG Low	0.14	0.13	0.20	0.14	0.15	0.13		0.17	0.21		
	Comparison	0.09	0.09	0.09	0.09	0.08	0.09		0.13	0.18		
PC and household substance use ^{d,e}	FTG High	3.29	3.12	3.10	3.02	3.02	3.09	ns	2.64	2.54	ns	
	FTG Low	2.82	3.00	2.88	2.91	2.92	3.08		2.52	2.70		
	Comparison	3.07	3.07	3.08	3.11	3.23	3.15		2.76	2.73		
	FTG High	3.78	3.75	3.65	3.69	3.63	3.59	ns	3.55	3.44	ns	
Family norms: getting drunk	FTG Low	3.59	3.66	3.58	3.54	3.58	3.70		3.47	3.51		
	Comparison	3.71	3.71	3.68	3.77	3.79	3.77		3.58	3.59		
		1,919	1,234	1,041	955	667	513		1,110	1,093		
	<i>n</i>											

^aThree high-dose sites with strict rule for “high” dose.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

^b Estimates for means adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Except where otherwise noted, higher scores indicate better outcomes.

^c Tests of interaction term adjust for primary caregiver: age, race, Hispanic ethnicity, risk index score, relationship to HS child, number of children at home, presence of infant at home, family structure, and child gender; site-level variables: community action agency, family service worker education, enrollment size, early HS, population density of city, SES factor score, percentage owner-occupied housing. Based on a Bonferroni adjustment, p values $< .001$ are treated as statistically significant and reported in the tables.

^d Dichotomous variable estimates reported as prevalences instead of means.

^e Lower scores indicate better outcomes.