RESEARCH ARTICLE

YOUTH AT RISK FOR GANG INVOLVEMENT

Reducing Crime Among Youth at Risk for Gang Involvement

A Randomized Trial

Terence P. Thornberry
University of Maryland—College Park
University at Albany, State University of New York

Brook Kearley
University of Maryland—Baltimore

Denise C. Gottfredson
Molly P. Slothower
University of Maryland—College Park

Deanna N. Devlin
CUNY Institute for State and Local Governance

Jamie J. Fader
Temple University

Research Summary

No gang prevention or intervention programs meet the standards for effectiveness promulgated by Blueprints for Healthy Youth Development. This randomized controlled trial of a well-known program—Functional Family Therapy—that was modified to address the needs of gang-involved adolescents yields two main findings. First, youth at

high risk for gang membership and their families engaged with and successfully completed the program at the same level as low-gang-risk youth. Second, the effectiveness results varied by gang-risk status. For youth at high risk for gang membership, the treatment group had significantly lower recidivism rates at the 18-month follow-up as compared with a "treatment as usual" control group. For youth at low risk for gang membership, however, no consistent differences were found between the treated and control groups.

Policy Implications

Modifying and extending evidence-based delinquency programs to gang-involved youth seems to be a reasonable strategy for developing a wider array of effective programs to respond to the challenge of street gangs. The differential findings by gang-risk status suggests that the juvenile justice system should expand the use of evidence-based community programs to higher risk youth, including those identified as being "at risk" because of their gang involvement.

Keywords

gangs, gang prevention intervention, family therapy for court-involved youth, randomized controlled trial, Functional Family Therapy

During the last 25 years, adolescent street gangs, once primarily a phenomenon of a few major metropolitan areas, have spread rapidly throughout the country. In its most recent survey, the National Youth Gang Center estimated that 30,700 gangs with 850,000 members were located in more than 3,100 jurisdictions throughout the United States. Gangs were found in 85% of larger cities, 50% of suburban counties, 32% of smaller cities, and 15% of rural counties. The number of gangs, gang members, and gang-related homicides are on the rise when compared with the previous 5-year average, and the problem of street gangs now reaches into all corners of American society (Egley, Howell, and Harris, 2014).

Given the prevalence of gangs and their negative impact on the behavior and development of gang members (Pyrooz, Turanivic, Decker, and Wu, 2015; Thornberry, Krohn, Lizotte, Smith, and Tobin, 2003), it is imperative to develop effective, evidence-based programs to prevent gang membership and to reduce the impact of gangs on the adolescents who do join them. In this area where effective programs are strongly needed, however, they are least available. Although evidence-based programs for a variety of other problem behaviors do exist, currently no known gang programs meet the rigorous standards of demonstrated effectiveness such as those promulgated by the Blueprints for Healthy Youth Development (blueprintsprograms.com). The purposes of this article are to (a) describe a strategy for developing an evidence-based gang program, (b) examine whether such a program can successfully engage gang members and their families, and (c) present the results of a randomized controlled trial (RCT) that evaluated its effectiveness. We begin

with a brief description of the impact of adolescent street gangs and the need for effective programming.

Impact of Adolescent Street Gangs

Since the pioneering work of Frederick Thrasher (1927), compelling scientific evidence has been presented that gang membership is associated with elevated involvement in delinquent and criminal behavior. The results of a recent meta-analysis of 179 empirical studies clearly demonstrate the association between gangs and crime: "[O]ur analyses reveal that gang membership is consistently and significantly related to offending regardless of the measurement of key variables, sampling approaches, and model specifications" (Pyrooz et al., 2015: 381–382). Gang members are involved at a higher level than are nongang members in almost all forms of criminal behavior including violent crime, property crime, drug use, drug sales, and gun crime (Thornberry et al., 2003). This higher rate of involvement is observed across racial and ethnic groups and across gender (Esbensen, Peterson, Taylor, and Freng, 2010; Pyrooz et al., 2015). The discrepancy between gang and nongang members is also evident at the most extreme end of the continuum. The National Youth Gang Center estimated that even though gang members represent only a fraction of 1% of the U.S. population, a total of 2,363 gang-related homicides occurred in 2012, representing approximately 16% of all homicides in the country (Egley et al., 2014). The impact of gang membership has also been linked to the recent upswing in homicide and violent crime in cities such as Philadelphia, the site of this study, and Chicago (Drug Enforcement Agency [DEA], 2015; University of Chicago Crime Lab, 2017).

Compelling evidence also exists that gang membership is not merely a risk factor for offending but seems to facilitate this high level of involvement in delinquency. In numerous individual studies where rigorous methods were used to control selection effects, scholars have found that crime and delinquency are elevated during time periods when adolescents are active gang members and directly exposed to gang influences as compared with periods either before or after membership (e.g., Haviland and Nagin, 2005; Melde and Esbensen, 2011, 2013; Thornberry, Krohn, Lizotte, and Chard-Wierschem, 1993; Thornberry et al., 2003). This pattern, which has been replicated in more than 20 studies in the United States and abroad using many different measures, samples, and analytical methods (Krohn and Thornberry, 2008), strongly suggests that it is unlikely that individual characteristics of gang members account for their elevated level of delinquency. Indeed, in no study of this issue have scholars found support for a pure selection model that hypothesizes that the gang effect is spurious and that gangs are merely a collection of individuals who already have a high propensity to offend. Instead, the data suggest that it is something about the gang itself—for example, its structure, norms, culture, and group processes—that increases the delinquency of gang members. It is also important to note that gang effects are not merely delinquent peer group effects. Even when gang members are compared only with nonmembers who have highly delinquent peer groups, the gang members still have significantly higher levels

of offending than do the nonmembers (Battin, Hill, Abbott, Catalano, and Hawkins, 1998; Huizinga, 1996; Thornberry et al., 2003).

Nor is the impact of gang membership limited to offending. Membership has a host of negative consequences that disrupt the normal course of development and create disorder in the individual's life course. For example, gang membership is associated with reduced school commitment and educational attainment, becoming a teen parent, experiencing unemployment, increased commitment to negative peers, and anger identity (Krohn, Ward, Thornberry, Lizotte, and Chu, 2011; Melde and Esbensen, 2011, 2013; Pyrooz, 2014). Although adolescents are typically only gang members for 1 or 2 years, the negative impact of their exposure to the gang lasts into adulthood. During their late 20s and early 30s, gang members evidence increased economic hardship and family problems, poorer physical and mental health, substance abuse, and continued involvement in crime including elevated rates of incarceration (Gilman, Hill, and Hawkins, 2014; Krohn et al., 2011). Similarly, gang membership is significantly related to the perpetration of child maltreatment against the next generation (Augustyn, Thornberry, and Krohn, 2014).

Need for Evidence-Based Gang Programs

Given the pernicious and lasting consequences of gang membership, it should be a high priority to have effective programs to prevent adolescents from joining gangs and, for those who do join, to reduce the negative impact of the gang on behavior and development. That is, communities should have available to them gang programs with known efficaciousness and cost-effectiveness so that they can respond to the gang problem with a reasonable likelihood that their efforts—often costly—will meet with success. The best way of accomplishing this objective is to identify "evidence-based programs"—programs with strong scientific evidence that they work.

Blueprints for Healthy Youth Development (Mihalic et al., 2002) is designed to address this need. It identifies what it calls "promising" and "model" programs and provides a "blueprint" of model programs that helps communities implement them properly.[1] Of the various initiatives that identify programs that work, Blueprints uses the most rigorous scientific standards in its assessment (Fagan and Buchanan, 2016). For a program to be considered a Blueprints Promising program for antisocial behavior, it must demonstrate a significant suppressing effect on violence, delinquency, or drug use in research conducted using an experimental design with random assignment to treatment and control groups or two high-quality quasi-experimental evaluations. For Model status, the program must also be replicated in at least one independent site using a randomized control trial and the intervention effect must be sustained for at least 1 year posttreatment. Although generally considered the most rigorous standards in the area of criminology, it should be

---

1.   The Blueprints initiative was formerly called "Blueprints for Violence Prevention," the program name when this project was initiated. We use both terms, as appropriate, in this article.

noted that these standards do little more than reflect the basic requirements of good scientific study. Thus, the standards are not unreasonable and they certainly seem to be the minimum needed to respond to a problem as serious and entrenched as adolescent street gangs.

Scores of programs have been developed and implemented to respond to the "gang problem" in American society (Klein, 1995; Klein and Maxson, 2006). Unfortunately, many (perhaps most) of them have never been evaluated, and of those that have been evaluated, the evaluations do not reach the level of scientific rigor required by Blueprints standards. As a result, in several systematic reviews of the gang intervention literature (e.g., Klein and Maxson, 2006; Thornberry et al., 2003), scholars concluded that there are no known evidence-based programs for gang prevention or reform. Consistent with this finding, the Blueprints for Healthy Youth Development initiative, which reviews hundreds of prevention and treatment programs, including those that focus on gangs and gang members, does not include any gang programs on its lists of Model or Promising programs. "If carefully done experimental and quasi-experimental designs are used as our benchmark of program effectiveness, there is simply no compelling evidence that [gang programs] are effective" (Thornberry et al., 2003: 197). Thus, in an area where we need the strongest and most effective programs, we ironically have the least.[2]

Evidence-based programs for other adolescent problem behaviors do exist, however. Indeed, when this project began, Blueprints for Violence Prevention had identified 11 Model programs and 19 Promising programs that met the standards summarized earlier. These programs demonstrated a significant effect on violence, delinquency, drug use, or a combination of these behaviors (Mihalic et al., 2002) and have been shown to be cost-effective. For example, in a recent report by the Washington State Institute for Public Policy (2017), researchers found an $8.35 benefit of Functional Family Therapy (FFT) services for every $1.00 spent.

Strategies to Develop Evidence-Based Gang Programs

Given the absence of strong, evidence-based programs either for youth at risk for gang membership or for active gang members, several strategies are available for moving forward. For example, one could rigorously evaluate existing gang programs that have not yet been adequately studied to see whether they are effective. A second strategy would be to develop and evaluate gang prevention/intervention programs *de novo* based on known risk factors for gang membership and on theories regarding the underlying mechanisms by which the gang generates such negative consequences for its members.

Other approaches would build on the demonstrated success of existing evidence-based programs for delinquency such as those in the Blueprints collection. There are at least two

---

2.   There are some promising gang programs (see Esbensen, Osgood, Peterson, Taylor, and Carson, 2013; Howell, 2000) although they do not achieve Blueprints standards.

strategies in this area. First, one could provide one of these programs specifically to youth at risk for gang membership or to active gang members and evaluate the outcomes. Several Blueprints model programs have gang members among their typical client populations, but little evidence has been found about whether the program works specifically for the gang-involved participants. One exception is the work of Boxer (2011) who examined whether Multisystemic Therapy (MST), a Blueprints model program in which a community-based treatment program for serious delinquents is provided (Henggeler, Schoenwald, Borduin, Rowland, and Cunningham, 2009), was effective for the subset of clients who were gang involved. Boxer, Kubik, Ostermann, and Veysey (2015) found that gang-involved youth were more likely to fail to complete treatment, either through a lack of engagement with the program or through rearrest. This outcome was particularly strong for active gang members where only 38% of the cases closed successfully, as compared with 78% of the other cases (Boxer et al., 2015). More recently, Boxer, Docherty, Ostermann, Kubik, and Veysey (2017) extended their examination of the effectiveness of MST for gang-involved youth to the study of rearrest rates 12 months posttreatment. For youth who received MST, they found no significant differences in the prevalence or frequency of rearrests by gang status. That is, the rates were not significantly different for those involved and not involved in gang activity, which suggests that the impact of MST was approximately equal for gang-involved and nongang-involved youth (Boxer et al., 2017).

A second approach that builds on the strengths of existing evidence-based programs is to work with program developers to enhance or make accommodations to the program to address the specific needs of gang-involved youth, the specific risk factors that are associated with gang membership, and specific gang processes. This accommodation strategy retains all of the basic components and approaches of the model program but places additional emphasis on issues identified in theory and empirical study that seem particularly relevant to explaining why street gangs have such a pervasive negative impact on their members.

Given the severity of the gang problem in contemporary American society, it would be valuable to explore all of these strategies for developing effective, evidence-based gang programs. In the current project, we considered these options and decided to adopt the accommodation strategy to see whether some model programs could be tailored to the particular needs of gang-involved youth.

Development of the Current Project

This project began in 2009 with two primary objectives. The first was to modify at least one evidence-based program to respond to the particular needs of gang-involved youth. In meeting this objective, we decided to focus on programs that met the standards of the Blueprints for Healthy Youth Development initiative given the strength of those standards and the centrality of the Blueprints initiative to current efforts in prevention science (Fagan and Buchanan, 2016). The second objective was to evaluate the implementation and effectiveness of the modified program with a research design that meets Blueprints standards.

To aid in accomplishing these objectives, especially the first one, an Advisory Board with representatives from both the gang and the prevention science research communities, two research traditions with little overlap or communication, was established. Once established, the Advisory Board worked together to review all of the existing Blueprint Model and Promising programs to identify those that had the greatest likelihood of working in the gang environment. A set of standards for reviewing the content of the Blueprints programs was established. The standards included such criteria as the consistency between the theoretical orientation of the program and the risk factors for and group processes of gangs, overlap in the developmental and behavioral characteristics of the program's targeted client population and the characteristics of gang members, the strength and comprehensiveness of the intervention given the number and severity of risk factors exhibited by gang members, and the level of program modifications likely to be needed to address gang issues. The Advisory Board narrowed the set of potential programs to three Model Blueprint programs that met these review criteria: FFT, Multidimensional Treatment Foster Care, and MST (Alexander and Parsons, 1982; Chamberlain, 2003; Henggeler et al., 2009, respectively). Each program was then visited by project staff and two members of the Advisory Board to gain more detailed information about the appropriateness of the program for our purposes and to gauge interest by the program developers in collaborating with this effort. After the site visits, the Advisory Board considered all three programs to be excellent candidates for modification to address the needs of gang-involved youth. Given available resources, however, only one program could continue to the next stage of investigation.

The Advisory Board decided to begin with FFT for two primary reasons. The first was its breadth. It was viewed as the most likely of the three programs to address both the prevention of gang membership and the treatment of active gang members. Second, although never studied specifically in a gang population, FFT—like the other two programs—has undergone numerous evaluations. The earliest study of FFT was an RCT involving random assignment of 86 court-involved youth to one of three conditions. Recidivism records collected 18 months after the end of treatment indicated moderate-to-large effect sizes (ES) favoring FFT (ESs ranging from .47 to .72) in comparison with each control condition (Alexander and Parsons, 1982). Subsequently, 15 English-language studies have been published, 8 of which included random assignment to conditions. The results from all but one of these studies (Darnell and Schuler, 2015) were recently summarized in a well-conducted meta-analysis. Hartnett, Carr, Hamilton, and O'Reilly (2017) separately meta-analyzed studies using randomly and nonrandomly assigned groups, and, within each of those groups, studies comprising untreated, treatment as usual, and well-defined alternative treatment control groups. All six analyses yielded an average ES favoring the FFT group. Average ESs ranged from .08 to .90 and were statistically significant in three of the six comparisons. The authors concluded that their results "provide support for the effectiveness of FFT compared with untreated controls and well defined alternative treatments such

as cognitive behavior therapy, other models of family therapy and individual and group therapy for adolescents" (Hartnett et al., 2017: 615).

Development of Functional Family Therapy-Gangs (FFT-G)

FFT is a brief and widely disseminated evidence-based treatment for youth presenting with problem behaviors including delinquency and substance abuse (Alexander and Parsons, 1982; Barton and Alexander, 1981; Waldron and Brody, 2010; Waldron, Slesnick, Brody, Turner, and Peterson, 2001). FFT cases are referred from a variety of sources, including child welfare and mental health agencies and schools, but the largest referral source is the juvenile justice system.[3] The typical program involves 12–15 face-to-face sessions of approximately 1 hour (delivered over a 3-month period) during which trained therapists work with the targeted youth as well as with his or her caregivers, usually in a home setting. FFT has been replicated across sites and settings (Alexander, Pugh, and Parsons, 1998; Barton, Alexander, Waldron, Turner, and Warburton, 1985; Waldron and Turner, 2008).

Although FFT therapists usually encounter youth who are gang-involved or deemed to be at risk for gang involvement, FFT has no specific module or approach focusing on the influence of gangs and no study to date has specifically been aimed at investigating its effectiveness for this population. In this study, we addressed these issues by developing an accommodation of FFT for use with a gang population. To accomplish this, the Principal Investigators and three members of the Advisory Board held a series of meetings with the developers of the FFT program, led by Dr. James Alexander, to discuss how the basic FFT approach could accommodate the needs of gang-involved youth. To aid in this process, the Advisory Board developed a list of issues that should be considered in the accommodation. They included such issues as risk factors and reasons for joining gangs, gang types, gang processes, understanding (and debunking) myths about gangs, the role of violence and guns in gang activities, and patterns of retaliatory violence. Based on these discussions, Dr. Alexander and his team modified the FFT manual and training materials to address issues that are likely to be more salient in a gang population than in a general delinquent population. It is important to note that this accommodation included all of the basic components and approaches of FFT. That is, no core aspect of FFT was removed. The FFT-G manual enhanced the core FFT model by emphasizing careful preparation and knowledge development prior to family engagement with respect to gangs in general, as well as the gang environment in the particular community where the youth and family live. Given the general severity of risk factors in a gang population, FFT-G was designed to involve more direct treatment to address ongoing pressure from neighborhood gang members as well as greater preparation prior to treatment. The FFT-G accommodation

---

3.   FFT referral mechanisms vary from state to state. In the most recent 3-year period included in FFT LLC reports, the percentage of cases referred from juvenile court ranged from 34% to 54% for Pennsylvania and from 75% to 80% for Maryland.

also anticipated that additional effort might be required to engage families of gang-involved youth.

While the FFT-G accommodation, manual, and training materials were being developed, other project activities focused on developing the randomized controlled trial to evaluate FFT-G. Primary among them was the selection of a study site. Several cities were considered, and Philadelphia was selected because it met two criteria. First, the FFT national offices indicated that Philadelphia had a sufficient number of trained teams to provide the required level of service and a strong infrastructure for implementing FFT as evidenced by high levels of model adherence and fidelity. Second, data from the National Gang Center's ongoing survey of law enforcement agencies identified Philadelphia as a city with chronically high levels of gang activity. This finding was corroborated in meetings with members of the Philadelphia juvenile justice system and service providers. Based on these criteria, Philadelphia was selected as the study site. We then began to refine the research design and work out the logistics of the study with the Philadelphia Juvenile and Family Court (hereafter "Family Court"); the FFT service providers; and Community Behavioral Health (CBH), the corporation contracted by the City of Philadelphia to provide mental health and substance abuse services for Philadelphia County Medicaid recipients.

During that 2-year period, all research procedures were developed, refined, and discussed with the Chief of Probation at Family Court. A faculty member in the Department of Criminal Justice at Temple University in Philadelphia joined the project to conduct data collection. The FFT team developed the FFT-G manual and training materials, and it identified and trained therapists in FFT-G. A Family Court Judge willing to work closely with the project was identified, and study procedures were tailored to her courtroom. The research team briefed probation officers on study procedures and addressed issues they identified. Working with the Department of Probation and related agencies, appropriate services were identified for the control group and the agency providing these services was briefed on the study. Letters of agreement with all participating agencies were developed along with a formal memorandum of understanding (MOU) with Family Court. This MOU specifically granted permission to randomize youth to study conditions and to access all needed court records. All study procedures (including random assignment) were tested with a few families during the pilot phase in 2012. Issues that needed to be addressed to improve study operations were resolved, and procedure manuals were revised before beginning the study in September 2013.

Implementation Challenges

The model originally developed for this project and just described was to modify an existing evidence-based program by adding a strong gang component and then to evaluate the modified program in a randomized controlled trial with active gang members and youth clearly at risk for gang membership. In implementing the project, however, we encountered

several real-world challenges that moved us away from the original model. To provide a context for interpreting the results that follow, we discuss two major issues here.

Level of Program Modification

As discussions with the FFT clinical team about gang modifications unfolded, it became clear that the developers were reluctant to make major changes to the core FFT model to address special issues related to gangs and gang members. In their minds, FFT routinely deals with gang members in its general population, and it does so successfully. In the end, the FFT-G accommodation was viewed as minor by the FFT developers. They noted that peer-focused strategies such as refusal skills were already a major focus of FFT and that the usual FFT behavior change strategies (e.g., modeling, directing, training parent(s) and youth, and applying specific techniques such as effective parenting, communication training, behavioral contracting, and contingency management) would be used similarly with the gang population as with any FFT population. Adding specific therapeutic approaches focused on gang issues was, therefore, viewed as somewhat redundant to standard FFT practice and one that ran the risk of interfering with the success of the basic FFT model. As a result, there were modest differences between the basic FFT manual and the FFT-G manual. Gang-related examples were added to the FFT-G manual to illustrate various therapeutic strategies, but the strategies themselves were not changed from the basic FFT manual.

In practice, all FFT-G families received the full FFT model. As noted, none of the core aspects of the FFT approach were eliminated or changed. Some additional material about how to respond to gangs and gang processes was added, but the core FFT approach remained in place. Also, the therapists did not view the families and adolescents in FFT-G as more dysfunctional or more serious cases compared with many in their regular caseloads. This reinforced their view that the basic FFT approach was adequate for addressing the specific needs of gang-involved youth. Relatedly, to maximize the chances of success in this initial test of an evidence-based program for gang members, we deliberately chose a city that had service agencies with a strong record of providing FFT with high fidelity, and within those agencies, we chose therapists who had extensive experience delivering FFT. As a result, the therapists were imbued with the FFT model and generally thought that the basic FFT approach was adequate to address the needs of gang members. In fact, in postprogram interviews, the therapists indicated that they rarely used the FFT-G manual once treatment began.

This is not to say that there were no differences between the FFT and FFT-G programs. The project's focus on street gangs, the impact of gang membership on behavior and development, and the lack of evidence-based gang programming was a constant theme in all our discussions with FFT developers and therapists. Also, as noted, the FFT-G manual included additional material about gangs and gang members, and the specific training for FFT-G, delivered by Dr. Alexander, focused on gang issues, for example, in the role-playing

aspects of the training. Perhaps the biggest difference, however, was the level of detailed supervision that the therapist received for the FFT-G families. Although these therapists carried a full FFT caseload for their agency, approximately 12 to 15 families, at any point in time, only 2 or 3 of them were FFT-G families. For these FFT-G families, the therapists received separate and much more extensive supervision during weekly supervision sessions provided by the national FFT office for this project. Even though the typical FFT case is "staffed" (e.g., the focus of extensive discussion and feedback during supervision meetings) less than once during the duration of treatment given the caseloads of the typical therapist, each FFT-G case was staffed between 1 and 10 times, with an average of 4.7 times. This level of staffing provided the FFT-G therapists with more intensive and frequent feedback about each of their cases than is typical in regular FFT supervision. Additional detail regarding FFT-G implementation can be found in a companion paper (Gottfredson et al., 2018).

Nevertheless, the overlap between regular FFT and FFT-G, both in content and in implementation, is noteworthy. In the end, it seems that the program, as delivered, represents something of a hybrid between two of the strategies described earlier—a hybrid between seeing whether a standard evidence-based program works with gang members and evaluating an accommodation of a standard evidence-based program. Although there were additions to the FFT model that reflect gang issues, the FFT-G program, as delivered, seemed primarily to reflect the basic FFT approach.

Recruiting Gang-Involved Youth

A second major challenge concerned the project's ability to recruit current gang members to evaluate FFT-G as a treatment program for active gang members. In general, the prevalence of gang membership is low, especially in a cross-sectional study in which scholars identified gang membership at a particular point in time, as was done in this project. For example, Melde and Esbensen (2013) reported annual prevalence rates of gang membership in the G.R.E.A.T. study that ranged from 2.8% to 5.3%, and Boxer et al. (2017) reported that 5% of the MST sample were current gang members (see also Melde and Esbensen, 2011; Thornberry et al., 2003). Recognizing that the annual prevalence of gang membership is quite low, we selected a city with a serious level of gang activity and decided to focus on the Family Court as a source of referrals as, by definition, the court deals with high-risk youth. Even so, for a variety of reasons, it proved difficult to recruit an adequate number of gang members.

Our first strategy was to recruit from the court-wide docket by receiving referrals from probation officers. The court personnel in Philadelphia, especially the probation officers, however, were reluctant to take part in a "gang project" or to label the youth in their charge as "gang members" or even as being at risk for gang membership. They were particularly concerned about the impact of that label for future arrests and dispositions and its potential negative consequences for school and work. Indeed, some court personnel even denied the existence of gangs in Philadelphia.

As a result, we were obliged to work with one Family Court judge who understood the importance of having gang members as participants and who valued rigorous evaluation of this program. Initially the judge agreed to refer all youth from her courtroom who met our eligibility criteria. It soon became clear, however, that she intended to refer only cases that she thought were suitable for community treatment. She sent many otherwise study-eligible youth to residential placement typically because she thought keeping them in the community posed a threat to public safety or, in some cases, to the youth. That further reduced the number of gang members in the study. This restriction made it difficult to implement our original plan of recruiting only families who lived in neighborhoods with high gang activity as identified by the Philadelphia police and courts. In the end, to have a rate of referral that generated the number of cases needed within the available time frame, recruitment was expanded to all neighborhoods. That said, most of the cases did come from the originally targeted neighborhoods, as might be expected given the normal caseload in an urban family court.

Because of these issues, it was difficult to achieve the purposes of this project in full, evaluating a program accommodated to focus specifically on gang issues for a population that included a substantial number of active gang members. First, as noted earlier, the program, as delivered, represents something of a hybrid between standard FFT and FFT-G. Second, although the gang risk measures did identify a substantial number of youth at risk for gang membership, it proved difficult to identify enough active, or even former, gang members to evaluate the program as a treatment intervention.

Although the project confronted challenges of this nature, it did successfully complete a randomized controlled trial of FFT-G for a high-risk population that contains a substantial proportion of youth who were either former or current gang members or who are at high risk for joining a gang (see the Measures section). Based on that, we test the following hypotheses:

H1: Youth at high risk for gang membership and their families will engage with FFT and successfully complete FFT at the same level as nongang-involved youth.

H2: FFT-G will significantly reduce the likelihood of joining and rejoining a gang.

H3: FFT-G will significantly reduce involvement in delinquency, violent delinquency, and drug offenses for youth at high risk for gang membership.

Method

We randomly assigned adjudicated youth from a single courtroom in the Philadelphia Family Court to FFT-G and a "treatment as usual" (TAU) condition. The TAU condition involved regular probation as well as referral to an alternative family therapy program, called the Family Therapy Treatment Program (FTTP). FTTP was a program also used by the Philadelphia Family Court, and its services were eligible for reimbursement through Medicaid. It was approximately of the same intensity and duration as FFT, but it was

not manualized and had not undergone rigorous evaluation. This study represents an independent evaluation of FFT-G; the national FFT organization had no involvement in data collection or analysis.

Participants

Participants were families of study-eligible youth whose cases were heard on the participating judge's docket between September 15, 2013 and February 4, 2016 and for whom the judge ordered family services. To be eligible for inclusion, youth had to be an 11–17-year-old male[4] and could not have been referred for FFT services in the past year.

A total of 129 families participated in the study. Based on responses collected from caregivers at the baseline interviews, families were disproportionately of lower income. The median household income was $17,500, and 44% of the sample had a household income of less than $13,000. Fifty-eight percent of caregivers were currently working at the time of the pretest, and 83% reported receiving public assistance. The caregiver sample was 79% female, 80% African American, 19% Hispanic/Latino, and 25% married, with mean age 41.1 (standard deviation [SD] = 8.4). The mean age of participating boys was 15.4 (SD = 1.4).

Consistent with our goal of targeting youth at risk for gang involvement, the study sample was slightly (4–5 months) younger and had been involved in a higher percentage of crimes against persons (40.3% vs. 35.4%) than was typical of cases disposed in Philadelphia Family Court during the same period. The sample was also more likely to be non-Hispanic Black (78.4% vs. 68.3%).

Study Recruitment and Randomization into Research Conditions

Upcoming court dockets were scanned 2 weeks before each hearing date to identify study-eligible youth. For those eligible youth whom the participating judge deemed suitable for community services, she ordered "family services" as a condition of probation. Two members of the research team, who were present each day in the courtroom, met in a private space with the parents/guardians and youth after the hearing to obtain parental consent and youth assent, and then they conducted the pretest interviews with the consenting parent and child in separate offices. If the family was unable to stay at the courthouse to complete the pretest interview, the researchers scheduled an appointment to administer the pretest at the home of the family within 2 weeks. As compensation for their time, the caregiver and youth were each paid $25 to complete the interview.

---

4. Approximately 20% of the FFT clients served by the Family Court are female (Commonwealth of Pennsylvania, 2015). Given that base rate, there were not enough female clients to generate the sample needed to test our hypotheses within the project's timeframe and available budget. Indeed, even for the much larger group of male clients (80% of the total), we fell slightly below the sample size called for by our power analysis (which will be introduced shortly). Because we could not enroll enough females to allow for disaggregated analysis, we limited the study to males.

After receiving consent and conducting the pretesting, we randomly assigned families to FFT-G or the alternative program (FTTP) using a list of random numbers previously computer generated by the PI. Only the research manager and the PI had access to the random assignment list, and only the research manager consulted the list to carry out the random assignment. Randomization results were never communicated to field staff. Researchers prepared referral forms for the appropriate treatment (FFT or FTTP) to be processed by CBH, which handles all such referrals for court services and reimburses providers using Medicaid funds. CBH then assigned each family to one of the three participating FFT agencies (for treatment cases) or to the agency that provides FTTP. Researchers then informed the judge and the probation officers (POs) of the specific assignment.

Figure 1 shows the flow of cases through the different phases of the study. Of the 66 families assigned to receive FFT-G, 53 (80.3%) received at least one session. Of the

FIGURE 1

Flowchart of Participant Enrollment and Study Participation [Color figure can be viewed at wileyonlinelibrary.com]
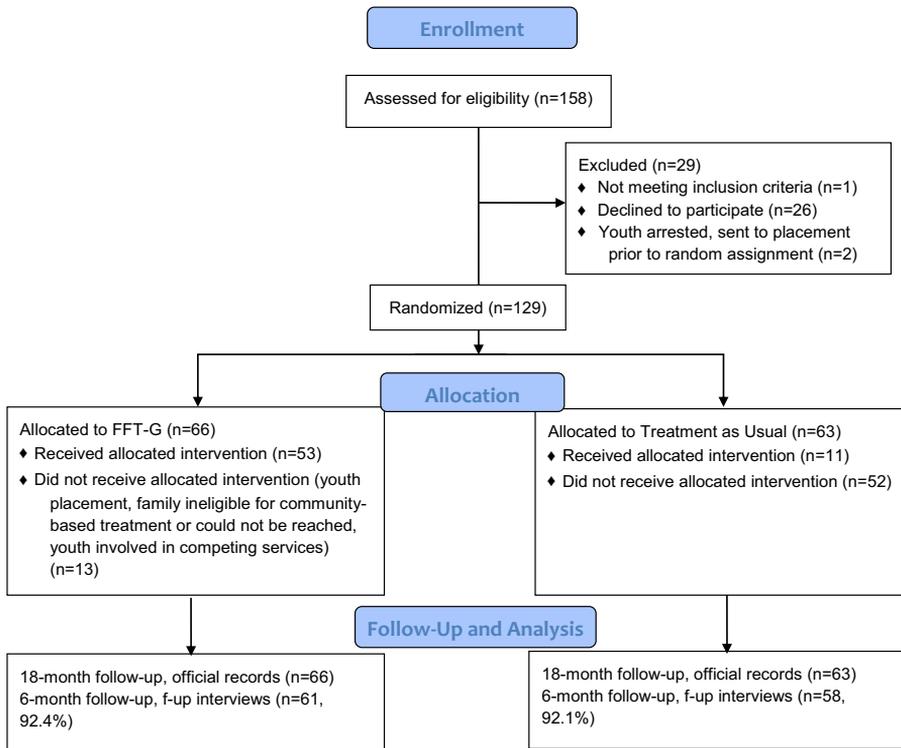
63 families assigned to the TAU condition, 11 (17.5%) enrolled in FTTP. The posttest response rate for the interview was 92% for both treatment and control youth.[5] All study participants were included in the follow-up data collection and analysis, regardless of program participation.

Data

Data come from four main sources. First, participants were interviewed at study intake and again at 6 months postrandomization.[6] Second, data on contacts with the juvenile justice system (the full history as well as subsequent contacts for the 18-month period after random assignment) were collected from Family Court records. Adult court records were also checked for the postrandom assignment period. Information on residential placements were collected from court records. Third, data on community services received during the first 6 months after random assignment as well as the costs of those services were obtained from CBH, the Family Court, and the Department of Human Services (for residential placements). Finally, data on FFT-G fidelity and adherence were obtained from a computerized tracking system into which therapists entered information about each client contact.

Measures

We discuss the measurement of two key concepts for the current analysis. The first is an indicator of gang risk. The second concerns the key outcome variables—gang membership and youth involvement in delinquency, violence, and substance use.

Gang risk. Gang involvement is measured by youth self-reports at the pretest interviews. The adolescent participants were asked whether they had ever been a gang member, they were currently a gang member, any of their family members currently belong to a gang, any of their family members had ever belonged to a gang, and any of their close friends are part of a gang. At pretest, 15% of the sample reported ever being a gang member and 7% reported current gang membership. Although in line with previous point estimates of the prevalence of gang membership (Boxer et al., 2017; Melde and Esbensen, 2011; Thornberry et al., 2003), the number of active gang members, about five per condition, is too small to support analysis of this important subpopulation separately. Instead, we developed a measure of "gang risk" by combining the available measures of gang involvement. If a youth reported ever being a gang member, currently being a gang member, having family members who were either current or past gang members, or who had close friends

---

5.   The characteristics of those who left the study were similar for treatment and control cases; that is, there were no significant treatment by attrition interactions on any pretreatment variable. The number of days elapsed from randomization to posttest was also similar for treatment and control participants, ranging from 215 to 224 days elapsed.

6.   Alpha reliability coefficients, percentage missing, and documentation of the source and item content for each scale used as outcome measures are available in Gottfredson et al. (2018).

who are part of a gang, they were coded as having high "gang risk." Youth who scored a 0 on all of these indicators are coded as having low gang risk. Of the 66 youth in the treatment group 34 (51.5%) were low gang risk and 32 (48.5%) were high gang risk; of the 63 youth in the control group, 31 (49.2%) were low and 32 (50.8%) were high gang risk.

In the analysis to follow, we test our hypotheses about the effectiveness of FFT-G separately for the low-gang-risk and the high-gang-risk participants. The aforementioned companion paper (Gottfredson et al., 2018) examined the impact of treatment on both the final outcomes and the mediators targeted by the FFT model for the full sample. The mediators include multiple measures of peer relationships, normative beliefs about the validity of rules/laws, constructive time use, family functioning, and parent behavior/parent substance use. In that paper, we assessed the costs and benefits of using a public funding stream (Medicaid) to implement an evidence-based practice and included fidelity and cost-effectiveness data as well as a comparison of treatment and control group differences on all measured outcomes and the full array of measured mediating variables. In brief, in the analysis, we found that FFT-G was not significantly related to any of the mediators but that it did have significant effects in the expected direction on several key outcome measures. All of the recidivism measures favored the treatment group at the 18-month follow-up, and significant differences ($p < .05$) were observed for the percent with drug charges, the percent adjudicated, and the percent with property charges. We have focused this article on different, albeit related questions. Is FFT-G effective for engaging and treating youth at risk for gang membership? Does it reduce gang membership? Is it as effective for reducing crime for youth at high risk for gang membership as it is for low-risk youth?

Outcome measures. The primary outcomes for this study are gang membership and youth offending. Posttreatment gang membership is measured by youth self-reports of whether he was a gang member at any point between the pretest and the 6-month follow-up interview. This measure is available at that interview but not at the 18-month follow-up when only official records were collected.

The variables related to criminal behavior are measured using both youth self-reports and official records. The 6-month follow-up interview included a 20-item general delinquency self-report index; we use a variety measure of *general delinquency*. We also ask about the frequency of involvement in *violent delinquency,* using a 5-item subset of the general delinquency index. Three indicators of substance use are included: a 2-item measure of frequency of *alcohol use*, a single item measure of frequency of *marijuana use*, and a variety measure of *substance use* based on a 13-item index of substance use ranging from alcohol through hard drugs. All of the self-report measures are based on measures from the Rochester Youth Development Study (Thornberry et al., 2003). At both the 6-month and 18-month follow-up periods, measures of criminal involvement from official records are available. Key indicators include the prevalence and frequency of arrest, the prevalence of felony charges,

crimes against the person charges, property crime charges, drug crime charges, the percent adjudicated delinquent, the prevalence of residential stays, and the number of days in residential placement. All data from court records were double coded by two researchers, and all discrepancies were resolved. The level of inter-rater reliability agreement prior to resolving discrepancies ranged from 94% to 100% across the different items coded and was above 97% for all but one of the variables coded.

Analysis

Statistical power. Based on an anticipated recidivism rate of 52% (Thornberry et al., 2003) and an expected ES of .58 (Alexander and Parsons, 1982; Lee et al., 2012), the total number of subjects needed to be able to reject the null hypothesis that the failure rates for experimental and control subjects are equal with probability (power) 0.8 and Type I error probability of 0.05 is 142. In these analyses, we assumed a continuity-corrected chi-squared statistic or Fisher's exact test to evaluate this null hypothesis. The final number of families included in the study was 129. Because the number of cases is slightly smaller than anticipated, and because more recent ES estimates of FFT have been somewhat lower than those obtained in earlier studies, the results from this study are regarded as meaningful if they reach the $p < .10$ level of statistical significance rather than the more conventional $p < .05$ level. At $p < .10$, the minimum detectable ES given the available number of cases is $d = .49$. Using a p-level of .10 is reasonable in this study given the prior expectation of a positive effect of the intervention. Although this practice is considered acceptable in experimental research in which the probability of Type I and Type II errors is balanced (Cohen, 1992; Weisburd, Petrosino, and Mason, 1993), exact p-levels are reported for those who wish to apply a more conservative test.

Outcome analysis. The distributional qualities of all outcome measures were examined, and extreme outliers were trimmed for a small number of scales. The pretreatment equivalence of the treatment and control group members was assessed using mean comparisons and t tests. An "intent to treat" (ITT) approach was used to compare outcomes for study groups. First, regression models were run using the model most appropriate for each outcome (e.g., logistic regression for binary outcomes, and negative binomial or Poisson regression for count outcomes). The model for each dependent variable included a dummy variable measuring assignment to the treatment condition, variables that differed significantly between the treatment groups at pretest, and the pretreatment measure of the outcome variable. These models were used to generate adjusted means for each outcome. For continuous outcome measures, standardized mean difference effects size statistics (d) were calculated using the difference between the treatment and the control group adjusted posttest mean in the numerator and the standard deviation for the corresponding unadjusted posttest measures in the denominator. Odds ratios are presented for binary outcomes. Interactions by gang risk status were tested by adding the gang risk indicator and the product of treatment and gang risk to each regression model.

Results

Program Implementation

FFT-G was delivered by six trained family therapists, two from each of the three partic-ipating agencies. All therapists were experienced FFT therapists, and their agencies had been previously certified to deliver FFT. The additional training for FFT-G consisted of 12 hours of training spread over 2 days provided by Dr. James Alexander, the developer of FFT, and assisted by the national consultant who provided weekly supervision during the implementation phase. All six therapists were in attendance for the entire training session. A 1-day follow-up training was provided by the national trainer approximately 2 months after the initial training.

Two trained research assistants documented the content of the initial training by coding, for each 10-minute period, what content had been covered in the period. The level of inter-rater agreement across time points for the initial training was 86.8%. The training consisted primarily of a review of the gang context in Philadelphia, gang research and risk and protective factors relevant to gang-involved populations, and a review of each stage of the FFT model with specific attention to special accommodations for gang-involved populations. The most common activity during the training was role playing how to address specific issues that were anticipated to come up in sessions involving gang-involved youth (30% of time periods). The FFT-G manual, developed by the FFT developer's team, formed the basis for the training and was provided to each therapist and supervisor. Three of the six trained therapists handled 85% of the cases.

FTTP, the family service to which control youths were referred, was far less successful at engaging families in therapy than was FTT-G. Only 11 families were successfully engaged.[7] The participating judge subsequently referred several control families to regular FFT when it became clear that they were not receiving services from FTTP. Of the 63 control families, 13 (20.6%) received FFT.[8] Thus, the contrast between the treatment and control groups is muted in this analysis. The findings from supplementary analysis for the full sample, but not done separately for low- and high-gang-risk youth because of small sample sizes, shows that the differences between the treatment and control groups are somewhat larger when the 13 control cases that received FFT are removed from the analysis (see Gottfredson et al., 2018). As a result, the ITT results presented here are conservative in regard to the effectiveness of FFT-G as a substantial portion of the control cases received regular FFT.

---

7.   The level of engagement observed for FTTP is probably typical for such programs when implemented with high-risk families. FFT's level of engagement is far higher than is typically obtained because FFT includes a major emphasis on engagement. Also, FFT therapists were always willing to meet the families in their homes, whereas FTTP used a combination of home- and office-based meetings.

8.   One of these families was inadvertently assigned to an FFT-G-trained therapist. The others were assigned to regular FFT therapists.

Outcomes

   Baseline comparisons. We compared the treatment and control groups—separately for the high-gang-risk and the low-gang-risk youth—on 36 baseline characteristics to see whether the groups differed prior to randomization (see Table S1 in the online supporting information). Starting with the low-gang-risk participants, only three comparisons (8.3%) were statistically significant at the .05 level. The FFT-G group has a lower percentage of children living with either their mother or father, were less likely to have been in residential placement during the previous 6 months, and self-reported a higher variety of substances use. For the high-gang-risk participants, three variables were significantly different: The FFT-G group had more felony charges, and the youth self-reported a higher variety of general delinquency and reported lower levels of attitudes unfavorable toward delinquency. In addition, two other variables were marginally significant ($p < .10$). Overall, the number of observed differences is about what is expected by chance and there is no discernible pattern to which variables were significantly different. The pretreatment scales that differed across groups are included as covariates in all outcome analyses.

Program Completion

The first hypothesis is aimed at examining whether FFT-G can be delivered as successfully to gang-involved youth as it is to youth who were not involved in street gangs. Table 1 presents the results about the level and quality of treatment received separately for the high-gang-risk and the low-gang-risk youth. These comparisons only involve the treatment group—participants who received FFT-G. Overall, there is no evidence that high-gang-risk youth and their families are more difficult to enroll and maintain in treatment.

   Of the low-gang-risk cases assigned to FFT-G, 85.3% ($n = 29$) began treatment as compared with 75.0% ($n = 24$) of the high-gang-risk cases. Although slightly fewer of the high-gang-risk cases started treatment, of those who did start, a higher percentage successfully completed treatment—70.1%—as compared with 62.1% of the low-gang-risk cases. In terms of the amount and fidelity of treatment, there are few differences between the low-gang-risk and high-gang-risk groups. There are three marginally significant differences ($p < .10$) in these comparisons: For all those who began FFT-G, the high-gang-risk cases have more face-to-face contact hours, and of those who successfully completed the program, the high-gang-risk cases have more face-to-face contact hours and more session hours. Overall, though, there are very few differences of any magnitude on these measures between the two groups, which is consistent with our first hypothesis that gang-involved youth and their families will engage with FFT and successfully complete FFT at the same level as nongang-involved youth.

Recidivism

Table 2 presents data on key outcomes measured in the 6-month youth interviews. For the low-gang-risk participants, none of the differences between the FFT-G group and the

T A B L E 1

Treatment Quality, by Gang Risk Status, FFT-G Participants Only

Variable	Low Gang Risk N = 34						High Gang Risk N = 32					
	Began FFT-G n = 29		Successfully Completed n = 18		Terminated n = 11		Began FFT-G n = 24		Successfully Completed n = 17		Terminated n = 7	
	Mean (SD)	Range or %	Mean (SD)	Range or %	Mean (SD)	Range or %	Mean (SD)	Range or %	Mean (SD)	Range or %	Mean (SD)	Range or %
Months Involved	3.2 (1.4)	.8–6.3	3.7 (1.2)	1.9–6.3	2.4 (1.4)	.8–5.1	3.2 (1.3)	.9–6.1	3.6 (1.1)	1.5–6.1	2.4 (1.4)	.9–4.1
Contacts	16.3 (9.4)	5–39	19.0 (8.2)	11–37	11.8 (10.0)	5–39	14.2 (6.3)	2–22	16.8 (4.2)	10–22	7.9 (6.0)	2–20
FFT-G sessions	8.5 (3.6)	1–14	10.8 (1.5)	9–14	4.6 (2.6)	1–10	9.0 (3.9)	1–14	11.2 (1.5)	9–14	3.7 (2.5)	1–7
Face-to-Face Contact Hours	13.2 (6.1)	2.3–23.5	17.1 (3.5)	7.8–23.5	6.8 (3.4)	2.3–14.0	16.6[a] (7.2)	2.0–29.0	19.4[a] (3.8)	13.7–25.4	9.9 (9.2)	2.0–29.0
Session Hours	11.3 (5.4)	2.2–21.2	14.8 (3.4)	7.5–21.2	5.7 (2.4)	2.2–10.2	13.6 (6.1)	2.0–23.5	16.8[a] (3.2)	11.3–23.5	5.8 (4.1)	2.0–11.5
FFT-G phase (N, %)												
Engagement/Motivation	3.3	49.7%	3.4	31.5%	3.1	79.6%	3.2	46.3%	3.5	30.9%	2.6	83.7%
Behavior Change	3.3	32.6%	4.4	40.7%	1.4	19.5%	3.4	32.6%	4.4	39.3%	1.1	16.3%
Generalization	1.9	17.6%	3.0	27.8%	.1	.9%	2.4	21.1%	3.4	29.8%	0.0	0.0%
Fidelity	4.2 (.7)	2.8–5.3	4.3 (.6)	2.8–5.0	4.1 (.7)	3.0–5.33	3.9 (.9)	2.0–5.0	4.1 (.8)	2.2–4.9	3.6 (1.0)	2.0–5.0

Notes. Only cases that had at least one FFT-G session are included. Ten cases were never contacted, and three contacted cases did not begin FFT-G sessions because they were placed in residential settings before the first session. Standard deviations in parentheses.

[a]This value is higher for high risk than low risk (*p* < .10).

T A B L E   2

Posttest Interview Outcomes, Adjusted, by Treatment Group and Gang Risk Status

Scale	Adjusted Posttest Mean or %					
	Treatment (*n*)		Control (*n*)		*p*-level	*Effect Size*	Odds Ratio
	Low Gang Risk					
Delinquency						
General delinquency variety	.06	32	.05	27	.271	.21	—
Violent delinquency frequency	3.43	32	2.90	27	.136	.08	—
% arrested or picked up by police in past 6 mos	37.50	32	25.92	27	.521	.24	1.56
% in residential past 6 mos	40.62	32	48.15	27	.782	−.15	.83
Substance Use of Adolescent						
Drug variety	1.19[b]	32	.81	27	.254	.30	—
Marijuana frequency	3.81	32	1.98	27	.762	.66	—
Alcohol frequency	1.08[c]	32	.86	27	.469	.14	—
Gang Involvement						
% currently in gang	.00	32	.00	28	—	—	—
	High Gang Risk					
Delinquency						
General delinquency variety	.12[*d]	28	.12	30	.015	−.04	—
Violent delinquency frequency[a]	3.34	28	3.21	30	.520	.02	—
% arrested or picked up by police in past 6 mos	39.28	28	43.33	30	.157	−.08	.38
% in residential past 6 mos	35.71[+]	28	46.67	30	.075	−.22	.30
Substance Use of Adolescent						
Drug variety	1.11[+b]	28	1.47	30	.064	−.24	—
Marijuana frequency	2.47	28	2.36	30	.701	.04	—
Alcohol frequency[a]	.41[*c]	28	1.03	30	.050	−.42	—
Gang Involvement						
% currently in gang	7.14	28	10.00	30	.134	−.10	.07

Notes. Adjusted means are predicted values from regressions that control for general delinquency variety, % in residential, hard drug variety, hard drug frequency, lives with mother or father, and attitudes unfavorable toward delinquency, as well as the pretreatment measure of the outcome variable. ES estimates are standardized mean differences for continuous outcomes and odds ratios for binary outcomes.
[a]One extreme outlier trimmed.
[b]Interaction of treatment by gang risk status, $p < .05$.
[c]Interaction of treatment by gang risk status, $p < .10$.
[d]The identical means are significantly different due to control for pretreatment scores.
[*]$p < .05$. [+]$p < .10$.

control group attain statistical significance. For the high-gang-risk participants, however, two of the nine comparisons are statistically significant and two others are marginally significant, all favoring the FFT-G group. Those who received the treatment have lower levels of general delinquency, drug use, and alcohol use, and a lower percentage of this group reported spending time in residential placement. Tests of statistical interaction by gang risk status are significant or marginally significant for drug use and for alcohol use.

T A B L E   3

Six-Month Recidivism Data from Official Records, Adjusted, by Treatment Group and Gang Risk Status

	Low Gang Risk						
	FFT-G		Treatment as Usual				
Variable	%/Mean	(*n*)	%/Mean	(*n*)	*p*-level	*Effect Size*	Odds Ratio
% arrested	11.76	34	20.00	30	.602	—	.65
# arrests	.15	34	.20	30	.792	−.12	—
% with felony charges	8.82	34	10.00	30	.884	—	.86
% with person charges	5.88	34	.00	30	.998	—	—[a]
% with property charges	2.94	34	3.33	30	.982	—	.00
% with drug charges	.00	34	13.33	30	.998	—	.00
% adjudicated delinquent	.00	34	13.33	30	.997	—	.00
% with residential stays	41.18[+]	34	26.67	30	.060	—	3.60
Days in residential placement	36.23	34	48.48	30	.742	−.19	—

	High Gang Risk						
	FFT-G		Treatment as Usual				
Variable	%/Mean	(*n*)	%/Mean	(*n*)	*p*-level	*Effect Size*	Odds Ratio
% arrested	18.75	32	16.67	30	.568	—	.60
# arrests	.19	32	.23	30	.306	−.06	—
% with felony charges	12.50[+]	32	16.67	30	.071	—	.11
% with person charges	9.38[*]	32	13.33	30	.047	—	.04
% with property charges	9.38[*]	32	16.67	30	.016	—	.02
% with drug charges	6.25	32	3.33	30	.391	—	3.45
% adjudicated delinquent	12.50	32	16.67	30	.261	—	.35
% with residential stays	28.12	32	40.00	30	.160	—	.38
Days in residential placement	29.78	32	84.99	30	.204	−.16	—

Notes. Adjusted means are predicted values from regressions that control for the three variables that differed significantly between the treatment groups at pretest (general delinquency variety, % in residential, and hard drug variety) and the pretreatment measure of the outcome variable. For comparisons among high-gang-risk cases, felony charges adjudicated delinquent and attitudes unfavorable toward delinquency are also controlled. ES estimates are standardized mean differences. No gang risk by treatment interactions were statistically significant at the $p < .10$ level.
[a] Cell sizes too small for meaningful OR calculation.
[*] $p < .05$. [+] $p < .10$.

Although the prevalence of gang membership at the 6-month mark is in the hypothesized direction (7.14% for FFT-G vs. 10% for the control condition), it does not attain statistical significance ($p = .13$).

The immediate posttreatment results based on official records are consistent with the interview results (Table 3). For the low-gang-risk participants, only one comparison between the treated and control conditions obtains marginal significance: the percent with residential

stays. For these participants, however, the difference favors the control group: 26.7% versus 41.2%. In contrast, for the high-gang-risk participants, seven of the nine comparisons favor the FFT-G group, and of these, two are statistically significant and one is marginally significant. The FFT-G cases, as compared with the control cases, have lower prevalence rates of felony charges, crimes against person charges, and property crime charges. Tests of statistical interaction by gang risk status did not reach statistical significance for any of these measures.

Next we examined recidivism data only between month 7 (immediate posttreatment) and month 18 (Table 4). Recall that the initial motivation for this study was to evaluate a gang prevention/intervention program using Blueprints standards so that, if effective, it could eventually be identified as a Model program. Blueprints criteria for Model programs require sustained effects postprogram. As such, Table 4 is broken out to examine findings independent of and sustained beyond the program period. None of the comparisons between the treatment group and the control group are statistically significant for the low-gang-risk youth, but three of them are significant for the high-gang-risk youth and one is marginally significant, all favoring the FFT-G group. Those who received FFT-G had a lower prevalence and frequency of arrests and were less likely to have felony charges and crimes against person charges. Tests of statistical interaction by gang risk status are significant for these same measures.

Finally, Table 5 presents results for the entire 18-month period after random assignment. The group differences for the entire time period gain in strength relative to the initial 6-month period but only for the high-gang-risk youth. For the low-gang-risk youth, the FFT-G cases have a marginally lower prevalence of drug charges, 11.8%, as compared with the control cases, 26.7%, at the 18-month mark. In contrast, for the high-gang-risk youth, all nine comparisons favor the FFT-G cases and five of the nine are statistically significant and one nearly so. By 18 months after randomization, and 12 months after the end of treatment, FFT-G participants have a lower prevalence of arrest; number of arrests; fewer felony charges, crimes against person charges, and property crime charges; and are less likely to be adjudicated delinquent. In general, the observed differences are large and clinically meaningful. For example, the arrest rate increases from 28% to 43% from the treatment to the control conditions. Also, although the difference is not statistically significant, the FFT-G cases spend less than half as many days (111 vs. 265) in residential placement as compared with the control cases. For three of these outcomes—frequency of arrests, percent with felony charges, and percent with person charges—there is a significant interaction of treatment by gang risk status.

Discussion

Street gangs remain a serious problem in American society. Gang membership is associated with increased involvement in criminal behavior, especially in serious and violent crimes, and has a host of negative consequences for the development of adolescents who become

T A B L E   4

Recidivism for Months 7–18 From Official Records, Adjusted, by Treatment Group and Gang Risk Status

| | Low Gang Risk | | | | | | |
| | FFT-G | | Treatment as Usual | | | | |
Variable	%/Mean	(n)	%/Mean	(n)	p-level	Effect Size	Odds Ratio
% arrested	35.29[a]	34	30.00	30	.709	—	1.27
# arrests	.44[a]	34	.33	30	.974	.18	—
% with felony charges	29.41[a]	34	16.67	30	.501	—	1.62
% with person charges	20.59[a]	34	13.33	30	.347	—	2.06
% with property charges	8.82	34	10.00	30	.976	—	.97
% with drug charges	11.76	34	13.33	30	.193	—	.20
% adjudicated delinquent	23.53	34	26.67	30	.873	—	.90
% with residential stays	64.70	34	63.33	30	.745	—	1.22
Days in residential placement	98.37	34	179.76	30	.402	−.46	—

| | High Gang Risk | | | | | | |
| | FFT-G | | Treatment as Usual | | | | |
Variable	%/Mean	(n)	%/Mean	(n)	p-level	Effect Size	Odds Ratio
% arrested	12.50[*][a]	32	30.00	30	.039	—	.15
# arrests	.16[*][a]	32	.43	30	.017	−.45	—
% with felony charges	6.25[*][a]	32	30.00	30	.032	—	.11
% with person charges	3.12[+][a]	32	23.33	30	.060	—	.08
% with property charges	6.25	32	20.00	30	.152	—	.23
% with drug charges	3.12	32	13.33	30	.131	—	.06
% adjudicated delinquent	12.50	32	23.33	30	.117	—	.26
% with residential stays	46.88	32	63.33	30	.154	—	.40
Days in residential placement	87.71	32	202.84	30	.213	−.18	—

Notes. Adjusted means are predicted values from regressions that control for the three variables that differed significantly between the treatment groups at pretest (general delinquency variety, % in residential, and hard drug variety) and the pretreatment measure of the outcome variable. For comparisons among high-gang-risk cases, felony charges adjudicated delinquent and attitudes unfavorable toward delinquency are also controlled. ES estimates are standardized mean differences.
[a]Interaction of treatment by gang risk status, $p < .05$.
[*]$p < .05$. [+]$p < .10$.

enmeshed in the world of gangs. Unfortunately, no evidence-based gang programs meet the rigorous standards promulgated by Blueprints for Healthy Youth Development to respond to this challenge. That is not the case for related problem behaviors, especially general delinquency and substance use. The purpose of this project was to build on the success of those programs to address the particular needs of gang members. In particular, by working with the developers of Functional Family Therapy, we established an accommodation of FFT that maintained all the basic components and approaches of FFT while enhancing the

T A B L E 5

Eighteen-Month Recidivism Data From Official Records, Adjusted, by Treatment Group and Gang Risk Status

	Low Gang Risk						
	FFT-G		Treatment as Usual				
Variable	%/Mean	(*n*)	%/Mean	(*n*)	*p*-level	*Effect Size*	Odds Ratio
% arrested	41.18	34	50.00	30	.813	—	.87
# arrests	.59[a]	34	.53	30	.887	.08	—
% with felony charges	32.35[a]	34	26.67	30	.742	—	1.24
% with person charges	23.53[a]	34	13.33	30	.242	—	2.45
% with property charges	11.76	34	13.33	30	.787	—	.78
% with drug charges	11.76[+]	34	26.67	30	.050	—	.12
% adjudicated delinquent	23.53	34	40.00	30	.262	—	.47
% with residential stays	67.65	34	63.33	30	.463	—	1.60
Days in residential placement	134.24	34	224.68	30	.618	−.68	—

	High Gang Risk						
	FFT-G		Treatment as Usual				
Variable	%/Mean	(*n*)	%/Mean	(*n*)	*p*-level	*Effect Size*	Odds Ratio
% arrested	28.12[*]	32	43.33	30	.064	—	.24
# arrests	.34[*a]	32	.67	30	.014	−.41	—
% with felony charges	18.75[*a]	32	43.33	30	.011	—	.10
% with person charges	12.50[*a]	32	33.33	30	.023	—	.12
% with property charges	15.62[*]	32	33.33	30	.031	—	.14
% with drug charges	9.38	32	16.67	30	.549	—	.58
% adjudicated delinquent	21.88[+]	32	36.67	30	.086	—	.30
% with residential stays	50.00	32	63.33	30	.281	—	.51
Days in residential placement	111.29	32	265.12	30	.190	−.31	—

Notes. Adjusted means are predicted values from regressions that control for the three variables that differed significantly between the treatment groups at pretest (general delinquency variety, % in residential, and hard drug variety) and the pretreatment measure of the outcome variable. For comparisons among high-gang-risk cases, felony charges adjudicated delinquent and attitudes unfavorable toward delinquency are also controlled. ES estimates are standardized mean differences.
[a]Interaction of treatment by gang risk status, $p < .05$.
[*]$p < .05$. [+]$p < .10$.

program to address specific risk factors and group processes associated with street gangs. The accommodation—FFT-G—was manualized and evaluated in an RCT using a sample of adjudicated youth in the Philadelphia Family Court containing a large proportion of youth at risk for gang involvement. We tested three hypotheses.

There is clear support for the first hypothesis that youth at high risk for gang membership and their families will engage with and successfully complete FFT-G at the same level as nongang-involved youth. There were no significant or discernible differences between the

low-gang-risk and high-gang-risk participants with respect to starting treatment in FFT-G, successfully completing it, or the amount and fidelity of treatment. Also, the national FFT office reported that the level of program participation and completion for the high-gang-risk youth is comparable with that observed for regular FFT families, both in Philadelphia and elsewhere. In general, high-gang-risk youth are no different than other youth in terms of their amenability to this program.

These results stand in contrast to those reported by Boxer and colleagues (Boxer, 2011; Boxer et al., 2015) with respect to MST where gang-involved youth were substantially less likely to complete treatment. In that case, however, there was no effort to change the basic MST model to address the specific needs of gang-involved youth and the sample was not selected with a focus on gang involvement. In the present case, because of the seriousness and extensiveness of the risk factors associated with the families of high-gang-risk youth (Hennigan, Kolnick, Vindel, and Maxson, 2015; Raby and Jones, 2016; Thornberry et al., 2003), the FFT-G accommodation emphasized the importance of "family engagement" in training and supervision sessions and in practice. That emphasis seems to have been rewarded as our results indicate that the FFT-G approach worked just as well with the high-gang-risk youth as with others.

In contrast, there is no support for the second hypothesis that FFT-G would reduce the level of gang membership. Our ability to assess this hypothesis, however, was seriously compromised. Because of funding limitations, we were unable to conduct interviews at the 18-month follow-up, and therefore, the only data on this outcome came from the interviews conducted 6 months after randomization. As noted earlier, point estimates of the prevalence of gang membership are low with an annual rate of approximately 5% (Boxer et al., 2017; Melde and Esbensen, 2013; Thornberry et al., 2003). This finding is consistent with our own; the rate of current gang membership at the 6-month time point was 4% for the full sample of participants. This rate is so low that it was not possible to assess group differences. As such, we did not find support for this hypothesis.

There is clear support for the third hypothesis that FFT-G will reduce involvement in criminal behavior, but this support is evident *only for the high-gang-risk youth*. For the low-gang-risk youth, there are few significant differences between the FFT-G group and the control group. For the high-gang-risk youth, however, there is a consistent pattern of differences between the treated and control participants that grows in magnitude over time. Differences in the expected direction are observed at the 6-month mark—reflecting during-treatment effects, between months 7 and 18—reflecting posttreatment effects, and for the entire 18-month study duration—reflecting the overall impact of FFT-G. The differences are most dramatic for the full 18-month period where all nine outcomes favor the FFT-G group, with five of them being statistically significant, and one nearly so. In general, for the high-gang-risk youth, FFT-G was able to reduce criminal offending according to official records.

The differences between the FFT-G and control participants are likely to comprise a conservative estimate of the program effect for two reasons. First, as noted, subsequent to

random assignment, the judge ordered 13 of the control cases to receive regular FFT. Thus, the difference between the two conditions was substantially muted. Indeed, in analyses not reported here, we found that when these 13 cases are removed from the control group, the treatment effect is enhanced. Second, the analysis for the high-gang-risk participants is based on only half the sample, and therefore, there is considerably less statistical power than initially estimated. Nevertheless, even for this smaller sample, substantial treatment effects emerge (Gottfredson and Devlin, 2017).

The fact that FFT-G was effective for the high-gang-risk youth, but not for the low-gang-risk youth, is intriguing. One possible explanation for this outcome is that the gang accommodation of FFT impacted the therapists and the delivery of services in the intended way. Even though all components and approaches of the basic FFT model were included in the program as delivered, the emphasis of FFT-G on gang members and gang issues permeated discussions with the agencies providing the services and with their therapists. This was especially so in the special training and supervision sessions provided as part of this effort. The FFT-G manual, which was used by the therapists at the training sessions, emphasized gang issues and used numerous gang illustrations, as did almost all of the role playing. Even though the FFT-G therapists reported reverting to core FFT principles in their interactions with the families, the overall emphasis of this program on gang issues may have permeated the process and may have led to the greater success with high-gang-risk youth. In other words, absent our emphasis on street gangs in developing and initiating this program, the success of the program with high-gang-risk youth may not have been evident.

This finding is also consistent with the "risk principle" in prevention science, which argues that intervention programs are most effective with higher risk youth, not, as is often assumed, with lower risk youth (Andrews et al., 1990). In several evaluation studies, scholars reported results consistent with this principle: "[W]hen actually explored, the effects of treatment typically are found to be greater among higher risk cases than among lower risk cases" (Andrews et al., 1990: 374). In his meta-analysis of interventions for juvenile offenders, Lipsey (2009) found that the strongest predictor of recidivism effect sizes was delinquency risk, with greater reductions of recidivism associated with higher risk youth. One explanation for this pattern of findings is that higher risk youth and their families have more and greater needs that effective services can improve, thereby yielding a larger impact on recidivism. Interestingly, in their evaluation of the G.R.E.A.T. program, Esbensen and colleagues (2013) also found a larger impact for youth at greater risk of gang membership but only during the earlier waves of their follow-up. This finding is consistent with our results, which ended at the 18-month mark. Thus, our finding that FFT-G was most effective for the high-gang-risk cases could be a reflection of the risk principle, of the gang emphasis in the FFT-G accommodation, or of some combination of the two.

It should also be noted that a detailed cost analysis indicated that FFT-G services replaced more expensive services to a greater extent for treatment than for control

youth, thereby making the cost per youth served nearly $2,800 less for FFT-G than for treatment-as-usual cases (Gottfredson et al., 2018). Thus, not only is FFT-G effective in reducing recidivism, but it is also a cost-effective intervention. Overall, the findings presented here have several implications for future research and for policy development. We briefly discuss some of the more salient issues here.

Future Research

One important topic for future research is to see if, and under what conditions, existing evidence-based (EB) programs for delinquency prevention are also effective when delivered to gang members and/or to youth at high risk for gang membership. A program of research is needed to provide systematic evaluation of existing EB programs such as FFT, MST, and MTFC without modification or accommodation for gang-involved populations. All of these programs have gang members in their typical client base, but there is not adequate evidence about their effectiveness for this important subpopulation. Indeed, we have somewhat conflicting results for this issue; Boxer and colleagues (2015) reported that gang-involved youth had low successful completion rates for MST, whereas we found that FFT-G can be delivered equally well to youth at low and high risk for gang membership and is effective for the high-gang-risk population. Finding out whether basic EB programs are effective when offered specifically to gang-involved youth should be a high priority.

This project represents the first effort to accommodate an EB program to focus on gang issues. In the early stages of the project, the Advisory Board developed a general strategy for identifying appropriate EB programs for this purpose and for working with the developers to modify their approach. This strategy should be expanded to other EB programs such as MST and MTFC, and the search for other potential partners should continue. Evaluating these modified programs will expand our knowledge base about potentially effective gang programs. Also, the modifications to some of these programs may be more extensive than the ones made to FFT and, therefore, offer a starker contrast between evaluating an existing evidence-based program and a modified one. The results of the current project are promising and imply that there is value in this strategy of modifying and evaluating EB programs.

It is not yet clear whether a standard EB program delivered to gang-involved youth differs in impact from an accommodation of the same program. For example, we do not know whether FFT-G would yield different (better?) results than basic FFT. An RCT with three conditions—FFT, FFT-G, and TAU—would address this question and provide fuller information about the use of the FFT model for a gang population. This strategy could, of course, be used with other EB programs. Obviously, this would be an ambitious research agenda, warranted only if the results from the prior two research topics indicate its value.

A second important topic for future research is to evaluate these programs using a sample of active gang members. As noted, we had difficulty recruiting a large number

of gang members into our study. We estimate that there would have been enough gang members for the project if we had had access to all (or even most) of the courtrooms in the Family Court, but we encountered several practical limitations to that access. Some judges seemed uninterested in the project in general, and as noted, many court personnel, especially probation officers, were reluctant to be involved in a "gang project" because of potential labeling effects. Philadelphia is certainly not alone in that regard; in many jurisdictions, there is a denial of gang activity and a reluctance to label youth in schools or Family Court settings as gang members or even as at risk of gang membership.

This raises an important challenge for future research. On the one hand, it is clear that we need to focus services and evaluations specifically on gang members given their disproportionate involvement in crime and serious crime. On the other hand, an overt focus on this issue may raise practical problems with access to referrals and create ethical problems with respect to the risks to which participants are exposed. To be labeled a "gang member" is not trivial and carries substantial risk, for example, U.S. immigration policy and the use of suspected gang ties to justify deportations. Researchers have an ethical responsibility to be sensitive to these concerns. Indeed, once we understood the prevailing views at the Philadelphia Family Court about the gang label, we began to downplay the gang focus of the project when dealing with court personnel and emphasized the fact that we were evaluating one of the basic EB programs that the court often used for its charges. The study-involved youth were never formally labeled as "gang members" in any court documents, and the judge's order simply read that they were referred to "family therapy services," not to a gang program. It would be helpful if researchers in this area as well as funding agencies specifically addressed this issue in the future to develop appropriate responses to resolve this dilemma. Future projects may also wish to consult with ethicists, youth advocates, and similar colleagues in addressing this question. However this issue is resolved, an important topic for future research is to examine the effectiveness of programs with active gang members.

Policy Implications

As noted in the beginning of the article, until recently, no EB programs met the strong standards of effectiveness such as those promulgated by the Blueprints for Healthy Youth Development. When asked what EB programs work for gang prevention and intervention, the answer was either that "nothing works" or that we do not know if anything works. Fortunately, this situation seems to be changing. The present results indicate that FFT-G has significant deterrent effects on criminal behavior for individuals at risk for gang membership. Positive effects were evident for several outcomes and persisted at least 18 months posttreatment. Boxer and colleagues (2017) reported that MST was as effective for gang-involved youth as for nongang-involved youth in terms of rearrest 12 months after discharge, and the findings from several studies have demonstrated the overall effectiveness of MST in RCTs

(Henggeler et al., 2009). Esbensen and colleagues (2013) found that the revised G.R.E.A.T. program, a universal prevention program, reduced gang membership and a variety of other important outcomes. All of these results are encouraging and, we hope, generate increased enthusiasm to continue the effort to develop and evaluate EB programs for gang-involved youth. It is also important to replicate the RCT conducted here in a different site, as called for in the Blueprints standards.

Our core finding that FFT-G had strong deterrent effects for high-gang-risk youth but not for low-gang-risk youth also has important policy implications. It suggests that the juvenile justice system can and should expand the use of effective EB community programs to higher risk youth. This would include youth identified as being "at risk" because of their gang involvement. All too often decision makers make conservative decisions and limit community-based programs to lower risk youth, for example, those who are not involved with gangs. This approach flies in the face of the present findings and of the risk principle (Andrews et al., 1990), which argues, with compelling empirical evidence behind it, that EB programs are most effective with higher, not lower, risk youth. An important task for future research would be to move up the risk continuum systematically in selecting clients for EB community programs and evaluate the results. We recognize that community-based programs are not appropriate for everyone, but given their efficaciousness and cost-effectiveness, it is important to expand their reach whenever possible.

We also note that FFT-G was effective in reducing the use of residential placements, especially for the high-gang-risk participants. Although the differences between the treated and control groups were not statistically significant, they are clinically meaningful. For example, for the high-gang-risk youth, the TAU participants spent more than twice as many days in residential confinement as the FFT-G participants. There is now a wealth of research findings that underscore the negative effects and outcomes for youth placed in lengthy confinement (Beck, Cantor, Hartge, and Smith, 2013; Gilman et al., 2014; Wasserman, Ko, and McReynolds, 2004). The National Research Council (NRC), Committee on Assessing Juvenile Justice Reform (2013), concluded that institutional treatment programs, by and large, do not produce reductions in reoffending. The findings from several studies reveal that in-home probation and services are equally or more effective in reducing recidivism than is out-of-home placement, even among violent adolescent offenders (Loughran et al., 2009; Ryan, Abrams, and Huang, 2014). Furthermore, youth in facilities that are underresourced and overcrowded often have high rates of recidivism and low rates of educational or vocational attainment. Even when residential placement is warranted and therapeutic services are evidence based and developmentally appropriate, study findings have shown that placement stays beyond the minimum needed to deliver intensive services have the potential to impede prosocial development and result in other potential harms (NRC, 2013, 2014).

Limitations

As is true of all projects, the current one has its limitations. The two major ones have already been discussed. The project would have been stronger if there had been a clearer delineation between regular FFT and FFT-G. Also, the proportion of active gang members in the sample was too small to support separate analysis of that important group.

In addition, the project would have been stronger if we had been able to conduct follow-up interviews at the 18-month mark. Doing so would have greatly enhanced our ability to assess the second hypotheses—that FFT-G would reduce gang membership— which is obviously a central issue in this line of research. Unfortunately, funding was not available for those interviews and we had to rely on official measures of recidivism that do not include any clear indicator of gang membership or gang involvement. Also, given the large impact observed at 18 months, it is important to know how these effects impacted criminal activity during the early adult years, as youth transition into adulthood. These longer term recidivism measures could provide important cost-effectiveness information, further supporting the value of preventive efforts with high risk youth.

Finally, the study would have been stronger if there had been less contamination of the study conditions. As noted, 20% of FFT-G subjects did not receive FFT-G and 21% of control subjects received FFT. This contamination most likely resulted in an understatement of FFT-G effects because subjects were treated in all analyses as they had been assigned rather than according to the services they received.

Conclusion

We believe that the present study makes a substantial contribution to the development of effective programming to respond to the problem of street gangs. By building on the strengths of existing EB programs for delinquency prevention, we developed a model for modifying existing programs to deal specifically with issues pertinent to gangs and gang members. We then worked with the developers of one of the most widely disseminated EB programs, Functional Family Therapy, to develop an accommodation of their approach to address gang-related issues. Finally, we successfully conducted an RCT to evaluate its effectiveness using a sample containing a large proportion of high-gang-risk youth. Two important conclusions can be drawn from this work. First, the rate at which high-gang-risk youth and their families successfully engage with and complete this community-based program is no different from that of low-gang-risk youth or the more general clientele of FFT programs. Although these families often have many serious and interwoven risk factors, they are clearly reachable by well-structured, community-based programs. Second, FFT-G demonstrated a clear deterrent effect on criminal offending for the high-gang-risk youth. Eighteen months after the program began, and 12 months after treatment ended, the FFT-G group had significantly lower recidivism rates as compared with those of the control group. Clearly, high-gang-risk youth benefited from FFT-G.

At a more general level, we hope that the results of this project stimulate increased efforts to develop effective, scientifically demonstrated programs to prevent gang membership and to reduce the impact of street gangs on its members and on society. For too long, gang programming has been divorced from the broader movement of prevention science. We believe that the present results, and those of Esbensen and colleagues (2013) and of Boxer and colleagues (2017), demonstrate that that should no longer be the case. Evidence-based gang programs based on rigorous scientific standards can and should be developed. The severity of the gang problem requires no less.

References

Alexander, James F. and Bruce V. Parsons. 1982. *Functional Family Therapy: Principles and Procedures*. Carmel, CA: Brooks and Cole.

Alexander, James F., Christie Pugh, and Bruce V. Parsons. 1998. Functional Family Therapy. In (Delbert S. Elliott, Series ed.), *Blueprints for Violence Prevention* (Book 3). Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado.

Andrews, Donald A., Ivan Zinger, Robert D. Hoge, James Bonta, Paul Gendreau, and Francis T. Cullen. 1990. Does correctional treatment work? A clincally relevant and psychologically informed meta-analysis. *Criminology*, 28: 369–404. https://doi.org/10.1111/j.1745-9125.1990.tb01330.x.

Augustyn, Megan B., Terence P. Thornberry, and Marvin D. Krohn. 2014. Gang membership and pathways to maladaptive parenting. *Journal of Research on Adolescence*, 24: 252–267.

Barton, Cole and James F. Alexander. 1981. Functional Family Therapy. In (Alan S. Gurman and David P. Kniskern, eds.), *Handbook of Family Therapy*. New York: Brunner.

Barton, Cole, James F. Alexander, Holly Waldron, Charles W. Turner, and Janet Warburton. 1985. Generalizing the treatment effects of Functional Family Therapy: Three replications. *American Journal of Family Therapy*, 13: 16–26.

Battin, Sara R., Karl G. Hill, Robert D. Abbott, Richard F. Catalano, and J. David Hawkins. 1998. The contribution of gang membership to delinquency beyond delinquent friends. *Criminology*, 36: 93–115.

Beck, Allen J., David Cantor, John Hartge, and Tim Smith. 2013. *Sexual Victimization in Juvenile Facilities Reported by Youth, 2012* (NCJ 241708). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.

Boxer, Paul. 2011. Negative peer involvement in Multisystemic Therapy for the treatment of youth problem behavior: Exploring outcome and process variables in "real-world" practice. *Journal of Clinical Child and Adolescent Psychology*, 40: 848–854.

Boxer, Paul, Meagan Docherty, Michael Ostermann, Joanna Kubik, and Bonita Veysey. 2017. Effectiveness of Multisystemic Therapy for gang-involved youth offenders: One year follow-up analysis of recidivism outcomes. *Children and Youth Services Review*, 73: 107–112.

Boxer, Paul, Joanna Kubik, Michael Ostermann, and Bonita Veysey. 2015. Gang involvement moderates the effectiveness of evidence-based intervention for justice-involved youth. *Children and Youth Services Review*, 52: 26–33.

Chamberlain, Patricia. 2003. *Treating Chronic Juvenile Offenders: Advances Made Through the Oregon Multidimensional Treatment Foster Care Model*. Washington, DC: American Psychological Association.

Cohen, Jacob. 1992. A power primer. *Psychological Bulletin*, 112: 155–159.

Commonwealth of Pennsylvania, Juvenile Court Judges' Commission. 2015. *2015 Pennsylvania Juvenile Court Dispositions*. Harrisburg, PA: Center for Juvenile Justice Training and Research. Retrieved from jcjc.pa.gov/Research-Statistics/Disposition%20 Reports/2015%20Pennsylvania%20Juvenile%20Court%20Disposition%20Report% 20(PDF).pdf.

Darnell, Adam J. and Megan S. Schuler. 2015. Quasi-experimental study of Functional Family Therapy effectiveness for juvenile justice aftercare in a racially and ethnically diverse community sample. *Children and Youth Services Review*, 50: 75–82.

Drug Enforcement Agency. 2015. *2015 National Drug Threat Assessment Summary*. Washington, DC: U.S. Department of Justice. Retrieved from dea.gov/docs/2015% 20NDTA%20Report.pdf.

Egley Jr., Arlen, James C. Howell, and Meena Harris. 2014. *Highlights of the 2012 National Youth Gang Survey*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. Retrieved from ojjdp.gov/pubs/248025.pdf.

Esbensen, Finn-Aage, D. Wayne Osgood, Dana Peterson, Terrance J. Taylor, and Dena C. Carson. 2013. Short- and long-term outcome results from a multisite evaluation of the G.R.E.A.T. Program. *Criminology & Public Policy*, 12: 375–411.

Esbensen, Finn-Aage, Dana Peterson, Terrance J. Taylor, and Adrienne Freng. 2010. *Youth Violence: Sex and Race Differences in Offending, Victimization, and Gang Membership*. Philadelphia, PA: Temple University Press.

Fagan, Abigail A. and Molly Buchanan. 2016. What works in crime prevention? Comparison and critical review of three crime prevention registries. *Criminology & Public Policy*, 15: 617–650.

Gilman, Amanda B., Karl G. Hill, and J. David Hawkins. 2014. Long-term consequences of adolescent gang membership for adult functioning. *American Journal of Public Health*, 104: 938–945.

Gottfredson, Denise C. and Deanna Devlin. 2017. *Functional Family Therapy for Gang Populations: Program Effects*. Paper presented at the annual meeting of the American Society of Criminology, Philadelphia.

Gottfredson, Denise C., Brook Kearley, Terence P. Thornberry, Molly Slothower, Deanna Devlin, and Jamie J. Fader. 2018. Scaling-up evidence-based programs using a public funding stream: A randomized trial of Functional Family Therapy for court-involved youth. Submitted for publication.

Hartnett, Dan, Alan Carr, Elena Hamilton, and Gary O'Reilly. 2017. The effectiveness of Functional Family Therapy: A meta-analysis. *Family Process*, 56: 607–619.

Haviland, Amelia M. and Daniel S. Nagin. 2005. Causal inferences with group based trajectory models. *Psychometrika*, 70: 1–22.

Henggeler, Scott W., Sonja K. Schoenwald, Charles M. Borduin, Melisa D. Rowland, and Phillippe B. Cunningham. 2009. *Multisystemic Therapy for Antisocial Behavior in Children and Adolescents*. New York: Guilford Press.

Hennigan, Karen M., Kathy A. Kolnick, Flor Vindel, and Cheryl L. Maxson. 2015. Targeting youth at risk for gang involvement: Validation of a gang risk assessment to support individualized secondary prevention. *Children and Youth Services Review*, 56: 86–96.

Howell, James C. 2000. *Youth Gang Programs and Strategies*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention. Retrieved from ncjrs.gov/pdffiles1/ojjdp/171154.pdf.

Huizinga, David H. 1996. *The Influence of Delinquent Peers, Gangs, and Co-offending on Violence* (Fact Sheet). Washington, DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.

Klein, Malcolm W. 1995. Attempting gang control by suppression: The misuse of deterrence principles. In (Malcolm W. Klein, Cheryl L. Maxson, and Jody Miller, eds.), *The Modern Gang Reader*. Los Angeles, CA: Roxbury.

Klein, Malcolm W. and Cheryl L. Maxson. 2006. *Street Gang Patterns and Policies*. New York: Oxford University Press.

Krohn, Marvin D. and Terence P. Thornberry. 2008. Longitudinal perspectives on adolescent street gangs. In (Akiva M. Liberman, ed.), *The Long View of Crime: A Synthesis of Longitudinal Research*. New York: Springer.

Krohn, Marvin D., Jeffrey T. Ward, Terence P. Thornberry, Alan J. Lizotte, and Rebekah Chu. 2011. The cascading effects of adolescent gang involvement across the life course. *Criminology*, 49: 991–1028.

Lee, Stephanie, Steve Aos, Elizabeth K. Drake, Annie Pennucci, Marna G. Miller, and Laurie Anderson. 2012. *Return on Investment: Evidence-Based Options to Reduce Statewide Outcomes*. Olympia: Washington State Institute for Public Policy.

Lipsey, Mark W. 2009. The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic review. *Victims & Offenders*, 4: 124–147.

Loughran, Thomas A., Edward P. Mulvey, Carol C. Schubert, Jeffrey Fagan, Alex R. Piquero, and Sandra H. Losoya. 2009. Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology*, 47: 699–740.

Melde, Chris and Finn-Aage Esbensen. 2011. Gang membership as a turning point in the life course. *Criminology*, 49: 513–552.

Melde, Chris and Finn-Aage Esbensen. 2013. The relative impact of gang status transitions: Identifying the mechanisms of change in delinquency. *Journal of Research in Crime and Delinquency*, 51: 349–376.

Mihalic, Sharon F., Diane Ballard, Amanda Michalski, Jonathan Tortorice, Linda Cunningham, and Susanne Argamaso. 2002. *Blueprints for Violence Prevention: Violence Initiative Final Process Evaluation Report*. Boulder: University of Colorado.

National Research Council. 2013. *Reforming Juvenile Justice: A Developmental Approach. Committee on Assessing Juvenile Justice Reform* (Richard J. Bonnie, Robert L. Johnson, Betty M. Chemers, and Julie A. Schuck, eds.; Committee on Law and Justice, Division of Behavioral and Social Sciences and Education). Washington, DC: The National Academies Press.

National Research Council. 2014. *Implementing Juvenile Justice Reform: The Federal Role*. Washington, DC: The National Academies Press.

Pyrooz, David C. 2014. From colors and guns to caps and gowns? The effects of gang membership on educational attainment. *Journal of Research in Crime and Delinquency*, 51: 56–87.

Pyrooz, David C., Jillian J. Turanovic, Scott H. Decker, and Jun Wu. 2015. Taking stock of the relationship between gang membership and offending: A meta-analysis. *Criminal Justice and Behavior*, 43: 365–397.

Raby, Carolotta and Fergal Jones. 2016. Identifying risks for male street gang affiliation: A systematic review and narrative synthesis. *Journal of Forensic Psychiatry and Psychology*, 27: 601–644.

Ryan, Joseph P., Laura S. Abrams, and Hui Huang. 2014. First-time violent juvenile offenders: Probation, placement, and recidivism. *Social Work Research*, 38: 7–18.

Thornberry, Terence P., Marvin D. Krohn, Alan J. Lizotte, and Deborah Chard-Wierschem. 1993. The role of juvenile gangs in facilitating delinquent behavior. *Journal of Research in Crime and Delinquency*, 30: 55–87.

Thornberry, Terence P., Marvin D. Krohn, Alan J. Lizotte, Carolyn A. Smith, and Kimberly Tobin. 2003. *Gangs and Delinquency in Developmental Perspective*. New York: Cambridge University Press.

Thrasher, Frederick M. 1927. *The Gang: A Study of 1,313 Gangs in Chicago*. Chicago, IL: The University of Chicago Press.

University of Chicago Crime Lab. 2017. *Gun Violence in Chicago, 2016*. Retrieved from urbanlabs.uchicago.edu/attachments/store/2435a5d4658e2ca19f4f225b810ce0dbdb92 31cbdb8d702e784087469ee3/UChicagoCrimeLab+Gun+Violence+in+Chicago+ 2016.pdf.

Waldron, Holly B. and Janet L. Brody. 2010. Functional Family Therapy for adolescent substance use disorders. In (John R. Weisz and Alan E. Kazdin, eds.), *Evidence-based Psychotherapies for Children and Adolescents*, 2nd Edition. New York: Guilford Press.

Waldron, Holly B., Natasha Slesnick, Janet L. Brody, Charles W. Turner, and Thomas R. Peterson. 2001. Treatment outcomes for adolescent substance abuse at 4- and 7-month assessments. *Journal of Consulting and Clinical Psychology*, 69: 802–813.

Waldron, Holly B. and Charles W. Turner. 2008. Evidence-based psychosocial treatments for adolescent substance abuse. *Journal of Clinical Child and Adolescent Psychology*, 37: 238–261.

Washington State Institute for Public Policy. 2017. *Benefit Cost Results*. Olympia: Washington State Institute for Public Policy. Retrieved from wsipp.wa.gov/ BenefitCost/Pdf/1/WSIPP_BenefitCost_Juvenile-Justice.

Wasserman, Gail S., Susan J. Ko, and Larkin S. McReynolds. 2004. *Assessing the Mental Health Status of Youth in Juvenile Justice Settings. Bulletin* (NCJ 202713). Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Weisburd, David L., Anthony Petrosino, and Gail Mason. 1993. Design sensitivity in criminal justice experiments. In (Michael Tonry, ed.), *Crime and Justice: A Review of Research.* vol. 17. Chicago, IL: University of Chicago Press.

Terence P. Thornberry, Ph.D., is the distinguished professor emeritus in the School of Criminal Justice at the University at Albany and a research professor in the Department of Criminology and Criminal Justice at the University of Maryland—College Park. He is the founding principal investigator of the Rochester Youth Development Study, a three-generation panel study begun in 1986 to examine the causes and consequences of delinquency and other forms of antisocial behaviors.

Brook Kearley, Ph.D., is a research assistant professor at The Institute for Innovation and Implementation, School of Social Work, University of Maryland—Baltimore. Her research interests include criminal and juvenile justice policy and program evaluation with a focus on substance use and delinquency prevention and intervention programs. With funding from the National Institute of Justice's dissertation fellowship program, she recently completed her dissertation titled *Long Term Effects of Drug Court Participation: Evidence from a 15-year Follow-up of a Randomized Controlled Trial.*

Denise C. Gottfredson, Ph.D., is a professor in the Department of Criminal Justice and Criminology at the University of Maryland—College Park. Gottfredson's research interests include delinquency prevention, and particularly the effects of school environments on youth behavior. Her recent publications include "Standards of Evidence for Efficacy, Effectiveness, and Scale up Research in Prevention Science: Next Generation" in *Prevention Science* and "Prevention Research in Schools: Past, Present, and Future" in *Criminology & Public Policy.*

Molly P. Slothower, M.A., is a Ph.D. candidate in the Department of Criminology and Criminal Justice, University of Maryland—College Park; a junior fellow at the Cambridge Centre for Evidence-Based Policing; and a research manager at the NYC Mayor's Office of Criminal Justice. Her research interests include behavior change interventions to reduce offender recidivism, restorative justice, and police management and discretion. Her recent publications include "Strengthening Police Professionalism with Decision Support: Bounded Discretion in Out-of-Court Disposals" in *Policing.*

Deanna N. Devlin, Ph.D., is a research associate for the CUNY Institute for State and Local Governance. Her research interests include juvenile delinquency, program evaluation,

and school crime prevention strategies. Her recent publications include "The Roles of Police Officers in Schools: Effects on the Recording and Reporting of Crime" in *Youth Violence and Juvenile Justice*.

Jamie J. Fader, Ph.D., is an associate professor in the Department of Criminal Justice at Temple University. Her research interests include youth justice, transitions to adulthood among vulnerable youth, urban crime, and qualitative research methods. She is the author of *Falling Back: Incarceration and Transitions to Adulthood among Urban Youth* (Rutgers University Press, 2013).