Justice Quarterly

Results from a Multi-Site Evaluation of the G.R.E.A.T. Program

Finn-Aage Esbensen, Dana Peterson, Terrance J. Taylor & D. Wayne Osgood

Available online: 15 Aug 2011

PLEASE SCROLL DOWN FOR ARTICLE

Results from a Multi-Site Evaluation of the G.R.E.A.T. Program

Finn-Aage Esbensen, Dana Peterson, Terrance J. Taylor and D. Wayne Osgood

Despite a long history of youth gang problems in the United States, there remains a paucity of evaluations identifying promising or effective gang prevention and intervention programs. One primary prevention program that has received limited support is Gang Resistance Education and Training (G.R.E.A.T.). An earlier national evaluation of the G.R.E.A.T. core middle school curriculum reported modest program effects but, importantly, found no programmatic effect on gang membership or delinquency. This manuscript presents results from a second national evaluation of the revised G.R.E.A.T. core curriculum that utilizes a randomized field trial in which classrooms were randomly assigned to treatment and control conditions. Approximately 4,000 students attending 31 schools in seven cities comprise the initial sample. Analyses of one-year post-treatment data indicate that students receiving the program had lower odds of gang membership compared to the control group. Additionally, the treatment group also reported more pro-social attitudes on a number of program-specific outcomes.

Keywords   youth gangs; gang prevention; program evaluation; risk factors

Finn-Aage Esbensen is the E. Desmond Lee Professor of Youth Crime and Violence and Chair of the Department of Criminology and Criminal Justice at the University of Missouri—St. Louis. He is co-author of *Youth Violence: Sex and Race Differences in Offending, Victimization, and Gang Membership* (with Dana Peterson, Terrance J. Taylor, and Adrienne Freng), Temple University Press. Dana Peterson, Ph.D., is an associate professor in the School of Criminal Justice at the University at Albany, State University of New York. She teaches and conducts research primarily in the areas of youth violence, youth gangs and gang prevention, and how sex/gender structure these behaviors and experiences. Terrance J. Taylor is an assistant professor in the Department of Criminology and Criminal Justice at the University of Missouri—St. Louis. His primary research interests involve juvenile victimization and offending, with an emphasis on the role of race/ethnicity. He received his Ph.D., in criminal justice from the University of Nebraska in 2002. D. Wayne Osgood is a professor of crime, law and justice and sociology at Pennsylvania State University. Professor Osgood does research on a broad range of topics concerning delinquency and other problem behaviors during adolescence and early adulthood. In addition, he has conducted research on a variety of programs for juvenile offenders (including prevention, diversion, and residential programs), and he has written about statistical issues for the analysis of deviant behaviors and of longitudinal data. Correspondence to: F. Esbensen Department of Criminology and Criminal Justice, University of Missouri—St. Louis, St. Louis, MO 63121-4499, USA. E-mail: esbensen@umsl.edu

Youth delinquent gangs received considerable academic and media attention during the 1990s. Much of this attention focused on the violence and drug dealing in which gang members are involved. To help combat this problem, a number of prevention, intervention, and suppression programs were developed (e.g. Decker, 2002; Klein & Maxson, 2006; Reed & Decker, 2002). Schools, one of the common grounds for American youth, have become a focal point for both general and specific prevention programing. In fact, Gottfredson and colleagues (2000) reported the average middle school offers 14 different and unique prevention programs that address violence, bullying, victimization, drug abuse, and other social problems, including gangs. Given the plethora of school-based prevention programs that have been designed to reduce an array of adolescent behaviors, school administrators face challenges in selecting a program that is optimal in light of the time and resource constraints of their facilities. Thus, it is imperative this choice be guided by a well-informed sense of program effectiveness. Several attempts in the past decade have sought to provide administrators with such knowledge. For example, the Blueprints Series (Mihalic, Fagan, Irwin, Ballard, & Elliott, 2002; Mihalic & Irwin, 2003) identified model violence prevention programs that have withstood rigorous scientific evaluations, and the Maryland Report (Sherman, Gottfredson, MacKenzie, Eck, Reuter, & Bushway, 1997) assessed the effectiveness of a broad range of projects. In 2005, the Helping America's Youth (HAY) Community Guide (Howell, 2009) rated programs identified by non-federal agencies on three levels: Level 1 (exemplary or model programs based on evaluation designs of the "highest quality"); Level 2 (effective programs based on quasi-experimental research); and Level 3 (promising programs). Similarly, the Office of Juvenile Justice and Delinquency Prevention provides a listing of effective or promising programs (OJJDP, 2010). One notable aspect of these reviews is the paucity of "model" or "effective" programs. This is not to say that most of the extant programs are ineffective; rather, the majority has not been evaluated in a manner that allows for assessment of their effectiveness (see, for instance, Dusenbury, Brannigan, Falco, & Hansen, 2003; Lillehoj, Griffin, & Spoth, 2004). In addition, some programs have experienced implementation failure that is then interpreted as program failure.

A second notable aspect of these reviews is that, in spite of the widespread concern with gangs and associated program development; there has been a paucity of research and evaluation of *gang-specific prevention* programs. One notable exception is the National Institute of Justice (NIJ)-funded evaluation of the Gang Resistance Education and Training (G.R.E.A.T.) program (Esbensen & Osgood, 1997, 1999; Esbensen, Osgood, Taylor, Peterson, & Freng, 2001). The G.R.E.A.T. program was developed in 1991 by law enforcement agencies in the greater Phoenix area (for a detailed accounting of the program history, consult Winfree, Peterson Lynskey, and Maupin, 1999) and experienced exponential growth calling for a national evaluation in 1994. That evaluation consisted of two separate studies: a cross-sectional design in which students receiving the G.R.E.A.T. program's core middle school curriculum were sur-

veyed one year after program delivery and that relied upon student self-report of program participation; and a five-year longitudinal study with matched control classrooms. The cross-sectional study identified favorable outcome results, including lower rates of gang membership among the treatment group (Esbensen and Osgood, 1997, Esbensen and Osgood, 1999) and held considerable promise for the program model. The findings from the more rigorous longitudinal design with matched classrooms and four-year follow-up were more ambiguous. No behavioral effects were found, but a lagged or sleeper effect was found for five mediating/proximal factors. That is, there were no differences between G.R.E.A.T. and comparison students in rates of gang membership or delinquency, but at three and four years post-program, G.R.E.A.T. students had lower risk-seeking tendencies, lower rates of victimization, more prosocial peers, more positive attitudes about police officers, and less positive attitudes about gangs (Esbensen et al., 2001).

Based in part on these modest findings, the G.R.E.A.T. program underwent a rigorous program review (see Esbensen, Freng, Taylor, Peterson, & Osgood, 2002; Esbensen, Peterson, Taylor, Freng, Osgood, Carson, and Matsuda, 2011, for a detailed account of the program review) that culminated in a redesign of the curriculum, expanding the core middle school component from 9 to 13 lessons, focusing more attention on skills building through interactive and cooperative learning strategies, and encouraging greater involvement of classroom teachers in program delivery. In addition, the revised curriculum took into account the extant research on risk factors for youth gang involvement, with lesson components targeting known risk factors or proximal influences for gang joining.

In this manuscript we report on the evaluation of the revised G.R.E.A.T. program, assessing the extent to which middle school students participating in this school-based gang prevention program express attitudes and engage in behaviors that are measurably different from those of a control group of students at one year post-program. Our findings contribute to the sparse body of knowledge about effective prevention strategies; the revised G.R.E.A.T. program is currently rated as "effective" by OJJDP and designated as "Level 2" in the Helping America's Youth rating scale.[1]

Review of Relevant Literature

Youth gangs and gang violence are community problems—that is, gangs and gang violence do not occur in a vacuum and must be considered within the larger contextual setting. When the G.R.E.A.T. program was initially developed in 1991, youth and gang violence were at "epidemic" proportions (Snyder &

1. In another article, we have reported on the high level of program fidelity associated with delivery of the G.R.E.A.T. program in classrooms participating in the current evaluation, allowing outcome evaluation results to be attributed with confidence to the program (Esbensen, Matsuda, Taylor, & Peterson, 2011).

Sickmund, 2006). Since then rates of youth and gang violence have decreased substantially, although the past few years have witnessed a new increase in gangs and gang membership (Egley, Howell, & Moore, 2010). A number of macro-level explanations have been offered for the decrease in violence and gang problems between 1995 and 2001, including a change in handgun availability, the crack market decline, an improved economy, and increased incarceration rates (Blumstein & Wallman, 2000). The increase in prevention and intervention programing during the 1990s may also have played a contributing role in this youth crime drop, addressing more proximal influences for gang involvement at the school, peer, family and/or individual level. Because virtually all American youths attend school, this setting has considerable potential for programs to prevent or intervene with gang joining and gang violence by attempting to ameliorate these negative proximal influences.

Risk Factors and Prevention Strategies

Research has identified a number of risk factors associated with gang affiliation and violent offending, and these risk factors can inform prevention programs. This growing body of research has categorized these influences within multiple domains, including community, school, peer, family, and individual. Representative of these risk factors are the following: community poverty and social disorganization, low commitment to school, poor school performance, association with few conventional or many delinquent peers, low parental monitoring, low attachment to parents, low involvement in conventional family activities, lack of empathy, impulsiveness, and moral disengagement (e.g. Battin, Hill, Abbott, Catalano, & Hawkins, 1998; Esbensen & Deschenes, 1998; Esbensen, Huizinga, & Weiher, 1993; Esbensen, Peterson, Taylor, & Freng, 2010; Hill, Howell, Hawkins, & Battin-Pearson, 1999; Klein & Maxson, 2006; Maxson & Whitlock, 2002; Maxson, Whitlock, & Klein, 1998; Pyrooz, Fox, & Decker, 2010; Thornberry, 1998; Thornberry, Krohn, Lizotte, Smith, & Tobin, 2003). Research has also demonstrated the deleterious cumulative effects of risk exposure: the greater the number of risk factors and/or the greater the number of risk domains experienced, the greater the odds of youth gang and violence involvement, with these increases in risk associated with exponential increases in odds (Esbensen et al., 2010; Thornberry et al., 2003). This collective body of risk factor research suggests that prevention programs should attempt to address risk factors in multiple domains and to do so earlier, rather than later, in adolescence, both before the factors accumulate and before the typical age of onset—i.e. age 14 for gang joining (Esbensen et al., 1993; Hill, Howell, Hawkins, & Battin-Pearson, 1999; Thornberry et al., 2003).

The developmental progression of behavior may also be important in planning prevention strategies. While many studies treat gangs as a phenomenon distinct from the general study of delinquency, there is considerable overlap between delinquency and gang involvement, as well as between risk factors

associated with delinquency, particularly violence, and gang membership (see, e.g. Esbensen et al., 2010). The works of Battin et al. (1998), Esbensen et al. (1993), Gatti, Tremblay, Vitaro, and McDuff (2005), and Thornberry, Krohn, Lizotte, and Chard-Wierschem (1993), Thornberry et al. (2003), for example, suggest that while the gang environment facilitates delinquency, many gang members are already delinquent prior to joining the gang (see also Melde & Esbensen, 2011). The rates of delinquent activity, however, increase dramatically during gang membership. This finding that delinquency generally precedes gang membership highlights the importance of universal gang prevention efforts during the early years of adolescence (i.e. programs that target all 11 to 12 year olds, that is, students in the 6th or 7th grade). Additionally, the link between risk factors associated with gang membership and delinquent behavior reinforces the relevance of two of the goals of the G.R.E.A.T. program: to reduce both delinquent (violent) activity and gang involvement. How does the G.R.E.A.T. program attempt to accomplish these goals?

The G.R.E.A.T. Program

The Gang Resistance Education and Training (G.R.E.A.T.) program is a school-based gang and violence prevention program with three primary goals: (1) teach youths to avoid gang membership; (2) prevent violence and criminal activity; and (3) assist youths to develop positive relationships with law enforcement. The original G.R.E.A.T. program[2] consisted of nine lessons and was modeled after the Drug Abuse Resistance Education (DARE) program. Developed by the Phoenix Police Department in 1991, G.R.E.A.T. was a cognitive-based program that taught students about crime and its effect on victims, cultural diversity, conflict resolution skills, meeting basic needs (without a gang), responsibility, and goal setting. Uniformed law enforcement officers taught the curriculum in schools, and teachers were requested to complement the program content during regular classes. The revised G.R.E.A.T. program contains much of the substance of the original program but, importantly, was also informed by the work of educators and prevention specialists and the growing body of risk factor research. As a result, the new G.R.E.A.T. program was expanded to 13 lessons; is still primarily taught by uniformed law enforcement officers (Federal agents from the US Marshalls and the Bureau of Alcohol, Tobacco, and Firearms as well as District Attorneys have also been trained and certified to teach G.R.E.A.T.); and incorporates classroom management training of officers and a focus on students' skill development through cooperative learning strategies: important pedagogical tools for educational settings (Gottfredson, 2001).[3]

2. The core program component of G.R.E.A.T. is its middle school curriculum, and this is often what is referred to with the term "G.R.E.A.T. program." Other optional components of G.R.E.A.T. are an elementary school curriculum, a summer program, and G.R.E.A.T. Families.
3. Information about the G.R.E.A.T. program and an overview of the G.R.E.A.T. lessons included in the middle school curriculum can be found at http://www.great-online.org/.

Two school-based programs guiding the revision of the G.R.E.A.T. program were the Seattle Social Development Model (SSDM) and Life Skills Training (LST). The SSDM is a comprehensive model that seeks to reduce delinquency and violence by building a positive learning environment incorporating several different classroom management components, including cooperative learning, proactive classroom management, and interactive teaching (Catalano, Arthur, Hawkins, Berglund, & Olson, 1998). The LST program is a three-year intervention in which two annual booster sessions supplement the initial program (Dusenbury & Botvin, 1992). LST consists of three components: (1) self-management skills; (2) social skills; and (3) information and skills that are directly related to the problem of drug abuse. The revised G.R.E.A.T. program has adopted some of the strategies from LST (in fact, some of the LST curriculum writers participated in the rewriting of the G.R.E.A.T. program), including an emphasis on the development of skills, rather than on the assimilation of knowledge, and has also incorporated problem-solving exercises and cooperative learning strategies. With this revised program fully implemented by 2003, there was renewed interest in the question of program effectiveness. In July 2006, the National Institute of Justice selected the University of Missouri-St. Louis to conduct a process and outcome evaluation of the revised G.R.E.A.T. program. This manuscript focuses upon sustained program effects one-year post-treatment (consistent with the Blueprints standard), while results from the process evaluation, which indicated a strong degree of implementation fidelity, are reported in Esbensen, Matsuda, Taylor, and Peterson (2011).

Methods

Site and School Selection

Site selection was driven by the presence of the G.R.E.A.T. program and willingness of the police departments and school districts to agree to the evaluation design. In addition, three main criteria guided site selection: (1) existence of an established G.R.E.A.T. program,[4] (2) geographic and demographic diver-

---

4. Length of time the locale had operated the program and the extent to which schools had been exposed to the program were assessed prior to site selection for the national evaluation. Sites where the program was just beginning were excluded because they were deemed likely to have had less time to "work out the kinks" associated with delivering the program with fidelity. Conversely, some sites with a long history of delivering the program were excluded from consideration because it was deemed likely that the program had saturated the entire school and/or community context. In the selected cities, G.R.E.A.T. had not been taught in all district schools which allowed us in some instances to include schools with little or no prior exposure to G.R.E.A.T. while at the same time having experienced officers teaching the program. The possibility for a contamination effect, however, is possible in some schools in which G.R.E.A.T. had been offered for several years.

sity, and (3) evidence of gang activity. The first step in the process was to secure a listing of potential program sites based upon the existence of the G.R.E.A.T. program. The research staff contacted the G.R.E.A.T. Regional Administrators[5] and Bureau of Justice Assistance[6] personnel to identify locales with institutionalized programs. Consideration was given to factors such as the length of time the program had been in operation, number of G.R.E.A.T.-trained officers, number of schools in which the program was offered, and the components of the G.R.E.A.T. program implemented. Also of interest were police department characteristics that could affect program delivery, including department size and organizational structure. Some G.R.E.A.T. programs, for instance, utilize School Resource Officers (SRO) to teach the program while others use the "Portland" model in which "street cops" teach the program on an overtime basis in schools on their beat. Once this list of potential agencies was constructed, the research staff contacted representatives in these cities to obtain more information about the delivery of the G.R.E.A.T. program (e.g. school district size, length of program history at a site, and degree of program implementation). Additional site characteristics (i.e. race and ethnic composition, and population size) were also taken into account at this time. A last criterion considered was the volume of youth crime (based on police reports) and gang activity (information was obtained from the National Gang Center) in each site. Ultimately, a list of seven cities varying in size, region, and level of gang activity were identified (Albuquerque, NM; Chicago, IL; a Dallas-Fort Worth area district; Greeley, CO; Nashville, TN; Philadelphia, PA; and Portland, OR).

Upon selection of the cities, the research staff worked with the primary local law enforcement agency and the school district in each city to secure their cooperation. Upon district approval, between four and six schools in each site were identified for study participation; the goal of the school selection was to identify schools that, taken as a whole, would be representative of the districts. Principals in these targeted schools were contacted to elicit their support and cooperation with the evaluation design. In two instances, the principals declined to participate.[7] These schools were then replaced with a

5. G.R.E.A.T. is a national program overseen by the G.R.E.A.T. National Policy Board (NPB). For administrative purposes, responsibilities for program oversight are held by (or "given to") agencies operating in different geographic regions: Midwest Atlantic, Southeast, Southwest, and West. Additionally, two federal partners—the Bureau of Alcohol, Tobacco, Firearms, and Explosives (BATF) and the Federal Law Enforcement Training Center (FLETC)—are involved in program training and oversight.

6. The Bureau of Justice Assistance (BJA) oversees the allocation of federal funds and grant compliance associated with the G.R.E.A.T. program.

7. Principals declined their schools' participation for different reasons. One principal indicated that he had previously been a police gang investigator, and, therefore, knew the program worked; the second principal would not agree to random assignment and withholding some students from the program.

comparable school in the district.[8] This process produced a final sample of 31 schools and 195 classrooms (102 received G.R.E.A.T. and 93 did not receive the program), and 4,905 students listed on the classroom rosters.

Following the principal's agreement to participate in the evaluation, more detailed discussions/meetings were scheduled with school administrators and grade-level teachers, G.R.E.A.T. officers, and the research team. Whenever possible, face-to-face meetings were held, but in some instances final arrangements were made via telephone. School and police personnel were informed of the purpose of the evaluation, issues related to the random assignment of classrooms to the treatment condition (i.e. receive G.R.E.A.T. or not receive G.R.E.A.T.), procedures for obtaining active parental consent for students in these classrooms to participate in the evaluation, scheduling the G.R.E.A.T. program delivery, and other logistical issues associated with the study design.

School configuration varied somewhat, with 20 schools having the traditional middle school organization of grades six through eight, five schools having grades five through eight, and six schools organized as kindergarten through eighth grade. For the evaluation, classes in the G.R.E.A.T. grade level were selected, and this varied slightly; while most officers taught the program to sixth-graders, some taught at the seventh-grade level. Thus, sixth grade students were included from 26 schools, and seventh grade students comprised the sample in the remaining five schools.

Active Parental Consent

Due to the nature of the evaluation, active parental consent was required for student participation. We utilized a strategy that had proven successful in prior studies (Ellickson & Hawes, 1989; Esbensen et al., 1996; McMorris et al., 2004; Unger et al., 2004). Specifically, teachers were recruited and compensated for their assistance collecting the consent forms from their students. Regardless of whether permission was granted or denied by the parent, teachers received $2.00 for each returned form. Additionally, for each classroom, there was an incentive for teachers based upon classroom-level return rates: the teachers would receive a $10 bonus if 70% or more of their students returned consent forms, $20 if the class reached 80% or more, and $30 if 90% or more of the students in the classroom returned a form. In three cities, the school districts would not allow direct compensation to teachers, but we were

8. One of the five originally-selected schools in Chicago (comprised of nearly 100 % African-American students) agreed to participate in the evaluation but was unable to meet the requirements of the study and was dropped from the sample. Given time constraints (i.e. too late in the school year to select a comparable school and implement the program with fidelity), we were unable to replace the excluded school during 2006–2007. Thus, the resulting sample was disproportionately Hispanic and not representative of the district. To increase sample representativeness, we added two primarily African American schools to the evaluation in the 2007–2008 school year, even though this meant that these schools would be one year behind other schools in the evaluation.

allowed to provide compensation to the school or district in the teachers' honor. In addition to compensating teachers, students were also provided with an incentive for returning the form—a small portable FM radio with head-phones (cost of approximately $3.00 wholesale).

Letters to parents and active consent forms were distributed to students and their return recorded on class rosters. This documentation allowed for fol-low-up forms to be sent home with students who failed to return the initial form. In addition to these incentives, teachers were contacted on a regular basis, in most instances daily, to monitor return rates. In most schools, this consent process was completed in less than two weeks, and in several instances, in just three days. (For more detailed description of the active con-sent process, consult Esbensen, Melde, Taylor, & Peterson, 2008).

This strategy of compensating teachers and students, while costly, is to be recommended because it rewards teachers and students for their assis-tance and allows the active consent process to be completed in a relatively short timeframe. Overall, 89.1% of youths (N = 4,372) returned a completed consent form, with 77.9% of parents/guardians (N = 3,820) allowing their child's participation. It should be noted that while Esbensen et al., 2008, reported a 79% consent rate, the addition of two schools to the evaluation after the publication of that article resulted in the 78% overall consent rate reported here. The direct cost of the teacher incentives was $12,894 and the cost of the 4,750 radios was $14,250 for a total of $27,144. This trans-lates into a cost of approximately $3,878 per city, $936 per school, $146 per classroom, and $7.39 per active consent participant. To summarize the results of the site selection and active parental consent process, this study includes an active consent sample of 3,820 students (77.9% of the 4,905 stu-dents listed on classroom rosters at the beginning of the study period) rep-resenting 195 classrooms in 31 schools in seven cities across the continental United States.

Research Design and Random Assignment of Classrooms

The outcome evaluation employs an experimental longitudinal panel design (a randomized control trial with long-term follow-up) in which classrooms in each of the participating schools were randomly assigned to the treatment (i.e. G.R.E.A.T.) or control condition. The G.R.E.A.T. program was taught in sixth grade in 26 of the 31 schools and in seventh grade in the remaining five schools. Once it was determined in which core subject area (commonly social Studies but also in English and Science classes) the program was to be taught, we enumerated all of the grade-level classes (ranging from 3 to 12). In situa-tions with an odd number of classes, we made the a-priori decision to over-sample treatment classes (in partial recognition of the fact that many of the principals were reluctant to "deprive" any of their students of the program). The list of classes was then numbered from one through highest and a table of

random numbers was consulted to select the classrooms in which G.R.E.A.T. would be taught. Unselected classrooms comprised the control group.

All students in the treatment and control classrooms were eligible to participate in the evaluation. All students for whom active parental consent was obtained (3,820) were then asked to participate in the evaluation by completing a confidential group-administered pre-test questionnaire. Upon completion of the G.R.E.A.T. program in each school, students were then requested to complete post-tests and four annual follow-up surveys. Retention rates across the three waves of data included in these outcome analyses were excellent: 98.3% completed the pre-test, 94.6% completed the post-test, and 87.3% completed the one-year post-program survey. These response rates reflect the diligent efforts of the research assistants working on this project. It is particularly challenging to track students through multiple schools and school districts, especially in a highly mobile sample: while initially enrolled in 31 middle schools at pre-test, students were surveyed in 121 different schools in Wave 3 (although we identified students enrolled in a total of 180 different schools, most of the schools in which students were not surveyed were outside the original seven districts). We obtained permission from principals at the new schools to survey the transfer students—clearly, a time and labor-intensive effort, but one well worth achieving these high response rates.

Student Sample Characteristics

The sample is evenly split between males and females; most (55%) youths reside with both biological parents; and the majority (88%) was born in the United States (see Table 1). The sample is racially/ethnically diverse, with Hispanic youths (37%), White youths (27%), and African-American (17%) youths accounting for 81% of the sample. Approximately two-thirds of the youths (61%) were aged 11 or younger at the pre-test, representing the fact that 26 of the 31 schools delivered the G.R.E.A.T. program in 6th grade. Three of the six Chicago schools and two of four schools in Albuquerque taught G.R.E.A.T. in 7th grade; thus, students in these sites were somewhat older than students in the other sites.

Measurement

Outcome Measures

To assess program effectiveness, it was essential that measures of the three program goals be included in the student surveys. Additionally, the G.R.E.A.T. lessons introduced a number of secondary (proximal) outcomes that sought to reduce known risk factors for delinquency and gang joining. We developed a student questionnaire that captured the essence of this skills building program;

Table 1 Sample Characteristics at Wave 1

	Full sample N = 3,820 (%)	ABQ N = 591 (%)	CHI N = 500 (%)	DFW area N = 614 (%)	GRE N = 582 (%)	NSH N = 590 (%)	PHL N = 457 (%)	POR N = 486 (%)
Sex								
Male	50	50	50	54	52	55	43	42
Female	50	50	50	46	48	46	57	58
Race/ethnicity								
White	27	16	7	20	34	45	12	51
African-American	18	4	29	21	2	23	44	7
Hispanic/Latino	37	49	56	46	50	17	20	13
American Indian	4	10	1	2	5	1	4	4
Asian	4	2	1	6	1	6	4	9
Multi-racial	8	14	2	5	4	4	12	13
Other	4	5	2	1	5	5	5	3
Age								
11 or younger	61	35	18	74	77	80	61	79
12	29	43	44	25	22	19	35	20
13 or older	10	23	38	2	2	<1	4	1
Mean	11.48	11.87	12.22	11.27	11.23	11.19	11.42	12.22

(Continued on next page)

Table 1 (Continued)

	Full sample N = 3,820 (%)	ABQ N = 591 (%)	CHI N = 500 (%)	DFW area N = 614 (%)	GRE N = 582 (%)	NSH N = 590 (%)	PHL N = 457 (%)	POR N = 486 (%)
Living arrangement								
Both biological parents	55	52	57	60	58	60	38	58
Single parent	20	20	19	15	14	18	24	15
1 Biological/1 step-parent	13	15	12	14	15	12	18	13
1 Biological/1 other adult	7	7	7	7	7	7	11	8
Other relatives	3	6	3	3	4	2	8	5
Other living arrangement	2	1	1	1	3	2	2	1
Immigration status								
Born outside US	12	10	9	13	11	15	11	15
Born in US	88	90	91	87	89	85	89	85

that is, identifying the mediating variables that could explain the mechanisms through which behavioral outcomes could be achieved. If the program is determined to reduce rates of gang membership and youth violence, it is important to understand how these goals are achieved. To reiterate, the G.R.E.A.T. program has three primary goals: (a) to help youths avoid gang membership, (b) to reduce violence and criminal activity, and (c) to help youths develop a positive relationship with law enforcement. In the current analyses, gang membership is measured by a single-item question that is part of a larger set of questions about youth gangs. Specifically, students were asked to answer the following question: "Are you now in a gang?" This self-nomination approach has been found to be a valid and robust measure of gang affiliation (e.g. Esbensen, Winfree, He, & Taylor, 2001; Thornberry et al., 2003). To measure delinquency and violent offending, students completed a 15-item self-reported delinquency inventory, including response categories that allowed for assessment of both ever and current prevalence as well as frequency of offending during the past six months. We treated this self-report inventory as a composite measure of general delinquency (examined both a variety and frequency score) but also created a separate measure of violent offending consisting of three items (attacked someone with a weapon, used a weapon or force to get money or things from people, been involved in gang fights). To measure the third specific program goal (improving relations with law enforcement), students were asked to respond to six questions tapping global attitudes to the police as well as two additional questions measuring students' attitudes about police officers as teachers.

In addition to these preceding three program goals, the 13 G.R.E.A.T. lessons are intended to teach youths the life-skills thought necessary to prevent involvement in gangs and delinquency (see, e.g. Hill et al., 1999; Klein & Maxson, 2006; Maxson & Whitlock, 2002; Maxson et al., 1998; Thornberry et al., 2003) by reducing the effect of a range of risk factors. These mediating or proximal variables are treated as implied program objectives and are included in our outcome analyses. We therefore examined the extent to which students exposed to G.R.E.A.T. had improved or enhanced skills that would enable them to better resist the lures of gang membership and resist peer pressure to engage in illegal activities. The G.R.E.A.T. lessons encourage students to make healthy choices such as being involved in more pro-social activities and associating more with pro-social peers and less with delinquent peers. The lessons also teach students to improve their communication skills by being active listeners and being better able to interpret verbal and non-verbal communication. The program targets these skills in order to improve students' empathy for others. Risk factors associated with youth violence and joining gangs are also addressed in the curriculum. The program, for example, seeks to increase the levels of guilt associated with norm violation and to reduce the neutralization of illegal acts (i.e. moral disengagement). For a full listing of scales and scale characteristics, see the Appendix.

Analysis Strategy

Our highly nested research design requires a multilevel analysis, which we implemented with the MLwiN software (Rasbash, Steele, Brown, and Goldstein, 2009). The design includes two waves (Waves 2 and 3) of outcome observations (level 1) for 3,702 individual students[9] (level 2), who are nested within 195 classrooms in which the program was or was not delivered (level 3), which are, in turn, nested within 31 schools (level 4) located in 7 cities (level 5). Given the small number of cities, we treated this level as a fixed effect through a set of dummy variables. The model included random effects for the remaining four levels. To insure that school differences were not confounded with the program effect, the treatment versus control contrast was centered within schools. The analysis controlled for the pre-test measure of the outcome and for the difference between Waves 2 and 3 (coded −.5 for Wave 2 and +.5 for Wave 3). The treatment effect was allowed to vary randomly across schools in order to insure a conservative test. A logistic model was applied to the dichotomous measure of gang membership and a negative binomial model was used for the highly skewed measures of self-reported general delinquency and violent offending. All other models were linear. For the linear models we express the magnitude of the program effect in terms of standard deviation units of difference between treatment and control (i.e. Cohen's *d*), with positive values reflecting beneficial impacts. For the logistic and negative binomial models, the value is the percentage difference between treatment and control.

Results

To assess program effectiveness, we compare responses from students in G.R.E.A.T. classes to students in control classrooms using the post-test and one-year follow-up questionnaires. Results presented here represent the average treatment effects over Waves 2 and 3.[10] However, prior to examination of

---

9. The analysis file includes data for 3,246 students with data for both Waves 2 and 3, another 368 for Wave 2 but not Wave 3, and 88 for Wave 3 but not Wave 2, for a total of 3,702 students with either or both. The 3,702 students represent an upper bound for the analyses because it counts youth with any data and does not take into account variable-specific missing data on any given outcome or cases lost when we control for Wave 1 (from being missing on the same variable). The analysis-specific counts of cases are for person/waves rather than people (as specified in MLwiN). In the basic model (without Wave 1 control) we lose cases only due to being missing on the outcome because the only other variables involved we have for everybody (wave, site, and treatment/control). With respect to missing data, the total dataset has 6,948 person/wave cases; the number included in the analyses with & without Wave 1 control varies from 6,611 and 6,180 (attitudes toward gangs) to 6,905 and 6,751 (school disorganization).

10. Analyses were also conducted separately by wave, to assess treatment effects at post-test and treatment effects at the one-year follow-up. For all but 5 measures, there was a significant treatment effect at both time points. For the five that differed, the difference in effect between Wave 2 and Wave 3 was not statistically significant, and there was a statistically-significant average treatment effect across the time periods.

outcomes, we examined the success of the random assignment of classrooms to produce comparable groups of treatment and control. We conclude that the random assignment process was moderately successful; there were three significant differences ($p < .05$) between the two groups (awareness of services, attitudes about gangs, and frequency of delinquency), with the treatment group being more pro-social at the pre-test. Five additional differences were noted at $p < .1$ (violent offending frequency, gang membership, pro-social peers, negative peer commitment, and delinquent peers). These pre-existing differences between the groups do not permit us to make strong claims of comparability because there may be a little more difference than one would expect by chance alone, and the differences that do arise tend to favor the treatment group. But overall the differences are quite small, and the biggest difference is well within the bounds of chance. Furthermore, controlling for pre-test measures, as we do in all of the analyses, has negligible impact on the size or significance of the group differences on outcomes.

With respect to the primary goal of reducing gang membership, it will be helpful to identify the number of gang-involved youth at Wave 2 and Wave 3 by treatment condition. At Wave 2, 177 youth answered yes to the question: "are you now in a gang?" Of these gang-involved youth, 105 were in the control group and 72 were G.R.E.A.T. students. At Wave 3, there were 172 gang members, 101 in the control group and 71 in the treatment group. This pattern of more gang members in the control group was found in all seven cities.

Program Goals

Our first concern is to determine if the three stated program goals (i.e. reduction in gang membership, reduction in violent offending, and improved attitudes towards the police) were achieved. The analyses reveal that there were statistically significant differences between the treatment and control groups on two of the three outcomes. First, with regard to gang membership, we note in Table 2 that the odds of gang membership were 39% lower[11] for students completing the G.R.E.A.T. program relative to the control sample.[12] Second, G.R.E.A.T. students reported more positive opinions of police officers than did the study participants in the control group (effect size (ES) of .076). This positive assessment of law enforcement was even more pronounced for the two-item scale measuring attitudes more specific to the G.R.E.A.T. program (G.R.E.A.T. ATP, ES = .204). While results were in the expected direction of a positive program effect (10% reduction in the frequency of offending, although 1% increase in the variety of offending), the third program goal of reducing

11. In an unpublished report submitted to NIJ and in Esbensen et al. (2011), we reported a 54% reduction in the odds of gang joining. The difference reported here is due to a change in the MLwiN program that now allowed the model to run with all variance terms included in the analysis.
12. In response to one reviewer's concerns, the Wave 2 specific program effect was a 38.7% reduction in the odds of gang membership and 40.6% for Wave 3.

Table 2 Program effect estimates for attitudinal and behavioral measures controlling for between city differences and overall change over time

	Program effect	B	SE	T
Attitudinal measures				
Impulsivity	0.015	−0.012	0.024	−0.513
Risk-seeking	0.041	−0.041	0.030	−1.360
Anger	0.057	−0.056	0.026	−2.123*
Self-centeredness	0.054	−0.046	0.022	−2.060*
Attitudes toward the police (ATP)	0.076	0.070	0.024	2.908*
GREAT ATP	0.204	0.190	0.033	5.720*
Prosocial peers	0.051	0.050	0.030	1.685*
Peer pressure	0.079	−0.050	0.020	−2.465*
Negative peer commitment	0.050	−0.047	0.029	−1.617
Positive peer commitment	−0.010	−0.011	0.037	−0.298
Delinquent peers	0.083	−0.051	0.021	−2.474*
Lying neutralizations	0.066	−0.066	0.034	−1.951**
Stealing neutralizations	0.018	−0.016	0.030	−0.543
Hitting neutralizations	0.105	−0.122	0.032	−3.800*
School commitment	0.020	0.015	0.021	0.733
Guilt	0.028	0.016	0.016	1.005
Conflict resolution	−0.018	−0.008	0.013	−0.646
Calming others	−0.004	−0.002	0.014	−0.135
Refusal skills	0.090	0.043	0.013	3.229*
Prosocial Involvement Index	0.047	0.056	0.030	1.856**
Empathy	−0.008	−0.005	0.022	−0.243
Active listening	0.028	0.019	0.020	0.940
Problem solving	0.027	0.025	0.024	1.048
Self-efficacy	−0.004	−0.003	0.024	−0.115
Awareness of services	0.015	0.012	0.021	0.539
Collective efficacy	0.125	0.075	0.021	3.554*
Attitudes about gangs	0.114	0.102	0.031	3.313*
Altruism	0.051	0.031	0.019	1.612
Behavioral[a]				
Delinquency (frequency)[b]	7.0%	−0.073	0.072	−1.019
Delinquency (variety)[b]	7.0%	−0.072	0.048	−1.495
Violent offending (frequency)[b]	10.0%	−0.107	0.179	−0.597
Violent offending (variety)[b]	−1.0%	0.007	0.108	0.060
Gang[c]	39.2%	−0.498	0.162	−3.069*

*Significant at $p < 0.05$. **significant at $p < 0.10$.

[a]Program effect as percent reduction.

[b]Negative binomial model.

[c]Logistic regression model.

violent offending, was not met. There were no differences between the two groups with respect to violent offending, or general delinquency for that matter.

Proximal Outcomes

With regard to more proximal outcome measures, a number of statistically significant differences was observed. These differences were all in the direction of a positive program effect. As discussed above, the G.R.E.A.T. program is intended to be a skills building curriculum that provides students with, for example, the ability to better resist peer pressure, to control their anger, and to view joining gangs as an unattractive choice. Our outcome analyses included 26 proximal outcome measures (in addition to the five program outcomes discussed above) that tapped the extent to which the students enrolled in the G.R.E.A.T. program developed skills and attitudes that were promoted throughout the G.R.E.A.T. curriculum. Of these 26 measures, G.R.E.A.T. students had significantly ($p < .05$) more positive responses to eight of these outcomes than did control students and marginal significance on another three (lie neutralization, pro-social activities, and pro-social peers). For instance, the G.R.E.A.T. students made better use of refusal skills (ES = .090), were better able to resist peer pressure (ES = .079), reported being less self-centered (ES = .054) and expressed less positive attitudes towards gangs (ES = .114). There were *no* statistically significant differences between the groups on 15 of the attitudinal measures: empathy, impulsivity, risk-seeking, negative peer commitment, positive peer commitment, neutralization for theft, school commitment, guilt, conflict resolution, calming others, active listening, problem solving, self-efficacy, awareness of services, and altruism.

Though program effects were somewhat larger at Wave 2 than at Wave 3, the difference was not substantial. For the 13 measures with program effects significant at $p < .10$ in Table 2, the mean effect size was .11 at Wave 2 and .07 at Wave 3. The program impact estimates reached significance with $p < .05$ for nine variables at Wave 2 versus 6 variables at Wave 3, and significance with $p < .10$ for 10 and 9 variables at the 2 waves. Though the program impact significantly declined over time for 2 of the measures (with $p < .05$, GREAT ATP and collective efficacy), the program impact remained significant at $p < .05$ for both waves in each case.

Summary of Program Outcomes

In sum, we examined a total of 33 outcome measures: five behavioral outcomes (variety and frequency of violent offending, variety and frequency of delinquency, and gang membership) and 28 attitudinal or perceptual outcomes. Of the 33 outcome measures included in the analyses, one behavioral

(gang membership) and 10 attitudinal/perceptual differences were found at the .05 significance level between the G.R.E.A.T. and non-G.R.E.A.T. students; an additional three attitudinal differences were marginally significant ($p < .10$). Specifically, the G.R.E.A.T. students compared to non-G.R.E.A.T. students reported (see Table 2):

- more positive attitudes to police (ES = .076);
- more positive attitudes about police in classrooms (ES = .204);
- less positive attitudes about gangs (ES = .114);
- more use of refusal skills (ES = .090);
- more resistance to peer pressure (ES = .079);
- higher collective efficacy (ES = .125);
- less use of hitting neutralizations (ES = .105);
- fewer associations with delinquent peers (ES = .083);
- less self-centeredness (ES = .054);
- less anger (ES = .057);
- lower rates of gang membership (39% reduction in odds);
- less use of lie neutralization (ES = .066; $p < .10$);
- more pro-social peers (ES = .051; $p < .10$);
- more pro-social involvement (ES = .047; $p < .10$).

In addition to knowing the overall magnitude of the program effects, it would also be useful to have information about how much that effect varies across schools and cities. In our multilevel analysis, the variance component for the treatment effect estimates this variation. For none of the significant program outcomes was this variation in program effect statistically significant, and for 6 of the 14 the maximum likelihood estimate of the variance was zero. This should not be taken as strong evidence of consistency, however, because this is not a very powerful test. Indeed, when the variance estimates were not zero, they typically corresponded to standard deviations about the size of the significant program effects. In that scenario, program impact would be negligible to slightly harmful in about 20% of schools.

Discussion

Schools have become a common setting in which delinquency prevention programs are delivered (Gottfredson, 2001). There is no shortage of available programs from which to choose, and schools—especially middle schools—often have multiple programs operating during the school year (Gottfredson, 2001). Given resource limitations, however, school administrators need to weigh the "costs and benefits" of each program when making their decisions. Research evolving from the movement toward "evidence-based practices" (e.g. Sherman et al., 1997) has provided a wealth of information regarding the implementation and effectiveness of specific prevention programs, although the evidence base on gang prevention programs is still insufficient.

During the past 20 years, there has been a commensurate increase in the inclusion of police officers on school campuses, as both School Resource Officers (e.g. Finn & McDevitt, 2005; Gottfredson & Na, 2010) and prevention program providers (e.g. DARE and G.R.E.A.T.). In this manuscript we have addressed the efficacy of one such program that utilizes law enforcement officers to deliver a gang prevention and violence reduction program. A third objective of this program is related to the program provider: that is, improving police—youth relationships.

The current manuscript highlights the key sustained outcome findings (average program effects for post-test and one-year follow-up) from the Process & Outcome Evaluation of G.R.E.A.T. Results from analyses of three waves of survey data collected from students in seven US public school districts indicate that the program is meeting its primary objective of preventing gang membership; the analyses indicate a 39% reduction in the odds of gang joining one year post-program. In spite of the research showing a number of shared risk factors between delinquency and gang membership (and few or no factors unique to gang membership), we did not find a significant program effect on rates of violent offending.[13] The third goal of the G.R.E.A.T. program, to improve youths' attitudes towards the police (ATP), was met, with an effect size of .11 for the global measure of ATP and an effect size of .20 for the more specific measure of ATP related to G.R.E.A.T.

These findings suggest that a relatively short-term (13 lessons) primary prevention program can have measurable effects on a diverse sample of students. The evaluation was conducted in seven cities representing a cross-section of the United States. The process evaluation indicated that the program was implemented with fidelity (Esbensen, Matsuda, et al., 2011), providing confidence that the outcomes can be attributed to the G.R.E.A.T. program. Active parental consent rates for the students' participation in the outcome evaluation were quite high, thereby reducing the potential bias of selective loss. The high retention rates from the Wave 1 to Wave 3 surveys also add confidence to the robustness of the outcome results.

In addition to examining direct effects of G.R.E.A.T. on the three main program goals, we explored a range of mediating or proximal factors. Our results identify positive program effects on many of these program objectives. Compared with students in the control classrooms, students in G.R.E.A.T. classrooms illustrated less susceptibility to peer pressure, better refusal skills, and less involvement with delinquent peers; lower support for neutralizations regarding violence; less favorable attitudes about gangs; lower levels of self-centeredness and anger; and a higher degree of collective efficacy. Thinking about these findings from a logical perspective, the results are quite promising: G.R.E.A.T. appears to reduce key underlying risk factors for gang

13. We acknowledge that given the findings reported for the other two program goals and proximate program goals, it is surprising that there was no reduction in offending associated with the program. This is especially so, given the overlap in risk factors associated with gang membership and offending.

membership and violent offending (e.g. self-centeredness, anger); reduce the situational contexts where delinquency and gang membership is most likely to flourish (i.e. associations with delinquent peers); and provide youth with the skills necessary to recognize and resist temptations of peer pressure (e.g. peer pressure susceptibility and use of refusal skills), including a greater belief that offending is universally "wrong" (i.e. fewer neutralizations).

It is important to place these findings in context. The one-year post-program results from the longitudinal component of the national evaluation of the original G.R.E.A.T. program (Esbensen et al., 2001) indicated no program effect. And, it was only three and four years post-treatment that a sleeper or lagged effect was found for five outcomes: more favorable attitudes to police, lower victimization, more negative attitudes about gangs, more pro-social peers, and less risk-seeking behavior. We can speculate that the revised curriculum with its emphasis on skills building and use of cooperative learning strategies (and other pedagogically sound practices) was more successful in achieving favorable outcomes than was the earlier program with its emphasis on cognitive elements that were delivered in a more "canned" and didactic delivery mode. Three of the five significant outcomes noted in G.R.E.A.T. 1, were replicated in the current evaluation (attitudes to police, negative attitudes to gangs, and pro-social peers). The fact that both evaluations produced more favorable attitudes toward the police among the G.R.E.A.T. students suggests that this kind of law enforcement-based prevention program can have a positive impact on youth—police relations. It is also interesting to note that both studies produced evidence that the G.R.E.A.T. program is associated with more negative views of gangs and greater association with pro-social peers. While there were no differences between the treatment and control students with regard to risk-seeking in the current study, two other elements of self-control theory (anger and self-centeredness) were significant. In the current evaluation we did not examine victimization as a potential outcome since it was not a stated program goal nor was it addressed in the lessons. We view these similarities in findings as suggestive of an overall consistency in the program but further speculate that the additional program effects of the revised G.R.E.A.T. program are likely an artifact of the revised and enhanced curriculum. Only time will tell if the delayed or sleeper effects reported in the earlier evaluation of G.R.E.A.T. will be replicated in the current evaluation.

Clearly, this program is no "silver bullet" but these findings suggest that G.R.E.A.T. can be effectively included as a primary prevention component of a larger community-wide effort to reduce gang membership and youth violence. It is important to note that the effect sizes were modest (ranging from .05 to .20) and that no differences were found between students in G.R.E.A.T. and non-G.R.E.A.T. classrooms for a number of important mediating factors. However, the fact that statistically significant differences were found for 11 outcome measures (and another three with marginal significance) should be considered very promising, especially in light of the fact that these effects were produced after just 13 class periods (approximately

40 minutes in length). We would also like to point out that in some of the study schools, there exists a small possibility of a contamination effect suggesting that the results presented here should be considered conservative estimates. The G.R.E.A.T. program, as discussed in the site selection section, had operated for multiple years in each of the participating school districts and in many of the selected schools. While we excluded from consideration sites in which there was a strong likelihood of contamination, it is still possible that in some schools, the presence of G.R.E.A.T. for several years may well underestimate program effectiveness.

Acknowledgments

References

Battin, S. R., Hill, K. J., Abbott, R. D., Catalano, R. F., & Hawkins, J. D. (1998). The contribution of gang membership to delinquency beyond delinquent friends. *Criminology, 36*(1), 93-115.

Blumstein, A., & Wallman, J. (2000). *The crime drop in America*. New York, NY: Cambridge University Press.

Catalano, R. F., Arthur, M. W., Hawkins, J. D., Berglund, L., & Olson, J. J. (1998). Comprehensive community and school based interventions to prevent antisocial behavior. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful inventions* (pp. 248-283). Thousand Oaks, CA: Sage.

Decker, S. H. (2002). *Policing gangs and youth violence*. Belmont, CA: Wadsworth/ Thomson Learning.

Dusenbury, L., & Botvin, G. J. (1992). Competence enhancement and the development of positive life options. *Journal of Addictive Diseases, 11*(3), 29-45.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237-256.

Egley, A. Jr., Howell, J. C., & Moore, J. P. (2010). *Highlights of the 2008 National Youth Gang Survey*. Washington, DC: Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Ellickson, P. L., & Hawes, J. (1989). An assessment of active versus passive methods of obtaining parental consent. *Evaluation Review, 13*, 45-55.

Esbensen, F.-A., & Deschenes, E. P. (1998). A multisite examination of youth gang membership: Does gender matter? *Criminology, 36*, 799-828.

Esbensen, F.-A., Deschenes, E. P., Vogel, R. E., West, J., Arboit, K., & Harris, L. (1996). Active parental consent in school-based research: An examination of ethical and methodological issues. *Evaluation Review, 20*, 737-753.

Esbensen, F.-A., Freng, A., Taylor, T. J., Peterson, D., & Osgood, D. W. (2002). Putting research into practice. The national evaluation of the Gang Resistance Education and Training (G.R.E.A.T.) program. In W. L. Reed & S. H. Decker (Eds.), *Responding to gangs: Evaluation and research* (pp. 139-167). Washington, DC: Department of Justice: National Institute of Justice.

Esbensen, F.-A., Huizinga, D., & Weiher, A. W. (1993). Gang and non-gang youth: Differences in explanatory factors. *Journal of Contemporary Criminal Justice, 9*, 94-116.

Esbensen, F.-A., Matsuda, K., Taylor, T. J., & Peterson, D. (2011). Multi-method strategy for assessing program fidelity: The national evaluation of the revised G.R.E.A.T. program. *Evaluation Review, 35*, 14-39.

Esbensen, F.-A., Melde, C., Taylor, T. J., & Peterson, D. (2008). Active parental consent in school-based research: How much is enough and how do we get it? *Evaluation Review, 32*, 335-362.

Esbensen, F.-A., & Osgood, D. W. (1997). *G.R.E.A.T. program effectiveness.* Research in Brief. Washington, DC: U.S. Department of Justice, National Institute of Justice.

Esbensen, F.-A., & Osgood, D. W. (1999). Gang Resistance Education and Training (G.R.E.A.T.): Results from the national evaluation. *Journal of Research in Crime and Delinquency, 36*, 194-225.

Esbensen, F.-A., Osgood, D. W., Taylor, T. J., Peterson, D., & Freng, A. (2001). How G.R.E.A.T. is G.R.E.A.T.?: Results from a quasi-experimental design. *Criminology & Public Policy, 1*, 87-118.

Esbensen, F.-A., Peterson, D., Taylor, T. J., & Freng, A. (2010). *Youth Violence: Sex and race differences in offending, victimization, and gang membership*. Philadelphia, PA: Temple University Press.

Esbensen, F.-A., Peterson, D., Taylor, T. J., Freng, A., Osgood, D. W., Carson, D., & Matsuda, K. (2011). Evaluation and evolution of the Gang Resistance Education and Training (G.R.E.A.T.) program. *Journal of School Violence, 10*, 53-70.

Esbensen, F.-A., Winfree, L. T. Jr., He, N., & Taylor, T. J. (2001). Youth gangs and definitional issues: When is a gang and why does it matter? *Crime & Delinquency, 47*, 105-130.

Finn, P., & McDevitt, J. (2005). *National assessment of school resource officer programs: Final project report*. Abt Associates. Retrieved from: http://www.ncjrs.gov/pdffiles1/nij/grants/209273.pdf

Gatti, U., Tremblay, R. E., Vitaro, F., & McDuff, P. (2005). Youth gangs, delinquency, and drug use: A test of the selection, facilitation, and enhancement hypotheses. *Journal of Child Psychology and Psychiatry, 46*(11), 1178-1190.

Gottfredson, D. C. (2001). *Schools and delinquency*. New York, NY: Cambridge University Press.

Gottfredson, D. C., & Na, C. (2010). Police officers in schools: Do they prevent crime? Paper presented at the annual meeting of the American Society of Criminology, San Francisco, CA.

Gottfredson, G. D., Gottfredson, D. C., & Czeh, E. R. (2000). *National study of delinquency prevention in schools, Final report*. Ellicott, MD: Gottfredson Associates. Inc.

Hill, K. G., Howell, J. C., Hawkins, J. D., & Battin-Pearson, S. (1999). Childhood risk factors for adolescent gang membership: Results from the Seattle Social Development Project. *Journal of Research in Crime and Delinquency, 36,* 300-322.

Howell, J. C. (2009). *Preventing and Reducing Juvenile Delinquency: A Comprehensive Framework*. Thousand Oaks, CA: Sage.

Klein, M., & Maxson, C. L. (2006). *Street gang patterns and policies*. New York: Oxford University Press.

Lillehoj, C. J., Griffin, K. W., & Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and relation to youth outcomes. *Health Education and Behavior, 31*(2), 242-257.

Maxson, C. L., & Whitlock, M. L. (2002). Joining the gang: Gender differences in risk factors for gang membership. In C. R. Huff (Ed.), *Gangs in America* (3rd ed.) (pp. 19-36). Thousand Oaks, CA: Sage.

Maxson, C. L., Whitlock, M. L., & Klein, M. W. (1998). Vulnerability to street gang membership: Implications for practice. *Social Service Review*, March, 70-91.

McMorris, B. J., Clements, J., Evans-Whipp, T., Gangnes, D., Bond, L., Toumbourou, J. W., & Catalano, R. F. (2004). A comparison of methods to obtain active parental consent for an international student survey. *Evaluation Review, 28*, 64-83.

Melde, C., & Esbensen, F.-A. (2011). Gang membership as a turning point in the life course. *Criminology, 49*, 513-552.

Mihalic, S. F., Fagan, A., Irwin, K., Ballard, D., & Elliott, D. S. (2002). *Blueprints for Violence Prevention replications: Factors for implementation success*. Colorado: University of Colorado, Center for the Study and Prevention of Violence.

Mihalic, S. F., & Irwin, K. (2003). Blueprints for Violence Prevention: From research to real-world settings — Factors influencing the successful replication of model programs. *Youth Violence and Juvenile Justice, 1,* 307-329.

Office of Juvenile Justice and Delinquency Prevention. (2010). Model Programs Guide. Retrieved from http://www.ojjdp.gov/mpg/

Pyrooz, D. C., Fox, A. M., & Decker, S. H. (2010). Racial and ethnic heterogeneity, economic disadvantage, and gangs: A macro-level study of gang membership in urban America. *Justice Quarterly, 27*, 867-892.

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2009). *A user's guide to MLwiN, version 2.10*. Bristol, UK: Bristol University.

Reed, W. L., & Decker, S. H. (2002). *Responding to gangs: Evaluation and research*. Washington, DC: US Department of Justice, National Institute of Justice.

Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: US Department of Justice, Office of Justice Programs.

Snyder, H. N., & Sickmund, M. (2006). *Juvenile offenders and victims: 2006 National Report*. Washington, DC: Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention (March).

Thornberry, T. P. (1998). Membership in youth gangs and involvement in serious and violent offending. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 147-66). Thousand Oaks, CA: Sage.

Thornberry, T. P., Krohn, M. D., Lizotte, A. J., & Chard-Wierschem, D. (1993). The role of juvenile gangs in facilitating delinquent behavior. *Journal of Research in Crime and Delinquency, 30*, 55-87.

Thornberry, T. P., Krohn, M. D., Lizotte, A. J., Smith, C. A., & Tobin, K. (2003). *Gangs and delinquency in developmental perspective*. New York: Cambridge University Press.

Unger, J. B., Gallaher, P., Palmer, P. H., Baezconde-Garbanati, L., Trinidad, D. R., Cen, S., & Johnson, C. A. (2004). No news is bad news: Characteristics of adolescents who provide neither parental consent nor refusal for participation in school-based survey research. *Evaluation Review, 28*, 52-63.

Winfree, L. T. Jr., Peterson Lynskey, D., & Maupin, J. R. (1999). Developing local police and federal law enforcement partnerships: G.R.E.A.T. As a case study of policy implementation. *Criminal Justice Review, 24*, 145-168.

Appendix. Scale Characteristics of Outcome Measures (Wave 1)

Impulsivity: Four items such as: I often act without stopping to think.
Scale Mean = 2.97 (0.81); α = 0.59
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Risk-Seeking: Four items including: I like to test myself every now and then by doing something a little risky.
Scale Mean = 2.60 (0.95); α = 0.77
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Anger: Four items including: I lose my temper pretty easily.
Scale Mean = 3.08 (0.96); α = 0.74
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Self-Centeredness: Four items such as: If things I do upset people, it's their problem not mine.
Scale Mean = 2.50 (0.82); α = 0.69
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Attitudes Toward Police: Six items such as: Police officers are honest.
Scale Mean = 3.81 (0.82); α = 0.86
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

GREAT ATP: Two items such as Police officers make good teachers.
Mean = 3.58 (0.95)
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Prosocial Peers: Four items, including: How many of your current friends have gotten along well with teachers and adults at school?
Scale Mean = 3.42 (0.97); α = 0.83
Response Categories: 1) None of them, 2) Few of them, 3) Half of them, 4) Most of them, 5) All of them

Peer Pressure: Seven items such as: How likely is it that you would go along with your current friends if they wanted you to bully another student at school?
Scale Mean = 1.27 (0.51); α = 0.82
Response Categories: 1) Not at All Likely to 5) Very Likely

Negative Peer Commitment: Three items including: If your group of friends was getting you into trouble at home, how likely is it that you would still hang out with them?
Scale Mean = 1.68 (0.85); α = 0.81
Response Categories: 1) Not at All Likely to 5) Very Likely

Positive Peer Commitment: Two items: If your friends told you not to do something because it was wrong, how likely is it that you would listen to them?
Scale Mean = 4.19 (1.17); α = 0.80
Response Categories: 1) Not at All Likely to 5) Very Likely

Delinquent Peers: Seven items including: During the last year, how many of your current friends have attacked someone with a weapon?
Scale Mean = 1.30 (0.54); α = 0.86
Response Categories: 1) None of them, 2) Few of them, 3) Half of them, 4) Most of them, 5) All of them

Lying Neutralizations: Three items including: It's okay to tell a small lie if it doesn't hurt anyone.
Scale Mean = 2.60 (0.98); α = 0.76
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Stealing Neutralizations: Three items such as: It's okay to steal something from someone who is rich and can easily replace it.
Scale Mean = 1.64 (0.80); α = 0.83
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Hitting Neutralizations: Three items such as: It's okay to beat up someone if they hit you first.
Scale Mean = 3.32 (1.11); α = 0.80
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

School Commitment: Seven items including: Homework is a waste of time.
Scale Mean = 3.92 (0.70); α = 0.77
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Guilt: Seven items such as: How guilty would you feel if you stole something worth less than $50?
Scale Mean = 2.66 (0.55); α = 0.93
Response Categories: 1) Not Very Guilty/Badly, 2) Somewhat Guilty/Badly, 3) Very Guilty/Badly

Conflict Resolution: Five items including: During the past year when you've gotten upset with someone, how often have you talked to the person about why I was upset.
Scale Mean = 2.17 (0.46); α = 0.66
Response Categories: 1) Never, 2) Sometimes, 3) Often

Calming Others: Three items including: When someone else was upset, how often have you asked the person why he/she was upset.
Scale Mean = 2.41 (0.51); α = 0.71
Response Categories: 1) Never, 2) Sometimes, 3) Often

Refusal Skills: Four items including: During the past year when you have tried to avoid doing something your friends tried to get you to do, how often have you told the person that I can't do it because my parents will get upset with me.
Scale Mean = 2.33 (0.51); α = 0.70
Response Categories: 1) Never, 2) Sometimes, 3) Often

Pro-social Involvement (Index): Four items including: During the past year have you been involved in school activities or athletics?
Mean = 2.38 (1.14)
Response Categories: 1) No, 2) Yes,

Empathy: Five item including: I would feel sorry for a lonely stranger in a group.
Scale Mean = 3.63 (0.65); α = 0.59
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Active Listening: Three items such as: I look at the person talking to me.
Scale Mean = 3.66 (0.72); α = 0.60
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Problem Solving: Two items including: I talk to my friends about my problems.
Scale Mean = 3.57 (0.91); α = 0.45
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Self-Efficacy: Five items such as: When I make plans, I am certain I can make them work.
Scale Mean = 3.76 (0.65); α = 0.72
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Awareness of Services: Four items including: You know where a person can go for help if he/she is victimized.
Scale Mean = 3.76 (0.65); α = 0.72
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Collective Efficacy: Three items including: It is my responsibility to do something about problems in our community.
Scale Mean = 3.25 (0.77); α = 0.62
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Attitudes about Gangs: Two items: Getting involved with gangs will interfere with reaching my goals.
Scale Mean = 3.72 (1.12); α = 0.71
Response Categories: 1) Strongly Disagree to 5) Strongly Agree

Altruism: Three items including: It feels good to do something without expecting anything in return.
Scale Mean = 3.60 (0.83); α = 0.66
Response Categories: 1) Strongly Disagree to 5) Strongly Agree