

Growing Against Gangs and Violence (GAGV): Findings from a Process and Outcome Evaluation

Abstract

Objective: The present study assesses program efficacy of Growing Against Gangs and Violence (GAGV), a primary prevention partnership with the UK Metropolitan Police Service, delivered in London schools with the aim of reducing gang involvement, delinquency, and violent offending and improving young people's confidence in police. GAGV is partially derived from an American program, Gangs Resistance Education and Training (GREAT).

Method: A qualitative process evaluation and randomized control trial outcomes study were undertaken.

Results: Findings indicate GAGV personnel were keen to enhance program fidelity and process implementation. The RCT did not demonstrate a statistically significant program effect. However, effect sizes indicate the program was effective in reducing levels of gang membership and the frequency and variety of delinquency and violence in the short- and longer term. More robust evidence indicated GAGV also improved students' attitudes toward police and reduced their adherence toward street code.

Conclusions: The use of cohort- (not individual-) level data and missing data in the one-year follow-up make it difficult to draw reliable and robust conclusions. However, results are encouraging. Several recommendations are suggested for GAGV, including curriculum design, regular evaluations, and expanding to include more schools. Limitations of this and similar evaluations also are discussed.

Keywords: Gangs; Delinquency; Randomized Control Trial; Primary Prevention; Youth Violence; Sexual Violence.

Introduction

Street-gang members are more likely to engage in various violent or non-violent offending than their demographically similar counterparts (Klein & Maxson, 2006; Melde & Esbensen, 2013). However, the strength of this relationship and the influence of mediating and moderating variables (e.g. demographic, contextual and communal) varies (Pyrooz, Turanovic, Decker & Wu, 2015). Once seen as being problematic only in the United States, gangs are a recognized global phenomenon (Hagedorn, 2005; Howell, 2015), subject to intensive scrutiny (Decker, Melde & Pyrooz, 2013). In the United Kingdom, gang proliferation and an increase in related crime has garnered interest among law enforcement, policy-makers, and researchers (Densley, 2013; Deuchar, 2013; Hallsworth, 2013; Harding, 2014; Pitts, 2008; Smithson, Ralphs, & Williams, 2013) and precipitated the need for expedient and effective intervention (see Densley, 2011). The present study presents the results of a process and longitudinal outcome evaluation of one such gang and juvenile delinquency program: Growing Against Gangs and Violence (GAGV).

Inspired by the schools-based Gangs Resistance Education and Training (GREAT) program in the United States (Esbensen, Osgood, Peterson, Taylor, & Carson, 2013), GAGV is a primary prevention program delivered in partnership with London's Metropolitan Police Service (MPS). GAGV delivers law enforcement-themed gangs resistance education in schools with three basic goals: (1) reduce gang membership, (2) reduce (violent) youth offending, and (3) develop a positive relationship between young people and police. The current study asks whether GAGV is effective in meeting these goals, by (1) evaluating its implementation processes, particularly the

extent to which curriculum delivery matched core program aims, and (2) evaluating outcomes on attitudinal change, self-report offending and gang activity.

Gangs Resistance Education and Training (GREAT)

Few gang prevention and intervention programs have undergone rigorous program evaluations and fewer still have undertaken randomized control trials (Esbensen et al., 2013; Klein & Maxson, 2006). Gangs Resistance Education and Training (GREAT) is an exception to this rule. No other school-based primary prevention program enjoys GREAT's longevity or "name recognition" (Gravel et al., 2013; Papachristos, 2013, p. 370). Uniformed police officers deliver 13 sessions to early adolescents, with three main goals: (1) reduce student involvement in crime and delinquency; (2) teach the consequences of gang affiliation; (3) improve relationships with and attitudes toward police (Esbensen, Peterson, Taylor, & Osgood, 2012).

The original incarnation of GREAT was subjected to two multisite evaluations. The first, a cross-sectional study, found some evidence of outcome success (Esbensen & Osgood, 1999). The second, more robust, study was less supportive of program efficacy, although showed a significant "sleeper" effect at years three and four follow up (Esbensen et al., 2001). In response to the questions that thus arose over efficacy (Klein & Maxson, 2006; Ludwig, 2005), key stakeholders revised and retooled GREAT (Maxson, 2013).

The new GREAT incorporated problem-solving strategies and cooperative learning activities, along with evidence-based risk factors for gang membership (Esbensen, et al 2011). Pre-rehearsed lecture style delivery was replaced with a more conversational tone. GREAT 2.0 was "proven effective" (Howell, 2013) in a seven city, randomized control trial with four-year follow-up, involving 3,820 students (Esbensen et al., 2012, 2013). Esbensen and colleagues

(2012, 2013) found the program reduced the odds of gang membership by 39 percent over one year. GREAT also improved youth-police relations and fostered more negative views of gangs. There was no concomitant reduction in recidivism, however, which was deemed “surprising” (Maxson, 2013, p. 423) and “inconvenient” (Pyrooz, 2013, p. 427) given the positive association between gangs and delinquency (Thornberry et al., 2003; Pyrooz et al., 2015).

Growing Against Gangs and Violence (GAGV)

GAGV is heavily influenced by, indeed partially replicates, GREAT.¹ GAGV is a primary prevention initiative (see Decker & Curry, 2003; Howell & Griffiths, 2015; Klein & Maxson, 2006; Spergel, 2007; Wyrick, 2006), thus designed to build resilience, not suppress gangs (Triplett & Ross, 1998) or target provision based on assessed risk (e.g., Hennigan, Kolnick, Vindel & Maxon, 2015; Higginson et al., 2014). Like GREAT, GAGV is universally delivered to entire school year groups and the well-documented “push” and “pull” factors that promote and sustain gang membership direct the curriculum scope and sequence (e.g., Decker et al., 2013; Harris, Turner, Garrett, & Aitkinson, 2011). GAGV seeks to challenge moral disengagement (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996) and to cultivate resilience and critical engagement with more pro-social and fewer anti-social groups (Tajfel & Turner, 1986), in part by increasing dissonance between gang myths (e.g., messages about ostensibly protective gangs) and realities (Esbensen & Matsuda, 2013; Howell & Griffiths, 2015; OJJDP, 2008). A key focus is naming and understanding “gang processes” conducive to violent offending (see Decker et al., 2013) so they can be avoided. Like GREAT, an informal,

¹ GAGV has a memorandum of understanding with GREAT and sent the first UK-based police officers to the United States for training.

conversational style is encouraged in sessions, and no formal script is used, although facilitators do use standardized core curriculum and slides and comprehensive supplementary materials.

GAGV also distinguishes itself from GREAT in several ways. Although it considers both skills development (e.g., refusal skills, conflict resolution, mood management) and issues awareness, for example, the GAGV curriculum is focused more on issues awareness, thus closer to GREAT 1 or “Educating Kids About Gun Violence” (Hayward et al., 2011). GAGV still encompasses gang resistance education and training, but because personal, social, health (including sexual relationships), and economic (PSHE) education is part of the National Curriculum for UK schools, some of GREAT’s life skills focus, necessary in the US-context, is excluded. GAGV thus comprises six lessons, compared to GREAT’s 13, but has the scope to rehearse police “stop and search” encounters in a non-confrontational setting and focus more on the legal, medical, social and emotional consequences of knife and gun crime, drug crime, cyber bullying, and peer-on-peer sexual violence (see online supplementary file for full curriculum outline).

GAGV’s revised curriculum both reflects and responds to growing understanding of young people’s agency and decisions in the London street gang context (Densley & Stevens, 2015) and the universality of correlates for gang participation (Haymoz, Maxson, & Killias, 2014). Recent empirical work has identified individual-level profiles and motivations and group-level structures and processes in London, specifically, that share more in common with the US than the rest of the UK (Alleyne & Wood, 2012; Densley, 2014; Harding 2014; Pitts, 2008). GAGV works in local authority areas prioritized under HM Government’s (2011) *Ending Gang and Youth Violence* initiative and deliberately places greater prominence on related public health

and policy concerns, including disproportionate rates of knife and gun crime, drug dealing, and sexual violence and child exploitation among gang-associated youth (Disley & Liddle, 2016), specifically multiple perpetrator rape (Densley, Davis, & Mason, 2013).

Bespoke, professionally produced, video content supplements and reinforces the GAGV curriculum and unlike GREAT, police officers deliver alongside paid facilitators, who are often teachers or youth workers. In some cases, community partners replace police officers entirely. Rape crisis workers, for example, aid in delivering sexual violence topics, which also include single sex groups. This enables engagement based on student group gender, interest, readiness, and learning styles (Densley, 2013). Further, it facilitates proper handling of any voluntary disclosures. In its most recent years of operation, GAGV worked closely with Local Safeguarding Children's Boards to identify referral pathways, signpost support, and identify emerging trends. Facilitators receive two days of intensive pre-service training, slightly less than that which GREAT officers typically receive, but engage in more team-teaching and ongoing, post-session debrief and evaluation than GREAT officers.

Current Study

According to internal data, GAGV was delivered to over 90,000 young people in approximately 450 London schools and Pupil Referral Units in the first six years since its founding in 2008.² Thus, it was vital to ask, is GAGV effective in meeting its aims? This paper documents short- and long-term results from a multisite, process and outcome evaluation of the GAGV program. Challenges for evaluative frameworks lie in the complexity of the phenomena under scrutiny, lack of sufficient funding, and methodological difficulties inherent in measuring

² It should be noted that the format and delivery of particular sessions developed over the course of the process evaluation. The outcome evaluation was completed after such changes had been implemented.

outcome and impact generally (Esbensen et al., 2012; Heckman, 1997; Heckman, Smith, & Clements, 1997). In approaching the evaluation of the GAGV program, we drew on best practice for evaluation standards (Project Oracle, 2015; Sherman et al., 1998; WSIPP, n.d.). The approach was set within an understanding of gang behaviors, etiology and interventions in the USA and UK, (Howell & Griffiths, 2015; OJJDP, 2008; Pyrooz et al., 2015) and was refined in close consultation with service users. This paper has two research purposes:

Research Aim (RA) 1: To evaluate implementation processes, particularly the extent to which curriculum delivery matched core program aims.

Research Aim (RA) 2: To evaluate outcomes on attitudinal change, self-report offending and gang activity.

Methods

RA 1. Process Evaluation

GAGV commissioned Forensic Psychological Services at Middlesex University to conduct some basic process evaluations in 2011 and 2012. Initially conceived as observational research, the remit was concerned mainly with asking whether the curriculum met its criminological intentions and was being delivered with fidelity. Given GAGV's willingness and alacrity in adopting recommendations made, the research may be better considered within a more traditional action research approach. Processes were changed as a result of recommendations provided at an interim phase, thus becoming the ongoing subject of evaluation.

Participants and Procedure. All syllabi, handbook, and supporting materials including films and presentation slides were reviewed in terms of their content and accessibility. They informed analysis of the observations also conducted. Eight observations were made in three

schools, one primary and two secondary. Horvath and Adler made the observations and conducted the analyses. The process team was independent of GAGV. Findings were presented to GAGV in two interim and one final report. This paper is the first public reporting of data extracted from those reports. The schools and sessions observed were selected opportunistically based on school and research team teaching schedules. Each session observed was delivered to a targeted age group, centering on the key topics elucidated further in the supplementary materials.

RA2. Outcome Evaluation

Following the process evaluation in 2011–12, and subsequent amendments to the curriculum materials in 2012–13, a three phase randomized control trial (RCT) was implemented in academic year 2013-14, in four locations. There was no overlap between schools in the two evaluations.

Participants. Sixteen Year 8 cohorts (age 12-14) from four schools (labeled as “S1”, “S2”, “S3”, and “S4”) in four of London’s 32 boroughs were recruited. The goal was to develop a geographically and demographically diverse sample of young people from schools that had never previously received GAGV (thus protecting against spillover or contamination) but were still located within local authority areas prioritized under HM Government’s (2011) *Ending Gang and Youth Violence* program. Within each school, two cohorts (year 8 class groups) were randomly assigned to the “intervention” group and two cohorts were randomly assigned to the “control” group.

In total, 391 students (198 intervention and 193 control) completed the pre-test survey and 361 students (193 intervention and 168 control), or 92.3%, completed the post-test survey. One-year follow-up was achieved with 184 intervention students, but only 67 control students because

of administrative issues in schools 1 and 3. With control group data for only two schools available (S2 and S4), there were 251 participants, or 64%, in wave three.

Sample demographics are shown in Table 1. There were no statistically significant differences between the intervention and control groups in terms of gender, age, ethnicity, school performance, or self-reported gang membership in wave one. In school level, however, there were statistically significant differences ($p < 0.05$) in students' age, ethnicity, and self-reported gang membership. Students in S2 (N=87) were younger and more likely to be Black. Students in S1 (N=111) were more likely to be Asian. And students in S3 (N=91) were more likely to be gang members.

Consistent with other studies (Klein & Maxson, 2006; Pyrooz & Sweeten, 2015), the prevalence of gang membership was small: 4.8% in the control group and 1.5% in the intervention group in wave one, a difference of 3.3 percentage points. Over time, gang involvement increased in both the control and intervention groups, although the changes were different and insignificant (albeit close to significance in wave two, see Table 1). Gang involvement was initially higher in the control group than in the treatment group, and the gap widened in waves two and three, confirming that analyses beyond exact test and direct comparison measures are required to uncover possible program effects.

(Insert Table 1 about here)

Measures. The surveys measured GAGV's three program goals to (1) reduce gang membership, (2) reduce levels of youth violent offending, and (3) develop a positive relationship between young people and police. The measures will be familiar to readers of the GREAT evaluation (Esbensen et al., 2013). *Gang membership* is measured by a single-item question,

“Are you now in a gang?” which was part of a larger set of questions about gangs. Self-nomination has been shown to be a reliable and valid measure of gang membership (Curry, 2000; Decker et al., 2014; Esbensen et al., 2001; Thornberry et al., 2003), “as reliable as, if not more reliable than, most social science measures” (Thornberry & Krohn, 2000, p. 72).³

Delinquency is measured by 14-item self-reported delinquency inventory that assess frequency of offending (e. g., sold illegal drugs) in the preceding six months (two months at Wave 2, the post-test). Response categories are (1) Never to (5) more than 10 times. Scale mean was 4.91 and Cronbach’s α [alpha] was 0.91. *Violent offending* is measured via a subset of three items from the delinquency inventory (i.e., attacked someone with a weapon, used a weapon or force to get money or things from people, been involved in gang fights; scale mean = 0.61; α = 0.85). These scales are widely used in gang research and demonstrate strong internal consistency (Esbensen et al., 2013). *Attitudes toward police* is derived from a six item inventory regarding perceptions of the honesty, work ethic, friendliness, courtesy, and respectfulness of police officers; and whether or not respondents felt safer when police officers were present in their schools (Response categories: (1) strongly disagree to (5) strongly agree; scale mean = 3.48; α = 0.90).

In addition to the three program goals, the GAGV curriculum aims to enhance life skills and resilience (Klein & Maxson, 2006). The following mediating variables are implied program objectives, therefore, included in the outcome evaluation:

³ We acknowledge that whether a “gang” is understood to be the same in different contexts is a matter of some debate (see Smithson, Monchuk, & Armitage, 2011). However, this research was conducted within one main context: London. As such, potential differences in rates of self-nomination between research and control groups are more likely to be associated with the intervention, than a contextual difference in recognizing gangs.

1. *Attitudes to gangs* (three items, including: “Getting involved with gangs will interfere with reaching my goals.” Response categories = 1) strongly disagree to 5) strongly agree; scale mean = 3.27; $\alpha = 0.67$);
2. *Adherence to street code* (seven items, including: “People tend to respect a person who is tough and aggressive.” Response categories = 1) strongly disagree to 5) strongly agree; scale mean = 3.06; $\alpha = 0.81$);
3. *Refusal skills* (five 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.” Response categories = 1) never, 2) sometimes, 3) often; scale mean = 2.45; $\alpha = 0.57$);
4. *Conflict resolution skills* (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.” Response categories = 1) never, 2) sometimes, 3) often; scale mean = 2.13; $\alpha = 0.59$);
5. *Resistance to 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?” Response categories = 1) not at all likely to 5) very likely; scale mean = 1.25; $\alpha = 0.85$);
6. *School commitment* (seven items, including: “Homework is a waste of time.” Response categories = 1) strongly disagree to 5) strongly agree; Scale mean = 3.65; $\alpha = 0.47$).

In total, three behavioral outcomes (gang membership, delinquency, violent offending) and six attitudinal measures are measured. Initial and post levels for all measures are presented in the online supplementary file.

Procedure. Intervention group students participated in the entire six-lesson GAGV curriculum over a five-week period. This delivery is atypical as sessions are targeted to different school year groups (from ages 10 to 15). Any one young person would normally receive them over five years, in at least two schools, rather than five weeks in the early part of their secondary education. This atypical delivery was the only way in which schools would agree to an RCT as they were so keen for delivery of the program to all pupils. The implications of this modified delivery are considered further in analytical strategies, limitations and discussion. Students from the control group did not receive GAGV sessions.

All students in the intervention and control groups voluntarily completed a confidential group-administered pre survey (wave one). After completion of the GAGV program, students in both the intervention and control groups were requested to complete a post-test survey (wave two) and one-year (wave three) follow up survey. In order to measure change in behaviors and attitudes, the same survey was used in each wave.

Ethics

Both the process and outcome evaluations were designed based on best principles outlined in the introduction, but also to facilitate access and utility to the participants, and to conform to professional and statutory codes of conduct. We worked closely with service users (schools) and with research commissioners during the design phase. There were several constraints that had to be met, these include limited resources available to pay for research time and effort, but most importantly, the concerns schools had regarding delivery of the program and recruitment to the research. After the processes for each evaluation had been designed, each

study was submitted to the relevant University ethical committees. In each case, the studies were approved for implementation without further amendment.

For the process evaluation, consent to observe sessions was obtained via the schools, acting *in loco parentis*. No children directly participated in the study and the sessions observed, were agreed, *a priori* as part of normal school arrangements with GAGV. Prior to audio recording of sessions, all young people, facilitators, and teachers were introduced to the researchers and the purpose of the observation was explained. Each was given the opportunity to have the session observed but not recorded if they preferred; all opted for the sessions to be recorded. Students were not compensated for their participation.

The outcome study involved direct child participation in the research and an intervention being either offered or not, to school-age children who would otherwise not have received that primary prevention program. It was therefore decided the consent of a parent or guardian was required for students' participation in the outcome evaluation.

School administrators and teachers identified "minimal risk" from the outcome study (see Pokorny et al., 2001) and felt it unnecessary to introduce additional safeguards to protect students from questions about subject matter they regularly confronted in school, the media, at home, and in their peer groups (Esbensen et al., 1996, 1999), hence passive parental consent was used, whereby parents had to specify if they wished their child to be excluded from the study. To ensure informed consent, schools sent multiple letters home about the evaluation, including an explanation of the risks and benefits of students' participation and sample questions. Parents and guardians were also invited (and many accepted) to attend one of a series of information sessions

about the evaluation. Passive consent procedures resulted in a very high participation rate (over 90% in all schools), thus a sample representative of the enrolled students.

Mindful of data protection duties and privacy, schools requested that students were tracked at only the classroom or cohort-level, not individually. This limitation was planned for accordingly and is considered further in the next section.

Analytical Strategies

The process evaluation was based on an inductive thematic analysis (Braun & Clarke, 2006). Recordings were transcribed,⁴ coded, and analyzed by the same researchers as observed the sessions. Horvath led on sexual violence related sessions and Adler led on the others. Each lead drew on the other, to assess rater reliability and interpretation. Each session observed, was reported separately first, using a grounded theory approach with no observational protocol. Themes emerged through consideration of all sessions observed and of the archival materials.

The outcome evaluation was designed in the knowledge there would only be classroom-, or cohort- level, matching information available. As such, Difference in Differences (DID) analyses were performed, a technique designed specifically to detect potential change at the higher cohort level (Abadie, 2005; Lechner, 2011). Basic DID models look for group-based differences where individual matching is not possible and assume two factors (each with two levels) influence the intervention effect: intervention (versus control) and time period (before versus after, or pre-test versus post-test). DID constructs a regression model with the form,

$$y = \alpha + \beta \times \textit{treatment} + \gamma \times \textit{period} + \delta \times \textit{treatment} \times \textit{period} + \varepsilon$$

⁴ As can be seen from the verbatim extracts in the supplementary file, budget constraints meant that full Jefferson transcription conventions could not be followed.

Where y denotes the outcome measure of interest, intervention and period are 0-1 dummy variables, where intervention group and post intervention measures were coded as 1, and control and pre intervention measures were coded as 0. The DID model divides the total effect into four parts, including impact of group (β , stable across time), impact of time (γ , stable across groups), impact of intervention (δ), and random noise. The effect size of intervention is measured by δ , which is the coefficient of the interaction term (see Table 2).

(Insert Table 2 about here)

The least square estimation for δ is equal to

$$\hat{\delta}_{LS} = (\bar{y}_{11} - \bar{y}_{10}) - (\bar{y}_{01} - \bar{y}_{00})$$

$$Var(\hat{\delta}_{LS}) = \delta^2 \left(\frac{1}{n_{00}} + \frac{1}{n_{01}} + \frac{1}{n_{10}} + \frac{1}{n_{11}} \right)$$

Where, \bar{y}_{ij} denotes the mean value for group defined by dummy variables *intervention* and *period*; δ^2 is variance of the regression model; n_{ij} is the number of observations in each group. DID accounts for how intervention and control students are nested within classrooms and within schools, and in an effort to improve the accuracy of effect size (i.e., reducing degrees of freedom for the sum of squared residuals), an elaborated model also was created to include additional variables (e.g., age, ethnicity, gender, grades, gang membership) and their interactions, but the coefficient δ changed very little. Part of the challenge is the rather small sample size in each group (e.g., 391 total participants corresponds to an average intervention or control group of 50 per school or 25 per class), hence we present the basic DID model with a logistic regression model for gang membership because it is coded as a binary variable.

Findings

RA1. Process Evaluation

The process evaluation aimed to evaluate implementation processes, particularly the extent to which curriculum delivery matched core program aims. Evaluation reports provided to GAGV included detailed assessments of materials used and interactions with young people, including confidential recommendations. The themes presented here are those most likely to be generalizable. It should first be noted the curriculum was clearly underpinned by theory and evidence based practice and was strongly tied to program aims. These findings therefore focus on its implementation and delivery.

Good Will and Positive Reception. In all sessions observed, we saw a willingness by facilitators and schools to embrace GAGV program ideals. From the schools' perspective, the curriculum augmented and complemented existing provision, particularly around bullying and sexual violence (see, "Positive extract..." in the online supplementary file). When schools shifted priorities and scheduling at the last minute, however, materials prepared for pupils in one year would be delivered to another and cohort sizes would exceed pre-agreed sizes.

There was an openness and willingness from police officers to talk about their work in an accessible, engaging way. Community support officers explained everything from how stop and account is supposed to work, or how to ask for a police officer's badge number, through to just how much mucus to expect when CS/pepper spray is deployed. The police officers were also flexible about where and how they fitted into sessions.

Young people were generally very well behaved and engaged in the sessions, showing creativity in drama activities, asking questions, and providing relevant insights after films. The length of the sessions may, however, have contributed to some instances of fidgeting and muttering but poor behavior never became an issue within the sessions observed.

Good Facilitation is Key. There were some observations of excellent practice. Some facilitators were adept at holding information generated by the group and returning to it at relevant later points. The best facilitators did so in neutral, non-judgmental ways. They also conveyed credibility, either through their previous experiences or through the expertise with which they answered questions. However, all facilitators demonstrated difficulty in managing timings. Balancing the quantity of material with the quality of interaction, allowing the flow to be influenced by the young people within a session, was the biggest practical challenge observed. Some facilitators also clearly had problems in gauging how much of the supporting materials had been provided for their benefit, how much should be shared with the young people, and when to share information with them (see supplementary file, “Negative example...”).

Facilitators’ general group management skills also varied and we noted there was no systematic training in group dynamics. Rather, facilitators were recruited because they had an education or youth work background. Although they were by no means thrown in at the deep end, the initial training, mentoring, and shadowing did not explicitly tackle group dynamics.

One of the strengths of the sexual violence sessions was evidenced in the detailed questions that the young people asked, but the quality of responses varied. In one sexual violence session, for example, questions included: “What happens when a suspect is arrested on suspicion of rape? What is a penile swab? Does it matter if you have had sex with someone else since the alleged rape?” These questions were largely skimmed over by facilitators, when observed. We also did not see evidence of appropriately tailored challenge being provided to young people when considering the sexualized roles of girls and young women within gangs.

Variable Quality of Materials. The supporting materials for facilitators and police officers were theoretically grounded and replete with academic and practitioner knowledge. However, they were difficult to implement systematically or consistently. The PowerPoint slides needed revision, for clarity, professionalism, and ease of use. Key messages were neither highlighted, nor repeated on the slides. Slides were revised prior to implementation of the outcome evaluation to address these problems.

The short films were always well received. They prompted lively discussion of complex, sensitive issues. Production values were extremely high and it was noticeable that the different cinematographic techniques and genres used helped in reaching the young audience. One film, *A Mother's Tear*, for example, was produced by local young people and professional filmmakers, many of whom were previously gang involved. Another film, *Knife City*, was produced like a virtual reality game, whereas *The Surgeons Speak* had *vox populi* foregrounded within actual casualty wards in which consultant trauma surgeons worked.

Poor Signposting and Unclear Onward Referral. Facilitators neither explained the purpose of the particular session they delivered nor did they directly consider the difficult nature of material that would be considered within it (even after initial feedback to this effect). There was an instance in one session of a young person who had been good friends with the younger sibling of a boy who had been murdered. This killing was focused on within a film and subsequent discussion with no additional sensitivity shown. Further, follow-up and onward referral information was not routinely provided either to teachers or young people within the sessions, although they were always provided to the schools at the outset. We saw instances of inappropriate behaviors, and pointers from young people who were either already involved in

gangs or likely to be at high risk of involvement, and signs that some participants had been victims of violence and were in need of support. When this point was first raised, GAGV responded that it was worried about “mission creep.” Once GAGV started to receive unsolicited disclosures of violent and sexual assaults (approx. 40 per year), however, they rapidly increased the training of facilitators and became fully embedded in local safeguarding children boards. GAGV also adjusted their mission statement to improve their response to sexual violence and child exploitation, both peer-on-peer and adult-perpetrated. These changes had been implemented by the time the outcome study was undertaken. At the time of this writing, moreover, GAGV was preparing online materials both for young people and their parents.

RA2. Outcome Evaluation

GAGV aims to reduce gang membership and levels of youth violent offending, while developing positive relationships between young people and police. These aims are the central focus of the outcome measures, although proximal outcomes also are briefly explored. The outcome evaluation considered both short-term (post-test) and long-term (one year after program implementation) program effect, but because one-year follow up data were available for only two of the four schools (S2 and S4, N=159), the long-term evaluation must be viewed with caution.

As already outlined, a DID regression model was used to measure the GAGV program effects. DID results are more reliable and robust than ordinary exact test and direct comparison methods (see online supplementary file), in accounting for both pre-existing differences between control and intervention groups and changes in measures across time. Table 3 shows the effect size (ES) in relation to each primary and proximal outcome measure (the intervention effect size is measured by δ , the coefficient of the interaction term). Of the 12 outcome measures, effect

sizes reached $p < 0.1$ ⁵ for only two measures in both the short- (post-test) and long- term (one year later) analyses. These are, positive attitudes toward police (ES=0.243 in short term, ES=0.467 in long term) and reduced adherence to street code (ES=-0.242 in short term, ES=-0.541 in the long term), the latter of which is related to gang- and delinquency-related outcomes (Matsuda et al., 2013; Stewart & Simons, 2010).

(Insert Table 3 about here)

The primary concern of the evaluation was whether the three main aims have been achieved. With regard levels of gang membership, the pre-post (wave 2) intervention odds ratio for those students in the intervention group was 6.88% smaller than those in control group, but too small to be significant. Regarding delinquency and violent offending, students in the intervention group reported a wider variety of delinquent activity (ES = 0.074), but less frequency of delinquency (ES = -1.427), less variety of violent offending (ES = -0.091), and less frequency of violent offending (ES = -0.549) compared to the control group. As previously discussed, the intervention group also expressed greater confidence in law enforcement and a reduction tended towards significance in subscription to the code of the street that promotes violence as a means of dispute resolution.

When comparing the results from the pre-test and one year follow up (wave 3), the odds ratio for the intervention group to be gang involved was 2.72% (ES = -2.72%) smaller than those control students (compared with -6.88% in pre and post survey), but the insignificant p value alerts this finding should not be overstated. Similarly, variety (ES=-0.382) and frequency (ES= -4.277) scores for delinquency and variety (ES= -0.113) and frequency (ES= -0.879) scores for

⁵ The significance level of 0.1 is used here to indicate a trend towards significance and to compensate for a possible lagged effect, as found in the original GREAT evaluation. Please note that in the one year follow up, significance at .05 is found.

violent offending were lower for the intervention group than the control group, which basically confirms the findings of the pre-post survey analysis. Consistent with the results of pre-post analysis, GAGV also showed its effectiveness in improving youth-police relationships and reducing subscription to the street code, but this time at the .05 significance level.

GAGV also showed some marginal gains relating to proximal goals, in terms of improving students' commitment to school (ES= 0.074), although the effect weakened with time (ES=0.027). Students' attitudes toward gangs, use of refusal skills, resistance to peer pressure, and use of conflict resolution skills evolved in different patterns in the short and the long term. Specifically, GAGV students reported more negative attitudes about gangs in the short term (ES=0.226) but more positive attitudes about gangs in the long term (ES=-0.097), which may reflect the actual proportional increase in gang involvement over time. Intervention students reported less use of refusal skills (ES=-0.345), less resistance to peer pressure (ES=0.057) and less conflict resolution skills (ES=1.0880), in the short-term, which was unexpected, however, those attitudes reversed one year after, a potential sleeper effect.

Discussion and Conclusions

Limitations

Both studies had strong design principles, but were implemented according to a number of constraints. In the process evaluation, these mainly related to the number of observations made and depth of analysis possible. In addition, no interviews were conducted. In the outcome evaluation, the absence of matching information on the individual level and missing data for control groups in schools 1 and 3 must curb enthusiasm for GAGV's achievement of primary and proximal goals. Despite the schools' objections, measuring program effects at the *individual*

level remains the most appropriate means of evaluation and is strongly recommended going forward. The absence of control cohorts in schools 1 and 3 in the one-year follow-up, moreover, had a major detrimental impact on retention and attrition rates, which otherwise were impressive.

Second, the process evaluation and outcome evaluation were not concurrent. The GAGV program made alterations to their implementation strategy *prior* to the outcome evaluation, which although indicative of the program's responsiveness to research findings, prevents us from concluding the program was taught with fidelity during the outcome evaluation; which, in turn, and along with other limitations discussed, could contribute to the lack of significant findings. The outcome evaluation was also based on an expedited, atypical program delivery, which influences dosage intensity and rate of exposure. While the curriculum content and administration was consistent, the focus on Year 8 students was a logistical decision that prevented differentiation based on students' age and readiness. The content in the sexual violence session, in particular, is more appropriate for older students, which, although schools, students and ethics committees consented, still should be acknowledged. Additionally, by concatenating presentation, only following up at peak times of selection into gangs (Pyrooz & Sweeten, 2015), and not engaging with young people over a number of years, it is likely that this design may have underestimated the impact of a more typical program delivery.

Another limitation pertains to the low alphas for some of the mediating variables (e.g., refusal skills, conflict resolution skills, school commitment), which, might have influenced the overall study findings. Discussion about using Cronbach's alpha to assess internal consistency for scales with very few items is beyond the scope of this paper (see Cortina, 1993; Dunn, Baguley, & Brunnsden, 2013). Suffice it to say, we are more cautious about those findings.

Research Implications

Further evaluation to remedy the limitations above is recommended. However, we can still conclude GAGV may be partially effective in meeting its aims, as follows.

RA1. Process Evaluation. Esbensen and Matsuda (2013) argue that a well-designed evaluation must first identify the goals or intended outcome of the program then confirm whether or not the program was implemented as designed. Only when a program is delivered with high fidelity, can any detected intervention effect be attributed to the program itself. The process evaluation broadly indicated that fidelity had been established although there was some variation in practice quality and the curriculum used in the outcome evaluation had been updated based on recommendations made after the process evaluation. If repeated, then a process evaluation would ideally observe more than one of each session and a richer, deeper analysis would also be facilitated by the direct participation of young people, GAGV personnel, and school administrators in the research.

RA2. Outcome Evaluation. The present study is only the second after the GREAT evaluation (Esbensen et al., 2013) to have attempted a randomized control trial in which some limited, but demonstrable impact has been found in terms of gang-related activity. Owing to the small sample size and the fact that statistical significance relates only to the likelihood results obtained were not due to chance (see Hubbard & Lindsay, 2008), we focus attention on the practical significance of effect size.

Basic descriptive analyses show overall levels of gang membership increased in both control and intervention groups in the post-test and one year follow-up stages (see Table 1). The DID regression analysis, however, lends support to GAGV's effectiveness in gang membership

reduction in the short and longer term, although the effect weakened over time. The results from descriptive analysis and the DID regression analysis thus seem to be contradictory, except that descriptive analysis considers gang membership only in one single group (i.e., intervention vs. control), whereas DID accounts for pre-existing differences between control and intervention groups and the changes in measures across time. Therefore, the increased overall level of gang membership may reflect the aging of cohorts towards a peak age of gang recruitment, or some unobserved contextual changes—GAGV curtailed such growth in the intervention groups.

Likewise, the overall frequency of delinquency and violent offending increased across time (see supplemental file), but the DID analysis suggests, increases in the intervention group were smaller than in the control group. Students in the control group had obviously more negative views toward police in the post-test stage and an even worse outlook on police one year later. In contrast, intervention group students viewed police officers more positively, something sustained one year later. At a time when police-community relations and public trust are under intense scrutiny, this is particularly salient (Maybin, 2014). Finally, GAGV may have achieved some proximal goals, improving students' commitment to schools, reducing their adherence to street code and gang norms, and promoting refusal skills, resistance to peer pressure, and non-violent conflict resolution.

Clinical and Policy Implications

RA1. Process Evaluation. None of the process observations were directly considered in terms of impact or outcome. For example, a powerful session observed related to the realities of prison life for male sex offenders. This is similar to a “scared straight” approach, which confers credibility but has not been shown to deter offending (e.g., Aos & Drake, 2013).

Most of the recommendations made to GAGV are relatively common within organizations that have started small and grown rapidly. They need to build capacity and maintain sustainability whilst also maintaining scrutiny of programming quality and efficacy. Practical recommendations made included advice regarding shaping materials, structuring sessions, training and management of facilitators and coordination of external relations. In particular we recommended that:

1. GAGV works more closely with schools for follow up and that it develops its website to host materials that are tied into those run within schools' extended activities;
2. Sessions be revised to include better signposting and to support participants better through sensitive material. The impact and salience of local victims is commendable yet, to use emotive, personally relevant cases requires better welfare consideration;
3. Sexual violence sessions needed to acknowledge that not all young people are heterosexual. Also checks need to be made that young people understood when material was being presented as problematic and not as acceptably normative, for example when considering the roles of girls and young women within gangs;
4. Facilitators need to engage with the young people who ask questions, whilst being aware of potential risks of providing too much information to a possible perpetrator, and acknowledging that some in the room may be victims of such behavior. To skim over awkward questions seems the least satisfactory form of response;
5. Better routes for referral should be identified, both within the school and beyond it. While acknowledging the point of a cohort program is not to specifically target those already known

to be at risk, when they present themselves as such, it would make sense for some form of onward referral to exist. Victims needed clear and robust safeguarding.

GAGV facilitators may be the person of first disclosure yet they are not employed to provide assistance, are not therapists, nor always police officers. Where safeguarding and, or other criminal matters are of concern, then facilitators need to know who to report information to, they need to explain to the young person what they can and what they cannot do to help support that young person and GAGV needs to have routes to demonstrate adherence to such processes.

RA2. Outcome Evaluation. To date, over 90,000 young people have had experience of an intervention that, based on the evidence gathered thus far, only works marginally effectively. At the same time, we have no real evidence of harmful effects and some indication that the methodology may have underestimated positive effects. GAGV has historically received funding and support both from local government and the Mayor's Office for Policing and Crime (which sets the direction and budget for the Metropolitan Police Service on behalf of the Mayor of London). The question remaining is whether marginal positive improvements are sufficient?

Findings considered here are similar to those from the most recent GREAT evaluation (Esbensen et al., 2013) in that they showed results in the predicted direction without a statistically significant program effect but with longer term "sleeper effects" on pro-social and anti-gang attitudes. As Pyrooz (2013, p. 430) rightly asked of GREAT, "was the base rate of gang membership too low for statistically and substantively meaningful differences in criminal offending to emerge?" Does GAGV focus on the "wrong kinds of people", or the "wrong kinds of mechanisms" to influence offending (Pyrooz, 2013), given what we now know about processes of gang embeddedness and gang selection? (Densley, 2012; Pyrooz & Densley, 2015;

Pyrooz, Sweeten, & Piquero, 2013) Whether further refinement to the curriculum and delivery would improve the likelihood of detecting significant change, or whether a more refined RCT would aid in this are important questions for future research, with potential implications for the direction of both programs.

References

- Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72, 1–19.
- Alleyne, E. & Wood, J. (2012). Gang membership: The psychological evidence. In F-A Esbensen & C. Maxson (Eds.), *Youth gangs in international perspective: Results from the Eurogang program of research* (pp. 151–168). New York: Springer.
- Aos, S. & Drake, E. (2013). Prison, police, and programs: Evidence-based options that reduce crime and save money. (Doc. No. 13-11-1901). Olympia: WSIPP.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71, 364.
- Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104
- Curry, G. D. (2000). Self-reported gang involvement and officially recorded delinquency. *Criminology* 38, 1253–74.
- Decker, S. H. & Curry, G. D. (2003). Suppression without prevention, prevention without suppression: Gang intervention in St. Louis. In S. H. Decker (Ed.), *Policing Gangs and Youth Violence* (pp. 191–213). Belmont, CA: Wadsworth.
- Decker, S. H., Melde, C., & Pyrooz, D. (2013). What do we know about gangs and gang members and where do we go from here? *Justice Quarterly*, 30, 369–402.
- Decker, S. H., Pyrooz, D. C., Sweeten, G., & Moule, R. Jr. (2014). Validating self-nomination in gang research: Assessing differences in gang embeddedness across non-, current, and former gang members. *Journal of Quantitative Criminology*, 30, 577–98.

- Densley, J. (2011). Ganging up on gangs: Why the gangs intervention industry needs an intervention. *The British Journal of Forensic Practice*, 13, 12–23.
- Densley, J. (2012). Street gang recruitment: Signaling, screening and selection. *Social Problems*, 59, 301–321.
- Densley, J. (2013). *How gangs work: An ethnography of youth violence*. New York: Palgrave Macmillan.
- Densley, J. (2014). It's gang life, but not as we know it: The evolution of gang business. *Crime & Delinquency*, 60, 517–546.
- Densley, J., Davis, A., & Mason, N. (2013). Girls and gangs: Preventing multiple perpetrator rape. In M. Horvath & J. Woodhams (Eds.), *Handbook on the study of multiple perpetrator rape* (pp. 255–281). London: Routledge.
- Densley, J. & Stevens, A. (2015). “We’ll show you gang”: The subterranean structuration of gang life in London. *Criminology & Criminal Justice*, 15, 102–120.
- Deuchar, R. (2013). *Policing youth violence: Transatlantic connections*. London: IOE Press.
- Disley, E. & Liddle, M. (2016). *Local perspectives in Ending Gang and Youth Violence Areas: Perceptions of the nature of urban street gangs*. London: Home Office.
- Dunn, T. J., Baguley, T., & Brunsdon, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105, 399–412.
- Esbensen, F.A., Deschenes, E., Vogel, R. West, J. Arboit, K. & Harris, L. (1996). Active parental consent in school-based research: An examination of ethical and methodological issues. *Evaluation Review*. 20, 737–53.
- Esbensen, F. A., & Matsuda, K. (2013). Program evaluation: How do we know if we are preventing gang membership?. In T. Simon, N., Ritter, & R. Mahendra, (2013). *Changing course: Preventing gang membership* (pp. 151–161). Washington DC: National Institute of Justice and Centers for Disease Control and Prevention.

- Esbensen F.A., Miller, M., Taylor T., He N., Freng, A. (1999). Differential attrition rates and active parental consent. *Evaluation Review*, 23, 316–335.
- Esbensen, F.A, & Osgood, D. W. (1999). Gangs Resistance Education and Training (GREAT): results from the national evaluation. *Journal of Research in Crime and Delinquency*, 36, 194–225.
- Esbensen, F.A, Osgood, D. W., Peterson, D., Taylor, T., & Carson, D. (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, F. A., Peterson, D., Taylor, T. J., & Osgood, D. W. (2012). Results from a multi-site evaluation of the G.R.E.A.T. program. *Justice Quarterly*, 29, 125-151.
- Esbensen, F. A., Peterson, D., Taylor, T., Freng, A., Osgood, D. W., Carson, D., & Matsuda, K. (2011). Evaluation and evolution of the Gang Resistance Education and Training (GREAT) program. *Journal of School Violence*, 10, 53–70.
- Esbensen, F. A., Winfree, L. T., He, N., & Taylor, T. (2001). Youth gangs and definitional issues: When is a gang a gang, and why does it matter? *Crime & Delinquency*, 47, 105–30.
- Gravel, J., Bouchard, M., Descormiers, J., Wong, J., & Morselli, C. (2013). Keeping promises: A systematic review and new classification of gang control strategies. *Journal of Crime Justice*, 41, 228–241.
- Hallsworth, S. (2013). *The street gang and beyond: interpreting violent street worlds*. Basingstoke: Palgrave Macmillan.
- Hagedorn, J. M. (2005). The global impact of gangs. *Journal of Contemporary Criminal Justice*, 21, 153–169.
- Harding, S. (2014). *The street casino: Survival in violent street gangs*. Bristol, UK: Policy Press.

- Harris, D, Turner, R., Garrett, I. & Aitkinson, S. (2011). *Understanding the Psychology of Gang Violence: Implications for designing effective violence interventions*. London: Ministry of Justice.
- Haymoz, S., Cheryl Maxson, C., & Killias, M. (2014). Street gang participation in Europe: A comparison of correlates. *European Journal of Criminology*, 11, 659–681.
- Hayward, T., Simons, C., St. John, W., Waymire, M., & Stucky, T. (2011). Impacting the problem of inner-city youth violence: “Educating Kids about Gun Violence” program. *The American Surgeon*, 77, 451–457.
- Heckman, J. (1997). Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *Journal of Human Resources*, 32, 441–462.
- Heckman, J. J., Smith, J., & Clements, N. (1997). Making the most out of program evaluations and social experiments: Accounting for heterogeneity in program impacts. *The Review of Economic Studies*, 64, 487–535.
- Hennigan, K. M., Kolnick, K. A., Vindel, F. & Maxson, C. L. (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.
- Higginson, A., Benier, K., Shenderovich, Y., Bedford, L., Mazerolle, L., & Murray, J. (2014). *Predictors of youth gang membership in low-and middle-income countries: A systematic review*. Oslo, Norway: Campbell Collaboration.
- HM Government (2011), *Ending Gang and Youth Violence: A cross-government report presented to Parliament by the Secretary of State for The Home Department by command of Her Majesty*. London: Author.
- Howell, J. C. (2013). GREAT results: Implications for PBIS in schools. *Criminology & Public Policy* 12, 413–420.
- Howell, J. C. (2015). *History of street gangs in the United States*. New York: Lexington.

- Howell, J. C., & Griffiths, E. (2015). *Gangs in America's communities*, 2nd ed. Thousand Oaks, CA: Sage.
- Hubbard, R., & Lindsay, R.M. (2008). Why p values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology*, 18, 69–88.
- Klein, M. & Maxson, C. (2006). *Street gang patterns and policies*. New York: OUP.
- Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4, 165–224.
- Ludwig, J. (2005). Better gun enforcement, less crime. *Criminology & Public Policy* 4, 677–716.
- Matsuda, K., Melde, C., Taylor, T., Freng, A., & Esbensen, F. A. (2013). Gang membership and adherence to the code of the street. *Justice Quarterly*, 30, 440–468.
- Maxson, C. (2013). Do not shoot the messenger: The utility of gang risk research in program targeting and content. *Criminology & Public Policy* 12, 421–26.
- Maybin, S. (2014, 25 March). Do the public still trust the police? *BBC*. Retrieved from <http://www.bbc.com/news/magazine-26730705>
- Melde, C., & Esbensen, F. A. (2013). Gangs and violence: Disentangling the impact of gang membership on the level and nature of offending. *Journal of Quantitative Criminology*, 29, 143-166.
- Office of Juvenile Justice and Delinquency Prevention. (2008). *Best practices to address community gang problems: OJJDP's comprehensive gang model*. Washington, DC: Author.
- Pokorny, J., Jason, L., Schoeny, M., Townsend, S., & Curie, C. (2001). Do participation rates change when active consent procedures replace passive consent? *Evaluation Review*, 25, 567-580.
- Papachristos, A. (2013). Two decades of G.R.E.A.T.: Considering the history and evaluation of one of the longest-running gang prevention programs. *Criminology & Public Policy*, 12, 367–371.

- Pitts, J. (2008) *Reluctant gangsters: The changing face of youth crime*. Cullompton, UK: Willan.
- Project Oracle. (2015). Validation against the standards. Retrieved from <http://project-oracle.com/support/for-youth-service-providers/validation-against-the-standards/>.
- Pyrooz, D. C. (2013). Gangs, criminal offending, and an inconvenient truth: Considerations for gang prevention and intervention in the lives of youth. *Criminology & Public Policy*, 12, 427–36.
- Pyrooz, D. & Densley, J. (2015). Selection into street gangs: Signaling theory, gang membership, and criminal offending. *Journal of Research in Crime and Delinquency*, ahead of print doi: 10.1177/0022427815619462.
- Pyrooz, D. C., Sweeten, G., & Piquero, A. R. (2013). Continuity and change in gang membership and gang embeddedness. *Journal of Research in Crime and Delinquency*, 50, 239–271.
- Pyrooz, D. C., Turanovic, J. J., Decker, S. H., & Wu, J. (2015). Taking Stock of the Relationship Between Gang Membership and Offending: A Meta-Analysis. *Criminal Justice and Behavior*, ahead of print doi:10.1177/0093854815605528.
- Pyrooz, D.C., & Sweeten, G. (2015). Gang membership between ages 5 and 17 years in the United States. *Journal of Adolescent Health*, 56, 414–419.
- Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P. & Bushway, S. (1998). *Preventing crime: what works, what doesn't, what's promising*. Washington D.C.: National Institute of Justice.
- Smithson, H., Monchuk, L., & Armitage, R. (2011). Gang member: Who says? Definitional and structural issues. In Esbensen & C. Maxson (Eds.), *Youth Gangs an International Perspective: Tales from the Eurogang Program of Research* (pp. 53-68). New York: Springer.
- Smithson, H., Ralphs, R. & Williams, P. (2013). Used and abused: The problematic usage of gang terminology in the UK and its implications for ethnic minority youth. *British Journal of Criminology*, 53, 113-128.

- Spergel, I. (2007). *Reducing youth gang violence: The little village project in Chicago*. Lanham, MD: AltaMira Press.
- Stewart, E., & Simons, R. (2010). Race, code of the street, and violent delinquency: A multilevel investigation of neighborhood street culture and individual norms of violence. *Criminology*, 48, 569–605.
- Tajfel, H. & Turner, J. C. (1986). The social identity theory of inter-group behavior. In S. Worchel & L. W. Austin (eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago, IL: Nelson-Hall.
- Triplett, R., & Ross, T. (1998). Developing partnership for gang intervention: The role for community corrections, *Perspectives*, 22, 4.
- Thornberry, T. & Krohn, M. (2000). The self-report method for measuring delinquency and crime. In D. Duffee (Ed.), *Measurement and analysis of crime and justice* (pp. 33–83). Washington DC: U.S. Department of Justice.
- Washington State Institute for Public Policy (n.d.). WSIPP’s decision tree for evidence-based, research-based, and promising practices inventories. Retrieved from http://www.wsipp.wa.gov/ReportFile/1594/Wsipp_Updated-Inventory-of-Evidence-Based-Research-Based-and-Promising-Practices-for-Prevention-and-Intervention-Services-for-Children-and-Juveniles-in-Child-Welfare-Juvenile-Justice-and-Mental-Health-Systems_Inventory-flowchart.pdf.
- Wyrick, P. (2006). Gang prevention: How to make the “front end” of your anti-gang effort work. *United States Attorneys’ Bulletin* 54, 52–60.

Table 2. Difference in Differences Model

	Before (<i>period=0</i>)	After (<i>period=1</i>)	Change
Control group (<i>intervention=0</i>)	\bar{y}_{00}, n_{00}	\bar{y}_{01}, n_{01}	$\bar{y}_{01} - \bar{y}_{00}$
Intervention group (<i>intervention=1</i>)	\bar{y}_{10}, n_{10}	\bar{y}_{11}, n_{11}	$\bar{y}_{11} - \bar{y}_{10}$
Difference	$\bar{y}_{10} - \bar{y}_{00}$	$\bar{y}_{11} - \bar{y}_{01}$	$(\bar{y}_{11} - \bar{y}_{10}) - (\bar{y}_{01} - \bar{y}_{00})$

Table 1. Sample Characteristics

	Wave 1			Wave 2			Wave 3		
	All N=391	Control N=193	Intervene N=198	All N=361	Control N=168	Intervene N=193	All N=251	Control N=67	Intervene N=184
<i>Gender</i>		p=0.126			p=0.207			p=0.826	
- Male	59.8%	63.7%	56.1%	59.0%	62.5%	56.0%	58.6%	59.7%	58.2%
- Female	40.2%	36.3%	43.9%	41.0%	37.5%	44.0%	41.4%	40.3%	41.8%
<i>Age</i>		p=0.244			p=0.272			p=0.01**	
- 12	33.5%	36.0%	31.1%	32.7%	36.9%	29.0%	6.8%	12.3%	4.9%
- 13	66.0%	64.0%	67.9%	66.2%	61.9%	69.9%	55.0%	63.1%	52.2%
- 14	0.5%	0.0%	1.0%	1.1%	1.2%	1.0%	38.2%	24.6%	42.9%
<i>Ethnicity</i>		p=0.186			P=0.680			p=0.239	
- White	28.4%	27.1%	29.9%	28.5%	25.6%	31.1%	31.6%	29.9%	32.2%
- Mixed	14.1%	16.7%	11.7%	11.9%	13.1%	10.9%	15.2%	19.4%	13.7%
- Asian	22.0%	25.5%	18.8%	20.2%	22.6%	18.1%	12.8%	6.0%	15.3%
- Black	30.7%	27.1%	34.5%	33.0%	32.7%	33.2%	35.2%	37.3%	34.4%
- Others	4.9%	3.6%	5.1%	6.4%	6.0%	6.7%	5.2%	7.5%	4.4%
<i>Grades</i>		p=0.916			p=0.572			p=0.772	
- A student	21.7%	23.1%	21.9%	26.9%	24.4%	29.0%	22.2%	20.0%	23.0%
- B student	44.0%	45.1%	45.9%	39.3%	44.0%	35.2%	32.5%	36.7%	31.1%
- C student	24.6%	24.7%	26.0%	24.7%	23.8%	25.4%	32.9%	30.0%	33.9%
- D student	4.9%	4.9%	5.1%	5.0%	3.6%	6.2%	7.8%	6.7%	8.2%
- Other	4.9%	2.2%	1.0%	4.2%	4.2%	4.1%	0.4%	0.0%	0.5%
<i>Gang member</i>		p=0.069*			p=0.055*			p=0.400	
- Yes	3.1%	4.8%	1.5%	3.9%	6.0%	2.1%	5.7%	7.8%	5.0%

Note: p * < .10 (tending towards significance) ** < .05.

Table 3. Difference in Differences Post-Test and One-Year Post-Program Effects

	Pre-test to Post-test Analysis				Pre-test to One Year After Analysis			
	Effect size	Std. Err.	t	p	Effect size	Std. Err.	t	p
Attitudinal and perceptual measures								
School commitment	0.074	0.095	0.790	0.431	0.027	0.116	0.230	0.817
<i>Attitudes to police</i>	<i>0.243</i>	<i>0.142</i>	<i>1.710</i>	<i>0.088*</i>	<i>0.467</i>	<i>0.167</i>	<i>2.790</i>	<i>0.005**</i>
Attitudes about gangs	0.226	0.159	1.420	0.155	-0.097	0.192	-0.510	0.611
<i>Adherence to street code</i>	<i>-0.242</i>	<i>0.134</i>	<i>-1.810</i>	<i>0.070*</i>	<i>-0.541</i>	<i>0.163</i>	<i>-3.320</i>	<i>0.001**</i>
Refusal skills	-0.345	0.831	-0.410	0.679	0.277	1.055	0.260	0.793
Resistance to peer pressure	0.057	0.081	0.700	0.482	-0.142	0.106	-1.340	0.182
Conflict Resolution	1.088	0.792	1.370	0.170	-0.231	0.999	-0.230	0.817
Behavioral measures								
Delinquency (frequency)	-1.427	2.083	-0.690	0.493	-4.277	3.369	-1.270	0.205
Delinquency (variety)	0.074	0.332	0.220	0.823	-0.382	0.464	-0.820	0.410
Violent offending (frequency)	-0.549	0.487	-1.130	0.260	-0.879	0.799	-1.100	0.272
Violent offending (variety)	-0.091	0.081	-1.110	0.266	-0.113	0.116	-0.980	0.327
Gang membership	-6.88%	0.904	0.07	0.941	-2.72%	0.116	0.230	0.817

Note: The pre-test and one year after DID regression results include all available data from pre-test and one year after surveys. Considering the control data from schools 1 and 3 were missing, we reanalyzed the data omitting schools 1 and 3 entirely in an effort to test the robustness of the analysis results. The results were not substantively different. The p value for item “attitude to police” shrunk to 0.136 and the effect size for that category changed to 0.301. As such, we continue to keep the dataset intact.

p * < .10 (tending towards significance) ** < .05.