

Understanding Adverse Effects in Gang-Focused Interventions: A Critical Review

Miriam Rubenson, M.A., Katharine Galbraith, B.S., and Stanley J. Huey, Jr., Ph.D.

University of Southern California

## Abstract

Programs that aim to reduce gang involvement and violence can unintentionally produce adverse outcomes. A recent systematic review identified 41 controlled evaluations of gang-focused interventions, eight of which produced statistically significant adverse effects (i.e., effects that favored control groups). Understanding what caused these effects and whether they indicate that programs actually harmed participants requires careful investigation. In this critical review, we provide an overview of how gang-focused interventions can yield adverse effects and specify when effects are more likely to indicate harmful programs. We then critically review four program evaluations with adverse effects that exemplify different implementation problems (i.e., *implementation failure*), faulty theories about behavior change (i.e., *theory failure*), and evaluation methods that did not adequately measure outcomes (i.e., *measurement failure*). We offer hypotheses about what may have caused these adverse effects and conclude by recommending ways to maximize confidence that measured outcomes reflect real intervention effects, rather than artifacts of research design, using both time-tested methods and new technologies.

*Keywords:* gang, intervention, adverse effect, antisocial behavior, controlled evaluation

### Understanding Adverse Effects in Gang-Focused Interventions: A Critical Review

Programs that aim to reduce gang involvement and violence can sometimes unintentionally make problems worse. Developers such as police, psychologists, and criminologists design interventions that try to steer youth away from gangs or reduce recidivism in adult gang offenders. The interventions may make intuitive sense, but when rigorously tested, sometimes people who received the interventions fare worse—with more arrests, say—than those in untreated control groups. Determining what caused adverse effects, or effects that favor control groups, requires careful investigation.

In recent years, scholars have argued for the importance of examining harmful effects in crime-prevention programs (Barnett & Howard, 2018; McCord, 2003; Welsh & Rocque, 2014). Several strategies for gang and non-gang offenders, including boot camps (Wilson, Mackenzie, & Mitchell, 2008) and direct street outreach to gangs (Braga, 2016; Wilson & Chermak, 2011), are associated with unintended increases in recidivism, shootings, and other negative outcomes. However, scholars have noted that not all adverse effects indicate that programs are harmful, since effects favoring control groups can emerge for complex reasons (Ekblom & Pease, 1995; Welsh & Rocque, 2014).

In this chapter, we present a review of controlled trials of gang-focused interventions that led to statistically significant adverse effects for individuals. We aim to explore what caused adverse effects in these studies, and whether they indicate that the programs were actually harmful. While Braga (2016) reviewed one type of gang-focused intervention (so-called “streetworker” programs), there have been no published reviews of the broader range of gang programs that produced adverse effects for individuals. This chapter fills a gap in the literature by exploring what caused adverse effects across a diverse array of gang-focused interventions.

### **What causes adverse effects?**

There are several ways interventions for people involved in crimes can go awry. Some programs are based on a flawed theory of behavior change, such that activities that program developers expect will be helpful to participants are actually harmful. This is sometimes called *theory failure* (Eckblom & Pease, 1995), and the well-known program Scared Straight is a good example (Petrosino, Turpin-Petrosino, & Finckenauer, 2000). Scared Straight is based on the theory that at-risk kids can be deterred from delinquency by visiting a prison to see what life behind bars with hardcore criminals is like. Unfortunately, youth in Scared Straight were more delinquent than controls at outcome, possibly because the experience encouraged them to commit more delinquent acts to show they were, ironically, not scared (Finckenauer, 1982). Thus, the theory that visiting prison would deter future crimes failed.

Theory failure can also occur when the target population for a program is inappropriate. For example, a residential group-based intervention may be effective for offenders who have a high risk of recidivating, but could actually be harmful for those at lower risk. For lower risk offenders, such a program might disrupt existing support networks that actually reduce their risk of reoffending (e.g., connections with prosocial friends, family, etc.), or expose them to new antisocial behaviors in an environment where peers are likely to reinforce them (Lowenkamp, Latessa, & Holsinger, 2006). This mismatch between offender needs and intervention type and intensity violates the Risk-Need-Responsivity (RNR) model, a well-supported set of principles for effective assessment and intervention in criminal justice populations (Andrews, Bonta, & Hoge, 1990). The RNR model posits that effective interventions should 1) match participants' risk of offending (with higher risk participants receiving more intensive services); 2) target criminogenic needs, or factors likely to cause criminal behavior, such as pro-criminal attitudes or

substance abuse; and 3) respond to offenders' strengths and skill deficits by using cognitive-behavioral principles (Public Safety Canada, 2007). In their review of ineffective interventions for adult offenders, Barnett and Howard (2018) argued that programs that failed to adhere to RNR principles were associated with increases in recidivism. Meta-analyses of crime-prevention programs also find those that do not adhere to RNR principles are often ineffective or harmful (Andrews & Bonta, 2010a, 2010b; Andrews, Zinger, et al., 1990).

Interventions can also fail if they are poorly implemented, even if they are grounded in good theory and RNR principles (Barnett & Howard, 2018; Welsh & Rocque, 2014). If a program's active ingredients are provided weakly or not at all, *implementation failure* may produce inconclusive or adverse effects (Ekblom & Pease, 1995). Poor monitoring and supervision of staff, high turnover, organizational mismanagement, and inconsistent collaboration with community partners may result in failures to implement particularly helpful intervention strategies, possibly leaving interventions to rely on less helpful elements. For example, a program that includes early release from incarceration and job placement could be harmful if participants are released but not successfully placed in jobs. Likewise, programs without an involved researcher, written program manuals, and regular clinical supervision of providers tend to be less effective than those with such characteristics (Andrews & Dowden, 2005). When a program produces an adverse effect, in the absence of clear theory failure or another plausible alternative explanation, and especially when other trials of the same intervention have been successful, implementation failure is a distinct possibility (Welsh & Rocque, 2014).

Theory and implementation failure may result in harmful effects, but other problems with research design and measurement validity can produce adverse effects that are less easily

interpreted. *Measurement failure* occurs when evaluation methods do not adequately detect an intervention's true effects (Ekblom & Pease, 1995). The likelihood of measurement failure increases if research methods change the way outcomes are detected across treatment and control groups. For example, if treated participants have more opportunity to offend than controls because they are incarcerated for less time, higher rates of offending in one group may reflect that time difference, confounding intervention effects. Likewise, if interview questions are asked differently for each group (e.g., Williams, Cohen, & Curry, 1999), or if differences between treatment and control groups at baseline create significant confounds (e.g., Spergel, Wa, & Sosa, 2005b), outcome differences between groups may be due to measurement failure, rather than the intervention.

One example of possible measurement failure involves using the average number of participant arrests as an outcome – a common index of recidivism in gang and crime prevention programs (e.g., Braga, Piehl, & Hureau, 2009; Spergel, Wa, & Sosa, 2006). Arrests for a given individual may not be independent events, since an arrest can result in incarceration, incapacitating the individual for significant portions of a study follow-up period (Bhati & Piquero, 2007). Criminally active probationers are often reincarcerated (Petersilia, 1990), so if number of arrests is used to measure program effects on recidivism, evaluators should account for the individual's time in the community (Petersilia & Turner, 1993), but often do not. Incarceration may be less likely following minor offenses than severe offenses, such that numbers of arrests could paradoxically increase for offenders who commit less severe crimes. In sum, number of arrests can be challenging to interpret as an indicator of program effectiveness or harm. In contrast, measures such as the proportion of study participants rearrested, self-reported offenses, or time to first re-arrest do not suffer the same validity problems.

Clearly, adverse effects can emerge for several reasons, but determining what caused an effect in any given study can be challenging. Our recent systematic review found that the presence of law enforcement officers as gang intervention providers increased the risk of adverse effects, possibly because police and probation officers *detect* more offending among intervention participants than untreated controls (Rubenson, Galbraith, Shin, Beam, & Huey, 2020). Thus, higher crime detection may lead to adverse effects in a program evaluation, even if an intervention is neutral or beneficial (Hyatt & Barnes, 2017; Petersilia & Turner, 1993).

### **The Present Study**

This chapter complements the findings in Rubenson et al. (2019) by critically reviewing four representative gang interventions with theoretical, methodological, and implementation failures that may have contributed to adverse effects. The studies are diverse in theoretical orientation, target population, and research design. Below we review their core theories, intervention implementation, and research methods, then hypothesize about what caused adverse effects and whether the interventions were harmful. We conclude with recommendations for increasing confidence that detected effects are real, and discuss how new technologies and social media affect data collection for intervention studies, with potential benefits and challenges for interpreting program effects.

### **Method**

The present study grew out of a review and meta-analysis of controlled evaluations of gang-focused interventions (Huey, Lewine, & Rubenson, 2016; Rubenson et al., 2020). We conducted a literature search as described in Huey et al. (2016) using online databases and references listed in gang-focused meta-analyses and reviews (e.g., Office of Juvenile Justice and Delinquency Prevention, 2000; Wong, Gravel, Bouchard, Morselli, & Descormiers, 2012).

Eligible studies included 1) predominantly gang-involved participants, or separate outcome data for gang-involved participants; 2) a control or comparison group; 3) assessment of gang involvement or antisocial behavior as intervention outcomes; and 4) posttreatment or follow-up outcomes for individual participants. We reviewed these for studies that reported at least one statistically significant effect favoring individuals in the control group for antisocial behavior or gang involvement (i.e., an adverse effect). We excluded studies that only reported outcomes at the neighborhood or community level, rather than for individuals, or if they were not written in English. Although some programs measured secondary treatment targets, such as changes in educational attainment or prosocial skills, we limit our discussion to outcomes related to antisocial behavior and gang involvement.

### **Results**

The literature search yielded 41 studies, eight of which produced one or more adverse effects (Agopian, 1990; Peters et al., 1996; Spergel, Wa, & Sosa, 2002; Spergel et al., 2005d; Spergel, Wa, & Sosa, 2005c; Wiebush, Wagner, McNulty, Wang, & Le, 2005; Willman & Snortum, 1982; Wodarski, Filipczak, McCombs, Koustenis, & Rusilko, 1979). We review four studies that exemplify a range of intervention strategies, research methods, and theories that may have contributed to adverse outcomes.<sup>1</sup> Table 1 provides brief descriptions of the four evaluations reviewed, including sample, content, and primary outcomes.

#### **Intensive Supervision Probation: Los Angeles, CA (1988-1989)**

Agopian (1990) evaluated the Los Angeles intensive supervision probation (ISP) program compared to regular probation in a retrospective quasi-experiment using matched controls. ISP is

---

<sup>1</sup> We chose not to review one study where adverse effects were likely due to a statistical error (Wodarski et al., 1979), one that provided no details about program implementation (Willman & Snortum, 1982), and we review only one of three trials of the same Spergel model program (Spergel et al., 2006) with adverse effects.



a community-based alternative to incarceration or regular probation designed to be “restrictive” and “invasive” (Hyatt & Barnes, 2017; Petersilia, 1990), that is still used throughout the United States (e.g., County of San Mateo Probation, 2019; Nevada Department of Public Safety, 2019; San Diego County, 2019). Probationers on ISP are supposed to have more frequent contacts with probation officers, drug testing, and home visits compared to standard probation (Hyatt & Barnes, 2017). The goals of the Los Angeles ISP were to “control high-risk offenders, reduce recidivism, and quickly return program violators to court” (p. 215). One hundred fifty-eight mostly male, repeat offenders were selected from the South-Central Gang Unit to participate in the study. Overall, there were no beneficial effects on recidivism, and at six-month follow-up, significantly more ISP participants were sentenced to prison than controls (24% compared to 4%, respectively).

**Was there theory failure?** Yes. ISP is based on deterrence theory, in the sense that “increased surveillance will act as a constraint on the probationer and the likelihood of detection will act as a deterrent to crime” (Petersilia & Turner, 1991; p. 651). Intensively supervising clients clearly succeeded in returning violators to court, since more ISP clients in Agopian’s (1990) evaluation were sent to prison. However, intensive supervision does not appear to deter new crimes; rather, it may simply increase the likelihood that law enforcement will detect violations and new crimes (Hyatt & Barnes, 2017; Mackenzie & De Li, 2002; Turner & Petersilia, 1992). Other ISP evaluations also detect a high rate of probation violations and find that the program increases recidivism, costs more, and is extremely aversive to probationers (e.g., Hyatt & Barnes, 2017; Petersilia, 1990). Likewise, systematic reviews of crime prevention programs find that deterrence-based programs that rely on surveillance without rehabilitative

support can increase recidivism (Barnett & Howard, 2018; Washington State Institute for Public Policy, 2007).

Additionally, the Los Angeles ISP did not appear to adhere to the RNR model. While it targeted high-risk offenders with multiple priors, it was not clearly responsive to criminogenic needs. Most ISP clients (> 90%) were ordered to refrain from contact with specified persons, observe a curfew, and submit to drug testing, electronic surveillance, and random searches. Far fewer, however, were ordered to drug treatment (38%), alcohol treatment (3%), educational programs (9%), psychological counseling (4%), or “antibuse” (*sic*) programs (1%). Some clients were also ordered to other potentially needs-responsive activities (i.e., to maintain or seek employment, perform community service, or abstain from alcohol). However, based on Agopian’s (1990) description, it is unclear whether or not these services were high-quality, nor whether ISP clients actually needed or received the described services.

**Was there implementation failure?** Yes. Although more ISP clients than controls were *ordered* to restrictions and intensive supervision, only drug testing was more intensive for ISP clients. Otherwise, the control group actually received slightly more contacts with probation officers than did ISP clients during the last six months of supervision, though it is unclear whether this difference was significant. Although ISP clients were supposed to receive more frequent electronic monitoring, home searches, and curfews, rates were similar for control clients. Implementation failure essentially made ISP similar to standard probation.

**Was there measurement failure?** Possibly. Two problems with study design and measurement may have contributed to adverse effects. First, although the study was designed with a retrospectively matched comparison group, significant differences between the groups may have skewed recidivism results to favor controls. In general, the ISP group had more

extensive criminal histories at baseline; a greater proportion had prior arrest histories, felony convictions, and misdemeanor convictions, all of which can predict subsequent arrests during program evaluations (e.g., Spengel et al., 2006).

Second, increased drug testing may have made detecting probation violations more likely for ISP clients than for controls, regardless of rates of actual drug offenses. ISP clients were essentially monitored more frequently, increasing their chances of being caught violating probation terms, so it is not surprising that more of them were sentenced to prison at follow-up. It is unknown whether the adverse effect in Agopian (1990) reflects an increase in offending or simply increased monitoring of one group.

**Conclusions.** Theory failure may have caused the adverse effect in Agopian's (1990) evaluation of ISP, although measurement failure may have contributed as well. Implementation failure seems unlikely to have caused the adverse effects, since it made ISP more similar to standard probation. Surveillance and supervision programs without real commitments to rehabilitative services tend to increase or fail to reduce recidivism in the broader crime prevention literature (Barnett & Howard, 2018), and this program for gang members found similar results. Although supervision in this evaluation was less intense than intended, ISP participants still received more drug testing than controls, giving law enforcement more opportunities to detect violations among ISP participants, which may have contributed to the adverse effect (Agopian, 1990). Baseline differences favoring controls may also have contributed.

It is unclear whether ISP made participants more likely to reoffend than control participants, and other ISP studies suggest the model is at least unhelpful, if not harmful. However, the result of spending more time in prison arguably was harmful in the long term,

since incarceration removes participants from families, employment opportunities, and prosocial bonds (Golembeski & Fullilove, 2008).

### **Intensive Aftercare Program: Las Vegas, NV (1995-2000)**

The Intensive Aftercare Program (IAP; Wiebush, Wagner, McNulty, Wang, & Le, 2005) served incarcerated juveniles in a special custodial unit, where they provided life skills instruction, case management, graduated rewards and sanctions, and then intensive supervision and community-based services following release. Wiebush and colleagues (2005) conducted three controlled evaluations of IAP in different cities; we focus our review on the Las Vegas site because it served predominantly gang members.<sup>2</sup> Two hundred forty-seven ethnically diverse male juvenile offenders at high-risk for reoffending were randomly assigned to Las Vegas IAP or traditional services. At the 12-month follow-up, there were no significant beneficial effects of IAP. However, IAP youth had significantly more technical violations of parole than controls, and a greater proportion of IAP youth showed major institutional misconduct.

**Was there theory failure?** Possibly. IAP is based on strain, social learning, and social control theories.<sup>3</sup> Together, these suggest that delinquency arises out of weak social controls, inadequate social opportunities and resources, and negative peer influences (Agnew, 1992; Akers & Jennings, 2016; Hirschi, 2017). Based on these theories, Wiebush and colleagues (2005) suggest that effective intervention requires several strategies: intensive supervision, social services, and a structured and gradual transition back into the community (Wiebush et al., 2005).

Intensive supervision may be an attempt to respond to weak social controls and negative peer

---

<sup>2</sup> Less than half the youth at the other two sites were identified as gang members.

<sup>3</sup> Strain theory posits that society pressures individuals to achieve socially accepted goals (e.g. buying a home). Individuals who lack means to achieve such goals experience strain, and commit crimes to gain financial security so they can achieve these goals (Agnew, 1992). Social learning theory posits that people learn pro-criminal attitudes and behaviors from deviant peers (Akers & Jennings, 2016). Social control theory asserts that people who feel they have a stake in legitimate society are more likely to obey the law, while those who engage in criminal behavior feel they do not have a stake in society (Hirschi, 2017).

influences, while life skills and community-based services may be responses to inadequate social opportunities and resources, although it is unclear whether these are effective responses.

It appears the IAP model adhered to RNR principles. For example, a key tenet of IAP was “careful assessment of youth needs and an appropriate matching of needs with services based on the assessment results” (pp. 70-71). IAP in Las Vegas targeted high-risk youth, used an empirically-based risk assessment tool, and attempted to match services to level of risk. Furthermore, IAP services addressed both criminogenic needs (e.g. substance use) and were tailored to respond to youth’s specific skills and deficits; for example, offering family counseling to youth who had difficulties managing family relationships.

IAP seemed to combine intensive supervision with treatment services that targeted offenders’ skill deficits. Is this an effective model for reducing recidivism? For adults, treatment-oriented intensive supervision may be beneficial (Petersilia & Turner, 1993; Washington State Institute for Public Policy, 2007). For juveniles, evidence is mixed. One systematic review of 18 evaluations of juvenile intensive probation supervision programs found no benefits overall in crime outcomes (Washington State Institute for Public Policy, 2007). A meta-analysis of juvenile intensive probation supervision and reentry/aftercare programs found contradictory results; such programs reduced *alleged* offenses (e.g., arrests, charges, police contacts, etc.) but increased *convicted* offenses (e.g., convictions, adjudications, incarcerations; Bouchard & Wong, 2018). On the other hand, one study of a more recent ISP program for juvenile gang offenders in Manitoba, Canada that included a mentorship component (not included in the aforementioned reviews) successfully reduced recidivism compared to usual probation services (Weinrath, Donatelli, & Murchison, 2016). The mixed evidence suggests it is not clear whether combining treatment services with intensive supervision for juveniles is an effective model. IAP was also

unsuccessful at the two sites with mainly non-gang offenders, producing significant adverse effects and no beneficial effects, suggesting the model may be flawed. Strain, social control, and social learning theories may be reasonable bases for intervention, but mixed outcomes from the juvenile literature and the failure of the other two IAP evaluations by Wiebush and colleagues (2005) suggest that intensive supervision plus rehabilitative services does not consistently reduce recidivism.

**Was there implementation failure?** Possibly. Wiebush and colleagues (2005) gave the Las Vegas IAP an overall rating of “moderate implementation” (p. 47). IAP successfully provided high-intensity community supervision and treatment services, especially compared to controls. All elements of community supervision and surveillance were strongly implemented; for example, parole officers had exclusively-IAP caseloads of fewer than 15, contacts with parolees were frequent, and field agents conducted weekend supervision as planned. The structured transition back into the community also adhered to plan, with formal step-downs in parole contact frequency.

However, several rehabilitative and administrative elements of the model were only moderately or weakly implemented, according to the researchers. A key element of the model involved creating “linkages with community resources and social networks,” which suffered from several barriers. For example, pre-release planning was hampered by poor family involvement and staffing vacancies, and the life skills curriculum was delivered only sporadically while youth were institutionalized. Staff turnover was “extensive,” positions responsible for coordinating community rehabilitative services were vacant, and planned pre-release trips into the community never happened. There was also “mixed support by a series of institutional administrators,” one particularly ineffective administrator “hurt the project,” and a

management team was never formed (p. 49). While supervision was reportedly intensive, treatment services seem to have been less well implemented. Since intensive supervision on its own is generally ineffective, the weak to moderate implementation of rehabilitative services may have been particularly damaging to the IAP model.

**Was there measurement failure?** Possibly. Intensive supervision itself may have increased detection of new offenses and violations in the treatment group. Parole officers in Nevada have the power to arrest and charge youth with offenses, and the authors suggest this could have contributed to the adverse effect for technical violations (Wiebush et al., 2005, p. 85). For example, parole officers can make drug and weapons charges if they see those items during a home visit. In sum, IAP participants may have been more likely to be caught offending because they were intensively supervised, compared to controls, and the intensive supervision was reportedly well-implemented.

**Conclusions.** Adverse effects in Wiebush and colleagues' (2005) Las Vegas IAP trial may have been caused by a combination of theory, measurement, and implementation failure. While IAP's intensive supervision appears to have been implemented as intended, treatment and community reentry services were less so. Wiebush and colleagues (2005) note there were difficulties effectively engaging peer and family networks in care, which may have undermined positive changes (p. 85). They speculate that community-based resources like drug treatment were not always high quality. These implementation weaknesses may have left the intervention relying primarily on supervision over rehabilitative services. Given the issues with measurement, it is not clear whether IAP increased reoffending among treated participants, or simply the rate of detecting offenses.

### **Comprehensive Community-Wide Approach to Gang Prevention, Intervention, and Suppression: Riverside, CA (1997-2000)**

The Comprehensive Community-Wide Approach to Gang Prevention, Intervention, and Suppression, known as the Spergel model, was implemented at six sites with heavy gang involvement in the United States (Spergel et al., 2006). The project targeted gang-involved youth offenders from high-crime neighborhoods, and provided individual counseling, suppression, group services, job services, school services, case planning, family counseling, and material support (Spergel et al., 2005d). “Youth outreach workers,” mainly former gang members, were the primary providers along with a mixed team that included probation and police. The project also required community grass-roots organization, with the goal of changing problematic social and economic factors that contributed to the gang problem in the communities.

Spergel and colleagues (2006) considered the trial in Riverside, CA to be one of the three “successful” implementations of the model (along with Chicago, IL, and Mesa, AZ; Spergel et al., 2003, 2002), since gang violence declined in the treatment area, relative to the comparison area, while the three other sites produced mainly null and a few adverse results (Bloomington-Normal, IL, Tucson, AZ, and San Antonio, TX; Spergel, Wa, & Sosa, 2005b, 2005a; Spergel et al., 2005c). However, the Riverside trial actually produced significant beneficial *and* adverse effects for individuals; arrests for serious violence dropped for intervention youths while they increased for comparisons, but intervention youth were also arrested more often for drug charges in the post-program period than control youth. We discuss the Riverside trial here because its mixed outcomes exemplify the Spergel model’s successes and failures.

**Was there theory failure?** Possibly. The Spergel model is based on multiple social, criminological, and psychological theories and uses several intervention strategies. The model



incorporates elements of ecological, social disorganization, anomie, opportunity, socialization, and maturational theories, since “no one theory sufficiently accounts for, nor can predict, the behavior of gang youth or provide adequate direction for policy or program development” (Spiegel, Wa, Grossman, et al., 2003; p. 8.4). Together, these theories suggest that effective intervention requires changes to the economy, neighborhood, police, peer network, and the individual. The Spiegel model sought to change problematic social and economic forces contributing to the gang problem by mobilizing law enforcement, local agencies, and citizens, and enhancing social, recreational, and occupational opportunities and treatment services for youth. The primary youth-facing strategy was direct street outreach by workers who were usually former gang members. With so many theories and intervention strategies, it is hard to tell whether there was theory failure, or to what extent the program’s strategies reasonably reflected the underpinning theories of crime prevention.

The Riverside program appeared to adhere to risk and need principles of the RNR model, but perhaps not to responsivity. An assessment team identified and planned services that targeted youth’s specific needs, and matched services to risk level, so that younger gang members and those with more serious arrest histories were provided with more services. Individual counseling was the most frequently used intervention strategy and was provided by youth outreach workers and occasionally probation officers. It is not clear what counseling entailed, but providers were probably not trained in cognitive behavioral principles in this 1990s intervention.

Klein (2011) notes that disaggregating the effects of the multiple strategies, settings, client types, and combinations thereof in comprehensive programs is challenging, although the Spiegel model programs attempt to do so. Spiegel and colleagues (2006) identified eight strategies as most important for success (city/county leadership, interagency street team

coordination, criminal justice participation, lead agency management/commitment, suppression, organizational change and development, balance of service, and targeting of appropriate youth), and three others as moderately important (community mobilization, crisis intervention/outreach, and social opportunities provision). The finding suggests that nearly everything is important, which may be true, and not very informative.

However, Braga (2016) argued that programs that rely on street outreach to gang youth as a primary intervention strategy may represent theory failure. Several so-called streetworker programs from the 2000s seemed to increase rates of shootings, homicides, and assaults (Fox, Katz, Choate, & Hedberg, 2015; J. M. Wilson & Chermak, 2011), although most evaluations only measured outcomes for communities, not for individual participants. Streetworker programs like the Spergel model may inadvertently increase gang cohesion. Streetworkers are often former gang members hired for their “street cred,” familiarity with gang life, and ability to connect well with current gang members. Klein (2011) and Braga (2016) describe how former gang members are often reluctant to share information with collaborating law enforcement, are arrested for committing crimes during interventions, and can share stories with youth that inadvertently glorify gang life and increase youths’ identification as gang members.

Braga (2016) notes that many streetworker programs from the 2000s were proclaimed “successes” even though they produced mixed effects. This seems to be the case with the Spergel model programs as well. Two other trials of the Spergel model produced mixed outcomes as well, reporting both significant beneficial and adverse effects (Spergel et al., 2002, 2005c), which could be evidence of theory failure.

**Was there implementation failure?** Yes. Klein and Maxson (2006) suggest that the Spergel programs were so complex that every element provided another opportunity for

implementation failure. Indeed, implementation strength at Riverside varied depending on the intervention strategy, with some significant weaknesses. Spergel and colleagues (2005) used progress reports, funding applications, and site visits to track service use and treatment fidelity throughout the intervention. They found that the Riverside project “did not begin to substantially meet the requirements” of the model in terms of community mobilization and organization until leadership changed in the latter half of the project (Spergel et al., 2005, p. 6.1).

Problems with the youth outreach workers’ performance were also apparent throughout the project. In addition to high turnover, outreach workers appeared to focus on less-delinquent youth who were more receptive to their services, rather than on the hardcore delinquent youth they were meant to target. Some avoided working at nights and in higher-risk areas, the very times and places they were needed most. Similar to other streetworker programs, outreach workers were also reluctant to share information about gangs with collaborating law enforcement.

However, other aspects of the Spergel model were reportedly well implemented, like employment training and suppression. Youth were placed in jobs and were given training by the City of Riverside, and the police ultimately collaborated with other agencies and helped mediate gang conflict situations.

**Was there measurement failure?** Yes. Measurement failure may have occurred in two ways: baseline differences between treatment and comparison areas and using number of arrests as an outcome measure may have skewed results. First, comparison youth for the Riverside evaluation were less delinquent and had a greater proportion of females than treatment youth at baseline, differences which may have skewed outcomes to favor controls at post-treatment and

follow-up. There were also different police arrest practices in the two communities, which may have affected the rates of detecting crimes across the two areas.

Second, a primary outcome was change over time across groups for number of arrests. Although Spergel and colleagues (2005) intended to measure self-reported delinquency, they were unable to obtain that data for many participants at follow-up, so they had to rely on official arrest data. It is possible that the adverse effect for drug arrests emerged in part due to issues described previously with measuring arrest frequency within a study period.

**Conclusions.** A combination of measurement, implementation, and theory failure may have contributed to the adverse effect seen in the Riverside Spergel program. Measurement validity may have suffered from significant differences between treatment and comparison groups' levels of prior offending, and implementation issues in the first year may have damaged the project. Theory may have been problematic as well since direct outreach to gangs has been associated with adverse effects in many interventions, including three of the six Spergel programs.

The Riverside evaluation corroborates Braga's (2016) observation that contemporary streetworker programs with mixed outcomes are sometimes declared "successes" (Braga, 2016; Spergel et al., 2006). While Braga attributes adverse effects in streetworker programs to theory failure, it seems difficult to distinguish between theory and implementation failure, when implementation fails repeatedly in streetworker programs. Do problems with worker effectiveness and interagency collaboration, common to many streetworker programs, indicate theory or implementation failure? Is the theory good if half the trials produce beneficial effects, but half also produce adverse effects? Was it helpful, harmful, or both? It is impossible to guess whether a more perfect implementation would have yielded fewer adverse effects since at no site

were all core elements of the model strongly implemented (Spergel et al., 2006). Although Spergel (2006) regarded Riverside as a success overall, the mixed outcomes, major implementation issues, measurement validity problems, and potential for theory failure due to street outreach suggest that any claims that the program was either successful or harmful are at best too simplistic.

### **Environmental Youth Corps: Mobile, AL (1992-1993)**

Peters and colleagues (1996) evaluated the Environmental Youth Corp (EYC) in Mobile, Alabama, a 90-day military-style boot camp, compared to standard probation or residential placement for nonviolent youth offenders. The sample was all male, mostly African American, about half gang-involved, and considered high risk for reoffending. EYC focused on discipline, educational remediation, and community service, and included six to nine months of reduced supervision aftercare. The evaluation found no overall differences in recidivism between EYC and control youth at six months; however, EYC youth were significantly faster to recidivate.

**Was there theory failure?** Yes. It is unclear whether any criminological, sociological, or psychological theory underpinned the intervention strategy. The authors argue that “factors contributing to a delinquent lifestyle can be overcome through military-style discipline and structure, life skills training, educational remediation, and community service” (p. iii), and suggest that completing the mentally and physically challenging boot camp “was to be regarded as a significant personal victory” for youth (p. I-3).

Welsh and Rocque (2014) argue that boot camps may suffer from theory failure, since there is no evidence that discipline reduces recidivism. Similarly, a systematic review of 32 military-style boot camps found no overall differences in recidivism between boot camp participants and controls, with some studies finding better outcomes for controls and some for

boot camp participants (Wilson et al., 2008). MacKenzie and Souryal (1994) suggested that boot camps may be harmful for the same reason as Scared Straight; the boot camp is meant to be a physically and emotionally unpleasant experience that deters future offending, but may backfire by provoking participants to prove they are “tough,” perhaps by reoffending (National Institute of Justice, 1994).

EYC appears to have attempted to target youths’ needs, but may not have provided appropriately responsive services. Youth were extensively assessed for criminal and social history, substance use, home life, school performance, employment and life skills, and gang involvement, then given an individualized treatment plan to address their specific needs. The plans were to be revisited regularly and maintained through aftercare. However, there does not appear to have been any use of cognitive-behavioral principles in response to needs assessments, and the program targeted nonviolent offenders using a custodial intervention with the same length of stay for all participants, which may not have matched participants’ level of risk (Lowenkamp et al., 2006).

**Was there implementation failure?** Probably not. EYC appears to have been largely implemented as intended, providing a 90-day military-style boot camp focused on discipline and education remediation. Although there were implementation weaknesses, these were not related to discipline, the core intervention strategy. For example, three-month reviews of individual treatment plans did not always occur as planned, the life skills coordinator did not use individual treatment plans to guide development of the life skills curriculum, and drill instructors did not feel they had input into treatment plan development (p. II-22). Program coordinators were “fairly successful” in involving family members of EYC youth, but many parents were unable to attend

the family aftercare program (p. II-22). EYC also suffered from high staff turnover due to poor pay and burnout.

**Was there measurement failure?** No. Peters and colleagues (1996) used a randomized experimental design, and no significant comparability problems emerged between groups. They measured EYC's effectiveness using youths' time to recidivate, defined as the number of days before youths' first court-adjudicated new offense or probation violation, with EYC youth recidivating significantly faster. There were no significant group differences in severity of new adjudicated offenses or violations.

**Conclusions.** Adverse effects in the Peters et al. (1996) EYC evaluation most likely reflect theory failure. The boot camp strategy seems to have relied on assumptions that discipline would reduce recidivism, which evidence does not support. In the absence of measurement and major implementation problems, theory failure seems likely, and suggests EYC may have actually increased participants' recidivism.

### **Discussion**

What lessons can we draw from studying adverse effects in gang-focused interventions? By exploring these effects, we aimed to shed light on their causes, and on the challenges of determining when adverse effects suggest programs are harmful or are simply artifacts of research design and measurement validity problems. Studying programs with adverse effects may also elucidate why effects occurred so that program developers and evaluators can reduce the likelihood of harm and flawed research designs in future studies. We were inspired by the growing body of research on harmful effects in crime prevention programs, including work by McCord (2003), Welsh and Rocque (2014), Barnett and Howard (2018), Braga (2016) and others.

Based on a systematic review of 41 controlled trials of individual-based gang interventions, we found eight studies with statistically significant effects that favored control groups. We examined whether failures of theory, implementation, or measurement caused adverse effects in four of these that exemplified a range of problems, and then hypothesized about whether they actually increased participants' antisocial behavior. We found that several factors contributed to adverse effects in gang-focused interventions, with most studies suffering from multiple types of failures. We relied both on what was reported in the studies themselves and on the relevant literature about similar programs to draw conclusions about each.

Consistent with the literature on intensive supervision, two intensive supervision programs for gang offenders in our sample found adverse effects that may have been due to measurement failure (Agopian, 1990; Wiebush et al., 2005). While these programs had other shortcomings as well, they both may have produced adverse effects because treatment groups were, in some key respects, monitored more closely than controls. For the Spergel program in Riverside (Spergel et al., 2005d), implementation and measurement validity problems in the evaluation may have led to the adverse effect for drug arrests. Finally, we speculated that the adverse effect in the EYC evaluation was due to theory failure (Peters et al., 1996), given similar results from other evaluations of boot camps for non-gang offenders.

By exploring what caused adverse effects, can we determine whether programs were actually harmful? We suggest that adverse effects due to implementation and theory failure provide stronger evidence for harm than measurement failure. When measurement failure is clear, adverse effects do not indicate either way whether a program was harmful.

**Moving forward: Adverse effects in the age of the internet and social media**



Lessons learned from these interventions are relevant even as evaluation strategies change with new technologies. The internet and social media provide new ways to collect data about criminal activity (Trottier, 2012; Walsh & O'Connor, 2019), but may also increase the potential for adverse effects. A recent study of gang outreach workers' social media use illustrates both the promise and potential pitfalls of this new intervention tool. Patton and colleagues (2016) interviewed outreach workers in Chicago who described using social media to monitor gang activities, build trust with youth, and intervene in escalating online beefs (Patton, Eschmann, Elsaesser, & Bocanegra, 2016). The workers described how gang members friending them online helped enhance the closeness of their relationships. It was also helpful for suppression, since social media often provided access to logs of youths' current and past illegal behaviors. The interviews suggest social media has potential as a tool to monitor behaviors and intervention outcomes.

However, social media may also increase the risk for adverse effects due measurement failure and by potentially encouraging unnecessary intervention. The outreach workers in Patton et al. (2016) noted that interpreting youths' online activity often required "insider knowledge" (p. 594). One described the fatal shooting of a boy who posted a picture of himself on Facebook with "his hat tilted to the side," which apparently indicated his gang membership and provoked the shooting. Workers also described how youth exaggerated their criminal behaviors online, taking credit for others' violence, or posting pictures and videos with fistfuls of cash, drugs, and guns they did not actually own. Workers who knew the gangs understood which posts reflected real events, and which seemingly-innocuous posts were real threats. It is easy to imagine how researchers could misinterpret social media posts (particularly for control groups not served by

providers who know them well) and risk unnecessary or harmful intervention, measurement failure, or both.

Future research should explore how to use social media for intervention evaluations while minimizing the potential for adverse effects. Questions remain about how to optimize the internet for good data collection: can viewing gangs' threatening videos for data collection socially reinforce youth by increasing the number of views (Storrod & Densley, 2017)? Can outreach workers erroneously intervene with youth on the street based on misleading online posts (Patton et al., 2016; Walsh & O'Connor, 2019)? Does supervising online gang behavior increase the likelihood of measurement failure? Future research should carefully consider these unknowns involved in using new technologies in interventions.

**Recommendations.** Determining which intervention effects are true effects in controlled studies can be challenging, and several kinds of evidence are needed. Here we provide a set of recommendations for designing evaluations that maximize confidence that measured outcomes reflect real intervention effects:

1. *Ground interventions in evidence-based theory* from psychology, sociology, criminology, or related fields, and avoid untested “common sense” strategies (Andrews & Dowden, 2005). If a theory underpins other successful interventions, adverse effects that emerge may be less likely due to theory failure than to another problem. Without good scientific theory, even when a program is successful, it is more difficult to determine *why* it worked. CrimeSolutions.gov is a good resource for checking intervention effects and theory (Crimesolutions.gov, 2019; Worrall, 2015). Experts at CrimeSolutions.gov rate programs based on theoretical framework,

- outcome evidence, design quality, and program fidelity, using multiple evaluations of the same program and meta-analyses (CrimeSolutions.gov, 2019).
2. *Measure intervention fidelity.* For example, national evaluators monitored the Spergel model programs at all sites, tracking services and contacts with program youth, and interviewing providers about implementation successes and challenges. As a result, Spergel and colleagues (2006) could test whether higher-risk youth in fact received more intensive services than lower-risk youth, as the model intended. Peters and colleagues (1996) also interviewed drill instructors from the EYC boot camp, which revealed several problems with fidelity (but none that clearly impacted the core intervention strategy).
  3. *Measure outcomes similarly for treatment and control groups.* Surveillance methods and intensity should be the same for treatment and control groups in outcome evaluations, since differences can change the likelihood of detecting outcomes across groups and risk producing false-positives. Social media can improve surveillance, since youth often post video evidence of illegal behavior on public platforms like YouTube and Instagram (Patton et al., 2016; Storrod & Densley, 2017). Monitoring gang activity online in controlled evaluations will likely pose challenges, and future research should take care that frequency, intensity, and also ability to interpret online activity is similar across groups.
  4. *Use multiple measures for the same outcome constructs,* such as time to first arrest, proportion arrested, and self-reported offenses for recidivism (Mackenzie & De Li, 2002). While the number of arrests is an often-used measure of recidivism, it primarily measures justice system contact and fails to capture the majority of offenses

(Gramlich, 2019; Mackenzie & De Li, 2002). Supplementing official arrest rates with other measures may help reduce validity problems associated with relying on the number of arrests as the primary outcome.

5. *Test mediators and confounds.* If researchers specify and track services and contacts (e.g., Spergel et al., 2006), they can test whether exposure to the hypothesized active ingredients improves the likelihood of positive outcomes, gaining stronger evidence that the intervention itself was effective. Testing known confounds for delinquency, such as age, helps isolate program effects as well.

In conclusion, multiple factors contribute to adverse effects in evaluations of gang-focused interventions, and these effects do not always indicate that programs are harmful. We hope that researchers, policy makers, and practitioners can use this information when making decisions about how to address gang problems in their communities and measure program outcomes.

## References

- Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology*, 30(1), 47–88.
- Agopian, M. W. (1990). The impact of Intensive Supervision Probation on gang-drug offenders. *Criminal Justice Policy Review*, 4(3), 214–222.
- Akers, R. L., & Jennings, W. G. (2016). Social learning theory. In A. R. Piquero (Ed.), *Handbook of criminological theory* (pp. 230–240). West Sussex, United Kingdom: John Wiley & Sons.
- Andrews, D. A., & Bonta, J. (2010a). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16(1), 39–55.
- Andrews, D. A., & Bonta, J. (2010b). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson Publishing Co.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17(1), 19–52.
- Andrews, D. A., & Dowden, C. (2005). Managing correctional treatment for reduced recidivism: A meta-analytic review of programme integrity. *Legal and Criminological Psychology*, 10(2), 173–187.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28(3), 369–404.
- Barnett, G. D., & Howard, F. F. (2018). What doesn't work to reduce reoffending? A review of reviews of ineffective interventions for adults convicted of crimes. *European Psychologist*,

23(2), 111–129.

Bhati, A. S., & Piquero, A. R. (2007). Estimating the impact of incarceration on subsequent offending trajectories: Deterrent, criminogenic, or null effect? *The Journal of Criminal Law and Criminology*, 98(1), 207–253.

Bouchard, J., & Wong, J. S. (2018). Examining the effects of intensive supervision and aftercare programs for at-risk youth: A systematic review and meta-analysis. *International Journal of Offender Therapy and Comparative Criminology*, 62(6), 1509–1534.

Braga, A. A. (2016). The continued importance of measuring potentially harmful impacts of crime prevention programs: The academy of experimental criminology 2014 Joan McCord lecture. *Journal of Experimental Criminology*, 12, 1–20.

Braga, A. A., Piehl, A. M., & Hureau, D. (2009). Controlling violent offenders released to the community: An evaluation of the boston reentry initiative. *Journal of Research in Crime and Delinquency*, 46(4), 411–436.

County of San Mateo Probation. (2019). Intensive supervision. Retrieved May 27, 2019, from <https://probation.smcgov.org/intensive-supervision>

Crimesolutions.gov. (2019). How we review and rate a program from start to finish.

Eckblom, P., & Pease, K. (1995). Evaluating crime prevention. *Crime and Justice*, 19, 585–662.

Finckenauer, J. O. (1982). *Scared straight! and the panacea phenomenon*. Edgewood Cliffs, NJ: Prentice-Hall.

Fox, A. M., Katz, C. M., Choate, D., & Hedberg, E. (2015). Evaluation of the phoenix TRUCE project: A replication of Chicago CeaseFire. *Justice Quarterly*, 31, 85–115.

Golembeski, C., & Fullilove, R. (2008). Criminal (in)justice in the city and its associated health

consequences. *American Journal of Public Health*, 98, S185–S190.

Gramlich, J. (2019). 5 facts about crime in the U.S. Retrieved May 31, 2019, from

<https://www.pewresearch.org/fact-tank/2019/01/03/5-facts-about-crime-in-the-u-s/>

Hirschi, T. (2017). *Causes of delinquency*. New York, NY: Routledge.

Huey, S. J. J., Lewine, G., & Rubenson, M. (2016). A brief review and meta-analysis of gang intervention trials in North America. In C. L. Maxson & F.-A. Esbensen (Eds.), *Gang transitions and transformations in an international context* (pp. 217–233). New York: Springer.

Hyatt, J. M., & Barnes, G. C. (2017). An experimental evaluation of the impact of intensive supervision on the recidivism of high-risk probationers. *Crime and Delinquency*, 63(1), 3–38.

Lowenkamp, C. T., Latessa, E., & Holsinger, A. M. (2006). The risk principle in action: What have we learned from 13,676 offenders and 97 correctional programs? *Crime & Delinquency*, 52(1), 77–93.

Mackenzie, D. L., & De Li, S. (2002). The impact of formal and informal social controls on the criminal activities of probationers. *Journal of Research in Crime and Delinquency*, 39(3), 243–276.

McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. *Annals of the American Academy of Political and Social Science*, 587, 16–30.

National Institute of Justice. (1994). *Multisite evaluation of shock incarceration: Executive summary*. Washington, DC.

Nevada Department of Public Safety, Parole and Probation. (2019). Intensive supervision

- programs. Retrieved May 27, 2019, from [http://npp.dps.nv.gov/Home/Offender\\_Supervision/Intensive\\_Supervision\\_Programs/](http://npp.dps.nv.gov/Home/Offender_Supervision/Intensive_Supervision_Programs/)
- Office of Juvenile Justice and Delinquency Prevention (2000). *Preventing adolescent gang involvement*. Washington, DC.
- Patton, D. U., Eschmann, R. D., Elsaesser, C., & Bocanegra, E. (2016). Sticks, stones and Facebook accounts: What violence outreach workers know about social media and urban-based gang violence in Chicago. *Computers in Human Behavior*, *65*, 591–600.
- Peters, M., Bullman, S., Gimbel, C., Thomas, D., Laxton, G., Opanga, M., ... Croan, G. (1996). *Evaluation of the impact of boot camps for juvenile offenders*.
- Petersilia, J. (1990). When probation becomes more dreaded than prison. *Federal Probation*, *23*, 23–27.
- Petersilia, J., & Turner, S. (1991). An evaluation of intensive probation in California. *The Journal of Criminal Law and Criminology (1973-)*, *82*(3), 610–658.
- Petersilia, J., & Turner, S. (1993). Intensive probation and parole. *Crime and Justice*, *17*, 281–335.
- Petrosino, A., Turpin-Petrosino, C., & Finckenauer, J. O. (2000). Well-meaning programs can have harmful effects! Lessons from experiments of programs such as Scared Straight. *Crime & Delinquency*, *46*(3), 354–379.
- Public Safety Canada. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation 2007-06*.
- Rubenson, M. P., Galbraith, K., Shin, O., Beam, C. R., & Huey, S. J. J. (2020). When helping hurts? Toward a nuanced interpretation of adverse effects in gang-focused interventions.



*Clinical Psychology: Science and Practice*, 1–14. <https://doi.org/10.1111/cpsp.12321>

San Diego County Probation Department. (2019). Adult community supervision. Retrieved May 27, 2019, from

[https://www.sandiegocounty.gov/content/sdc/probation/adult\\_information\\_community\\_supervision.html](https://www.sandiegocounty.gov/content/sdc/probation/adult_information_community_supervision.html)

Spergel, I. A., Wa, K. M., Grossman, S., Sosa, R. V., Anderson, L., Arata, L., ... Lyndes, K. (2003). *The Little Village gang violence reduction project in Chicago*. Chicago, IL.

Spergel, I. A., Wa, K. M., & Sosa, R. V. (2006). The comprehensive, community-wide gang program model: Success and failure. In J. F. Short & L. A. Hughes (Eds.), *Studying youth gangs* (pp. 203–224). Lanham, MD: AltaMira Press.

Spergel, I. A., Wa, K. M., & Sosa, R. V. (2002). *Evaluation of the Mesa Gang Intervention Program (MGIP)*. Washington, DC.

Spergel, I. A., Wa, K. M., & Sosa, R. V. (2005a). *Evaluation of the San Antonio Comprehensive community-wide approach to gang prevention, intervention and suppression program*. Washington, DC.

Spergel, I. A., Wa, K. M., & Sosa, R. V. (2005b). *Evaluation of the Tucson comprehensive community-wide approach to gang prevention, intervention and suppression program*. Washington, DC.

Spergel, I. A., Wa, K. M., & Sosa, R. V. (2005c). *Evaluation of the Bloomington-Normal comprehensive gang program*. Washington, DC.

Spergel, I. A., Wa, K. M., & Sosa, R. V. (2005d). *Evaluation of the Riverside comprehensive community-wide approach to gang prevention, intervention & suppression*. Washington,

DC.

- Storrod, M. L., & Densley, J. A. (2017). 'Going viral' and 'going country': The expressive and instrumental activities of street gangs on social media. *Journal of Youth Studies, 20*(6), 677–696.
- Trottier, D. (2012). Policing social media. *Canadian Review of Sociology/Revue Canadienne de Sociologie, 49*(4), 411–425.
- Turner, S., & Petersilia, J. (1992). Focusing on high-risk parolees: An experiment to reduce commitments to the Texas Department of Corrections. *Journal of Research in Crime and Delinquency, 29*(1), 34–61.
- Walsh, J. P., & O'Connor, C. (2019). Social media and policing: A review of recent research. *Sociology Compass, 13*(1), e12648.
- Washington State Institute for Public Policy. (2007). Evidence-based public policy options to reduce future prison construction, criminal justice costs, and crime rates. *Federal Sentencing Reporter, 19*(4), 275–290.
- Weinrath, M., Donatelli, G., & Murchison, M. J. (2016). Mentorship: A missing piece to manage juvenile intensive supervision programs and youth gangs? *Canadian Journal of Criminology and Criminal Justice, 58*(3), 291–321.
- Welsh, B. C., & Rocque, M. (2014). When crime prevention harms: A review of systematic reviews. *Journal of Experimental Criminology, 10*(3), 245–266.
- Wiebush, R. G., Wagner, D., McNulty, B., Wang, Y., & Le, T. N. (2005). *Implementation and outcome evaluation of the intensive aftercare program: Final report*. Washington, DC.
- Williams, K., Cohen, M. I., & Curry, D. (1999). *Evaluation of youth gang drug*

*intervention/prevention programs for female adolescents*. Washington, DC.

- Willman, M. T., & Snortum, J. R. (1982). A police program for employment of youth gang members. *International Journal of Offender Therapy and Comparative Criminology*, 26(3), 207–214.
- Wilson, D. B., Mackenzie, D. L., & Mitchell, F. N. (2008). Effects of correctional boot camps on offending. *Campbell Systematic Reviews*, 6, 126–143.
- Wilson, J. M., & Chermak, S. (2011). Community-driven violence reduction programs. *Criminology & Public Policy*, 10(4), 993–1027.
- Wodarski, J. S., Filipczak, J., McCombs, D., Koustenis, G., & Rusilko, S. (1979). Follow-up on behavioral intervention with troublesome adolescents. *Journal of Behavior Therapy and Experimental Psychiatry*, 10(3), 181–188.
- Wong, J., Gravel, J., Bouchard, M., Morselli, C., & Descormiers, K. (2012). *Effectiveness of street gang control strategies: A systematic review and meta-analysis of evaluation studies*. Ottawa, ON.
- Worrall, J. L. (2015). Validating peer review in criminal justice evaluation research: Evidence from CrimeSolutions.gov. *Journal of Criminal Justice Education*, 26(4), 507–529.

Table

Table 1

Gang-Focused Interventions with Adverse Effects

<u>Program Name, Year, &amp; Setting</u>	<u>Sample</u>	<u>Design &amp; Intervention</u>	<u>Significant Effects</u>
Agopian (1990) Intensive Supervision Probation (ISP), Los Angeles, CA	158 adult gang offenders with histories of drug use or sales.	Retrospective quasi-experimental. Received ISP or regular probation. Groups matched by gender, education, marital status, and drug offenses.	At 6-month and 1-year follow-up, more ISP participants were serving time in jail or prison than controls.
Wiebush et al. (2005) Intensive Aftercare Program (IAP), Las Vegas, NV	247 incarcerated juvenile offenders rated “extremely high risk” for recidivism, 55% gang members.	Randomly assigned to IAP or traditional services. IAP included case management, intensive supervision with community-based services, and graduated rewards and sanctions.	A greater proportion of IAP youth had major institutional misconduct and technical violations of parole.
Spergel et al. (2005) Comprehensive Community-Wide Approach to Gang Prevention, Intervention, and Suppression, Riverside, CA	369 youth gang members (48%), associate gang members (11%) or at high risk of gang involvement (14%). 60% offenders.	Prospective quasi-experimental. Treatment group received (from most to least used) individual counseling, suppression, group services, job services, case planning, family counseling, material support, and school services. Demographically matched controls received no treatment.	Treated youth increased drug arrests, but decreased arrests for serious violence in post-program period relative to controls.
Peters et al. (1996) Environmental Youth Corps (EYC), Mobile, AL	374 non-violent adjudicated youth, categorized as “high risk” of continuing delinquency. 44% gang-involved.	Randomly assigned to EYC, a 90-day military-style boot camp focused on discipline, educational remediation, community service, and 6-9 month reduced supervision aftercare, or controls. Controls received usual probation or residential services, plus one or more services for substance abuse, education, anger management, or mental health.	EYC youth were faster to recidivate.