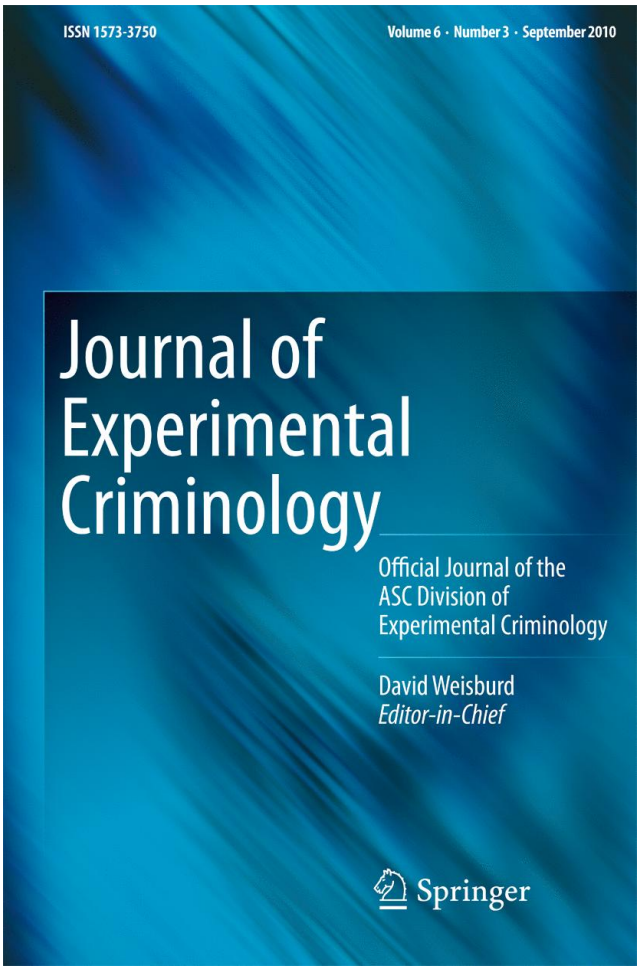
ISSN 1573-3750, Volume 6, Number 3



J Exp Criminol (2010) 6:229–243
DOI 10.1007/s11292-010-9101-9

Deviancy training: understanding how preventive interventions harm

The Academy of Experimental Criminology 2009 Joan McCord Award Lecture

Denise C. Gottfredson

Published online: 8 July 2010

Abstract Joan McCord's follow-up study of the Cambridge–Somerville Youth Project showed that even well-intentioned, well-implemented prevention programs sometimes have harmful effects on participants. She reported that peer reinforcement of delinquent behaviors or bragging about delinquent behaviors that occurred during summer camp experiences provided as part of the project might explain the negative outcomes observed for treatment boys. We explored this "deviancy training" mechanism in the context of an evaluation of an after-school program. The study found that peer deviancy training does occur in the context of after-school programs, that it is more likely to occur during less structured activities, and that more violent behavior also occurs during these less structured times. Also, the amount of adult supervision that is afforded in after-school programs did not counteract the reinforcing effect of peers. Finally, we showed that while teaching a prevention curriculum that was part of the after-school program, the most effective group leaders provided positive reinforcement for students' pro-deviancy expressions. A scale assessing beliefs that illegal, violent, and risky behaviors are common and acceptable in the peer group favored the control students in the programs in which group leaders were observed providing this positive reinforcement. Implications for prevention programming are discussed.

This work was supported in part through grant number R305F050069 from the U.S. Department of Education, Institute of Educational Sciences, to the University of Maryland. I wish to thank Amanda B. Cross, Denise M. Wilson, and Melissa Rorie, former or current graduate students in the Department of Criminology & Criminal Justice, University of Maryland, and Nadine Connell, Law and Justice Studies Department, Rowan University, whose work is summarized herein. I also thank Elise Andrews of the Baltimore County Local Management Board and Beahta Davis and the staff of the Baltimore County Dept. of Recreation and Parks for implementing and managing the after-school programs. Thanks also to Matthew Brigham, Stephanie DiPietro, Matt Gugino, Lynda Okeke, Freshta Rahimi, and Jaynie Trageser for research assistance, and to Gary Gottfredson for advice on an earlier draft.

D. C. Gottfredson (✉)
Department of Criminology & Criminal Justice, University of Maryland, 2220D LeFrak Hall, College Park, MD 20742-8235, USA
e-mail: dgottfredson@crim.umd.edu

Keywords Deviancy training · Prevention · Harmful preventive interventions · Experimental research

Joan McCord had a long and extremely productive research career. She authored or edited 12 volumes and published 127 journal articles and book chapters on a wide range of topics in criminology (Academy of Experimental Criminology, AEC, n.d.). Professor McCord helped us understand that even well-intentioned, well-implemented prevention programs (such as the Cambridge–Somerville Youth Project, CSYP) sometimes have harmful effects on participants. Her work on this randomized study left her with an enduring respect for the value of experiments for identifying effective and ineffective delinquency prevention practices. At the time of her death, McCord was serving as the third President of the AEC.

In this paper, I reflect on the work of Joan McCord and discuss how her work influenced my own. Specifically, I show that a mechanism called "deviancy training" is central to both bodies of research and forms at least part of the explanation for why some presumably helpful preventive interventions actually harm participants. I begin with a summary of some of Professor McCord's contributions. Then, I provide examples of how my work has built upon hers. I end with a few recommendations for future research in this vein.

The Cambridge-Somerville Youth Project

The CSYP was designed to test the idea that boys who seemed to be headed for trouble could be steered away from a delinquent career if "a devoted individual outside his own family gives him consistent emotional support, friendship, and timely guidance" (Allport 1951, p. vi). Each of the boys in the treatment condition, whose ages ranged from 5 to 13 (median = 10.5) at the start of the project and who were living in the urban centers of Cambridge and Somerville, Massachusetts, were assigned to a counselor (most of whom were professionally trained social workers). These counselors visited the homes of the boys twice per month (on average) for 5.5 years. They provided "friendly guidance to the boys, counseled parents, assisted the families in a variety of ways, and referred the boys to specialists when that seemed advisable" (McCord 2003, p. 19). The boys were tutored, received medical and psychiatric attention, and participated in recreational activities with the counselors. Almost half the boys also attended summer camp. The intervention was an "ameliorative" type intervention in that the counselors sought to understand the boys' problems and to help improve their situations. They did not attempt to alter the social conditions that might have given rise to these situations.

During 1937 and 1938, nearly 2,000 referrals were received for the program from schools, social service agencies, churches, police, and other sources. The program staff collected extensive pre-treatment data to assist in the selection process. A screening process narrowed the pool of referrals to 782 boys who seemed to meet the selection criteria for the program in terms of age, location, and the availability of information. A selection committee then used the information to produce a "prognostic indicator of the probabilities of good or bad behavior" (Powers and

Witmer 1951, p. 54). The ratings ranged from "−5" for boys clearly headed for delinquency to "+4" for boys thought to be unlikely to engage in antisocial acts. Approximately half the rated boys were considered pre-delinquent. Of these 782 boys, 650 were grouped into matched pairs on the basis of age, social background, somatype, and temperament, and then randomly assigned[1] either to receive the treatment or to become a member of a no-treatment control group. This classification and randomization process occurred on a rolling basis during 1937 and 1938, and by May of 1939, all 325 treatment boys had been assigned to a counselor's case load. A post-randomization check comparing a subset of 264 treatment boys who remained in the project by 1942 and their matched controls revealed no significant differences across 20 different variables that had been used in the matching process.[2]

Although the program was designed to last ten years, the provision of services was disrupted in major ways by World War II. The program was terminated in 1945, after treatment boys had received between 2.5 and 8 years of services. The early assessment of the outcomes of the program (Powers and Witmer 1951) reported no overall difference on a global measure of adjustment between the treatment and control boys.[3] Also, a follow-up of police and court records for all 325 pairs originally randomly assigned showed a slight disadvantage for the treatment youths.[4] The tone of the initial evaluation report, however, made clear that people involved with the study thought it likely that beneficial effects of the intervention would appear later as the boys matured.

Between 1975 and 1976, McCord and her research team traced the members of 253 of the matched pairs[5] through agency records and by using a variety of tracking strategies. They managed to locate 95% of the 506 men. Questionnaires eliciting

[1] Actually, all pairs were not randomly assigned. Eight pairs were assigned non-randomly due to pressures to begin treatment before the matches were finalized. Comparison cases were found later for these treatment cases. Twelve pairs were randomly assigned within pair. The remaining boys were grouped into approximately 45 homogeneous sets of matched pairs. These sets usually contained 8 or 9 pairs per set, but ranged as high as 25 matched pairs. The sets rather than the pairs within them were randomly assigned. In other words, if a set contained 10 pairs, with the pairs arranged in two columns, and the set was assigned to receive the treatment, all the boys in column one would be assigned to treatment and all in column two to control. This "set matching" strategy was intended to increase flexibility in later comparisons, but the strategy was later abandoned (Powers and Witmer 1951, p. 78).

[2] This post-randomization check was conducted after 60 pairs of boys were "retired" from the project in order to reduce case loads, and an additional boy had died. Cases were dropped non-randomly, based on the perceived needs of the treatment boys. Only boys with "relatively few needs" were dropped. No "difficult" boys were dropped. Of course, this adjustment had the potential to bias the randomized groups in favor of the control cases because it used information about only the treatment boys' post-randomization healthy adjustment to drop cases. Powers and Witmer's (1951, pp. 79–81) compared the groups on pre-treatment measures after this adjustment and reported no significant differences, but close examination of their data reveals that although the mean differences between the treatment and control groups were small relative to their standard deviations, the direction of the differences favored the control cases on 19 of the 20 variables.

[3] Based on the 148 pairs for whom outcome data could be obtained for both members.

[4] Police and court appearances were slightly higher among treatment youths, although a slightly higher percentage of control youths were found among the most serious offenders.

[5] The number of pairs used in different reports from the CSYP research differs slightly. The initial reports of treatment-control differences on official records are based on all 325 pairs. Witmer's initial outcome evaluation, relying primarily on case by case judgments of adjustment, was based on 254 pairs. Excluded from this analysis were 65 treatment cases that had been "retired" from the study in 1941 because their social adjustment had been uniformly good in the early period of the study, plus six additional boys who either died or moved away shortly after program services started. McCord's follow-up was based upon the 253 men "who had been in the treatment program after 1942" and their matched control (McCord 1978, p. 284).

information about marriage, children, occupations, drinking, health, and attitudes were sent to 410[6] of these men. The treatment cases were also asked to report how, if at all, the program helped them. Questionnaire responses were obtained from 113 treatment and 122 control men, or slightly less than 50% from each group. According to McCord (1978), two-thirds of the treatment subjects reported that the program had helped them in various ways, including keeping them out of trouble. She showed, however, that the court records, death records, and hospital records collected for all 253 pairs did not support this positive assessment. Equal numbers of treatment and control subjects had been convicted of a crime. Controlling for the initial delinquency prediction score or for the seriousness of the juvenile record did not produce group differences. The only significant difference that emerged on official criminal records was on the percentage of men with records who had multiple offenses, and this difference favored the control group: 78% of treatment but only 67% of control men had a criminal record indicating at least two crimes. McCord's follow-up showed remarkable similarity across groups in the number who had died (9.5% of each group), been treated for alcoholism (7% and 8%), and received treatment in mental hospitals for diseases other than alcoholism (8.3% of each group). But more subtle differences were found, all favoring the control group: Treatment men who died did so an average of six years earlier than controls, and the diagnoses of treatment men who were treated in mental hospitals were more serious than those of control men. Self-reports also showed treatment men had significantly higher rates of alcoholism and stress-related diseases than control men (McCord 1978). Similarly, self-reports of socioeconomic status indicators showed a considerable advantage for control men: 43% of the control men versus 29% of the treatment men reported occupations that were classified as white-collar or professional. The groups did not differ on a host of other outcomes measured, including general life satisfaction, use of leisure time, political orientation, and authoritarian attitudes.

McCord's conclusion from the 30-year follow-up study of CSYP was that although the program was "successful in achieving the short-term goal of establishing rapport between social workers and teenage clients . . . none of the objective measures confirmed hopes that the treatment had improved the lives of those in the treatment group" (McCord 1978, p. 288). Further, the program appears to have had subtle negative side effects.

What mechanism could have produced these subtle negative effects? Interestingly, Witmer, the original researcher for the study, noted that the overall null findings seemed to mask heterogeneity of effects across different individuals in the study. She noted that the boys who seemed to benefit most tended to have other personal and social resources already in place to compliment the CSYP services. She also noted that many of the boys initially judged as seriously maladjusted seemed to be harmed by the program. McCord's (1978) data showed no evidence that program outcomes interacted with the initial delinquency status of the boys. She proposed several alternative mechanisms to explain the subtle harmful effects of the program: (1) interaction with adults with different values may have caused internal conflict that

[6] McCord (1978) reports that 48 of the 480 men successfully located had died. She does not explain why questionnaires were sent to 410 men rather than to the 432 still living.

manifested later in disease and dysfunction; (2) agency intervention may have created dependency on assistance which, when removed, created lasting dependency and resentment; (3) the treatment program may have generated high expectations for later accomplishments that never panned out; or (4) receiving welfare services may have altered the youths' self-perceptions such that they were more likely than controls to view themselves as requiring help.[7]

Later in her career, McCord proposed a different mechanism to explain the possibly harmful effects of the CSYP. The term, "deviancy training," had been used to describe a social learning process of "contingent positive reactions to rule-breaking discussion" (Dishion et al. 1999, p. 756) that often occurred among boys and their friends. Dishion and others had documented that such deviancy training occurs when youths spend time in groups with other youths, and that it seems to increase future problem behaviors (Dishion et al. 1996; Patterson et al. 2000; Snyder et al. 2005). McCord teamed up with Dishion to investigate the possible role of deviancy training in producing the harmful effects observed in the CSYP. Recall that most of the services provided during CSYP were delivered to individuals, not groups. But approximately half the boys were sent to summer camps, which provided ample time during which youths, relatively unsupervised, could reinforce each other's bragging about deviant behaviors. Dishion et al. (1999) hypothesized that this reinforcement during summer camp might have increased subsequent problem behaviors among the CSYP treatment youths who attended these camps. They reported that the likelihood of observing negative outcomes, although not related to having been sent to summer camp only once, was extremely elevated relative to their matched controls (10:1) for treatment boys who were sent to summer camp more than once. Of course, CSYP treatment boys were not randomly assigned to the summer camp experience. Although those who attended summer camp were compared with their matched and randomized controls, the summer camp decision was made after randomization, and it is therefore possible that those chosen to attend summer camp were already showing signs of problems before they ever attended the camps.[8] Also, the study afforded no direct measure of the deviancy training mechanism. The finding is nonetheless provocative.

Positive Action Through Holistic Education (PATHE)

As I mentioned, my work has to some extent replicated and extended Professor McCord's. One of my first research projects (Gottfredson 1986a) began shortly after McCord published her 30-year follow-up of CSYP. Project PATHE was implemented between 1980 and 1983 by the Charleston, South Carolina, school district with funds provided through the Office of Juvenile Justice and Delinquency

[7] McCord (1981) attempted to test these explanations to the extent possible using data from the follow-up questionnaires and found some support only for the "high expectations" explanation in that treatment cases were more likely to report being dissatisfied with their jobs and marriages. However, lack of available data to directly measure the mechanisms hindered this effort.

[8] Dishion et al. (1999) reported that 41% versus 33% (ns) of those ever versus never sent to summer camp had a prior prediction of delinquency, but they do not report the more relevant comparison of those sent twice versus the other groups.

Prevention's Alternative Education Initiative. Like the CSYP, PATHE involved an "ameliorative" intervention. But, unlike CSYP, this "direct service" component was embedded within an environmental change intervention aimed at altering the school environment for all the youths in the school. Program staff identified youths in each school who were considered at elevated risk for academic or behavioral problems. They collected data on each youth and developed a problem diagnosis based on the data. Individualized objectives were developed and specific counseling and academic services were prescribed for each youth. Two full-time specialists per school, usually a teacher and a counselor, provided these services. The intensity of the individualized services varied from school to school as well as from year to year, but during the most intensive year of the program, the typical treatment youth had between 8 and 33 contacts with a specialist, depending upon the school (mean = 18, or about two contacts per month as in the CSYP). The intervention lasted for one school year for 75% of treatment students and two school years for the remainder. Another interesting parallel between CSYP and PATHE is that both were regarded by their implementers to be effective programs. PATHE teacher surveys administered in the last year of the program at each school showed that high percentages of the teachers in each school agreed that PATHE had a positive effect on discipline problems (range = 62–94%) as well as academic achievement (range = 80–100%).

While the environmental-change aspect of the program was evaluated using a quasi-experimental design, the evaluation of the direct services component used an experimental design. Referred youths (n=1,161, roughly 200 per school) were randomly assigned to treatment and control conditions, separately within each school. Post-randomization checks comparing the groups' pre-treatment standardized test scores, disciplinary referrals, suspensions, gender, age, and parental education levels suggested that the groups were equivalent at the start of the intervention period, at least on the measured variables.

The study is most often cited for the positive effects reported for the school-level intervention.[9] But Gottfredson (1986a) also reported results from the evaluation of the direct services component. The program had some positive effects on the school performance of treatment students: 76% of twelfth grade targeted students graduated, compared with 42% of twelfth grade students not targeted (p<.01), and during the first year of the program, 77% of targeted students were promoted to the next grade, compared with 70% of controls (p<.05). But the program did not reduce delinquent behavior. Instead, in six of the seven schools, measures of problem behavior showed higher levels for the treatment than for the control group (Gottfredson 1986b). The only significant difference overall between the two groups across seven measures of problem behavior also favored the controls: they self-reported less drug involvement than the treatment cases (p<.05).

Significant treatment by school interactions on problem behaviors were also observed. Although the treatment students in six of the seven schools reported higher levels of drug use than controls, the difference was large (δ=.52) and statistically significant (p<.01) in one school. In that school, a much larger percentage of

[9] For example, students in the participating high schools reported significantly less delinquent behavior and drug use, had fewer suspensions, and fewer school punishments after the first year of the program. Students in the comparison high school did not change significantly on these outcomes (Gottfredson 1986a).

treatment than control students also withdrew from school (30 vs 17%, $p<.05$).[10] Records of implementation showed that the number of contacts with target students was lower in this school ($n=8$) than the others (range = 11–32), but it is hard to imagine how occasional counseling and tutoring services could actually *increase* substance use and drop-out. I noted (Gottfredson 1986b, p. 30) considerable heterogeneity across schools in the extent to which specialists were guided by the underlying PATHE theory that had been developed by the program managers. In exit interviews with staff in the school which produced the negative effects, the specialists were unable to recite the guiding program principles that were to have formed the basis for the services to students. These specialists, more than others, seemed to be guided by their personal theories about how best to help the young people. Witmer (Powers and Witmer 1951) warned of the potential negative effects of assigning relatively untrained case workers to provide services to high risk children. After reviewing the case files describing the interactions between case workers and several boys who seemed to have been harmed by the CSYP intervention, she concluded, ". . . organizations that provide services of the 'friendship' type (that is, that do not have psychiatrically trained staff) must proceed very cautiously in encouraging the development of close relationships with seriously maladjusted individuals . . . an untrained worker always runs the risk of breaking trust with a neurotic boy whose sensitivities he does not understand or of increasing his anxiety and guilt" (pp. 517–518). It may be that the PATHE specialist intervention also served to exacerbate students' pre-existing problems in some cases. It is also possible, as McCord suggested, that the services had the unintentional side effect of reinforcing the youths' problem behaviors.

The deviancy training mechanism seems especially likely in the case of the PATHE intervention because the counseling and tutoring services were always delivered to small groups of high risk youths who were pulled out of their regular classes to receive these services. Relatively unsystematic "help" provided over the course of a school year in a small group setting to especially high risk youths is the type of intervention that might facilitate deviancy training. It is reasonable to hypothesize that this mechanism diminished any positive effects that might have resulted from the contacts with specialists and tipped the problem behavior outcomes in the unfavorable direction. As in the CSYP, however, we had no measure of deviancy training and thus could not test this possibility.

An enhanced after-school program

Our most recent work on after-school programs (ASPs) afforded an opportunity to measure the deviancy training process directly. We were particularly interested in studying the effectiveness of ASPs for adolescents because previous research had

[10] The differential drop-out rate presents a challenge to the internal validity of the treatment-control comparison of self-reported substance use, as surveys were completed only by subjects who remained in school. However, the direction of the likely bias (more at-risk students dropping out of the treatment than the control group would bias the findings in favor of the treatment group) operates in the opposite direction of the observed finding and therefore is not a likely alternative explanation for the difference reported.

shown mixed results for participation in such programs, especially for problem behavior outcomes. Gottfredson et al. (in press) summarized results from prior research on the effectiveness of ASPs serving middle-school-aged youths. We noted that prior studies varied considerably in their level of scientific rigor. Only half the existing studies could be considered reasonably rigorous in terms of their research designs.[11] These stronger studies provided inconsistent findings regarding the benefits of ASP participation for middle-school youths on school attendance, academic performance, and student misconduct. Although two studies of the effects of ASP participation on misconduct found positive effects (Huang et al. 2005; Smith and Kennedy 1991), two others found that ASP participation *increased* misconduct (Dynarski et al. 2004; Weisman et al. 2002).

The two studies that reported undesirable effects of ASP participation on conduct outcomes were of large, relatively unstructured programs. The evaluation of the nation's largest ASP, 21st Century Community Learning Centers (Dynarski et al. 2003, 2004), found that participating middle-school students were more likely to have had their property damaged, more likely to report they had used or sold drugs, and less likely to rate themselves positively at working out conflicts with others (Dynarski et al. 2003). The 21st Century programs served an average of 60 youths each per day and provided mostly homework assistance and recreational activities. Similar negative results were uncovered by Weisman et al. (2002) in a quasi-experimental evaluation of 22 ASPs. Participants in the programs that Weisman et al. studied reported more conduct problems (including a wider variety of drug use) and more drug-using peers than did non-participants. Weisman et al. noted that the evaluated programs, like the 21st Century programs, were large and offered much unstructured programming. In another study relating program characteristics to effects across 35 ASPs, Gottfredson et al. (2007) found that smaller programs had more beneficial effects on delinquency outcomes. Based on this prior research, we surmised that harmful effects of ASP participation on delinquency might be due to an increased opportunity for deviancy training afforded when youth socialize with their friends at higher rates in unstructured programs.

Recently, my students and I completed a study (Gottfredson et al. in press) designed to test the effects of an "enhanced" ASP model on problem behavior and learning outcomes. The enhanced model incorporated several structured components[12] that had been demonstrated to reduce problem behaviors or increase learning in other studies. By requiring adherence to this standardized model, we expected to demonstrate that a well-implemented, evidence-based, structured ASP program would reduce delinquency. Previous experience had taught us that programs are seldom implemented as intended, however, especially in "real-world" settings. We therefore carefully measured implementation quality and program structure, expecting variability both across sites and time within site, and also measured deviancy

[11] "Reasonably rigorous" was defined as (1) using either a randomized design or sufficient controls on identified pretest differences between groups, (2) having more than 100 participants per experimental condition, and (3) either having attrition of less than 20%, or evidence demonstrating that differential attrition between the study groups was not present.

[12] The plan for the enhanced ASP included attendance monitoring and reinforcement, structured tutoring, and a structured prevention curriculum. See Gottfredson et al. 2010a.

training directly in order to test the association between the levels of structure observed and the amount of deviancy training.

Five urban middle schools were recruited to participate in this study. All students who attended these schools were invited to register for the ASP, and those who registered were randomly assigned either to attend the ASP or be invited to recreational events after school once per month. Gottfredson et al. (in press) report that treatment and control students did not differ in terms of demographics, and they differed significantly on only 1 of the 18 pretreatment measures of the projects' goals and objectives. One difference out of 18 tests conducted is approximately what would be expected by chance using a critical value of $p<.05$. The randomization produced groups of 224 treatment and 223 control students, about half of whom were male (54%), and the majority of whom were African American (70%). The average age of participants was 12.2, and 59% received subsidized meals at school.

We contracted the Baltimore County Department of Recreation and Parks to deliver the program. After the program had run for one school year, we found that the treatment cases still resembled the control cases on a large number of outcomes. No significant posttest differences were found on measures of conduct problems, academic performance, school attendance, prosocial or antidrug attitudes, social competence, school bonding, or positive peer influences. Youths in the treatment group engaged in less unsupervised socializing after school than did youths in the control group, but not as much less as would be expected if the after-school program were providing consistent supervision to youth who would otherwise be unsupervised. Our reports on the project explored various reasons for the poor showing of the program.[13] In the remainder of this talk, I will focus on deviancy training as one potential mechanism.

Although we did not observe overall negative outcomes due to participation in the ASP, we thought it possible that the anticipated positive effects due to the evidence-based programming were offset by negative effects due to deviancy training. Our measures of deviancy training provided some evidence that this may have occurred. Our deviancy training data came from direct observations of approximately 3,000 five-minute intervals during 398 discrete activities that occurred in the ASPs (Rorie et al. in press). These observation occasions were deliberately selected to include activities that varied according to the extent to which expectations for how youths should spend their time were clearly defined and communicated, which we defined as "structure" and measured on a five-point scale. Activities included a prevention curriculum and academic activities (both of which were expected to be more structured), as well as recreational and other activities (e.g., snack time) that were expected to be less structured. On any given observation day, research assistants selected two more- and two less-structured activities for observation. For each five-

[13] Gottfredson et al. (in press) showed that attendance was low overall (students attended an average of 36 days during the year), but that attending more days was not generally related to better outcomes. Cross et al. (2009) reported that the program did not attract many delinquency-prone youths who were unsupervised, suggesting that the students most in need of the program did not attend. Further, data obtained from a mid-year activity survey revealed that youths in the study were engaged in many additional after-school activities. The addition of our experimental after-school program to the mix of available activities made little impact on the frequency with which students participated in organized activities after school.

minute interval in each selected activity, the observers recorded on a coding sheet the level of structure, the problem behaviors observed, and responses to those behaviors by both peers and group leaders. They coded instances of violence as well as defiant or inappropriate behavior, and recorded talking about and pretending to engage in problem behavior as well as actual misconduct.

Analyses of these responses indicated, first, that while peer responses generally reinforce the undesirable behaviors, the predominant response of group leaders was neutral rather than chastising (see Fig. 1). Further, correlations between the observed structure in each five-minute interval and responses to the problem behavior showed, as expected, that peer and group leader responses to problem behaviors are less reinforcing in more highly structured activities. Figure 2 shows the percentage of peer responses to defiant or inappropriate behaviors (which we labeled "counter-normative behaviors") that were reinforcing, by activity type. The activities in the two left-most bars are more structured, while those in the three right-most bars are less structured.[14] Peers clearly reinforce these behaviors at higher rates during less structured activities ($p<.01$). Finally, Rorie et al. (in press) used multi-level analyses that took account of both within-activity and between-activity variation in structure to show that higher levels of structure in the activity as a whole and across five-minute intervals within an activity are significantly related to lower levels of violent behavior.[15]

These analyses provided evidence that what might be considered "peer deviancy training" does occur in the context of ASPs, that it is more likely to occur during less structured activities, and that more violent behavior also occurs during these less structured times. They also showed that the adult supervision afforded in ASPs does not counteract the reinforcing effect of peers. Hence, this mechanism might explain the finding that large, unstructured programs are more likely to increase conduct problems.

We also used the deviancy training observations to explore the unexpected null effects of the prevention curriculum, All Stars (www.allstarsprevention.com), that was embedded in the ASP. All Stars is intended to prevent substance use and to reduce bullying, violence, and other conduct problems. Prior research had demonstrated that the program, when delivered in school settings, can decrease substance use (Hansen and Dusenbury 2004; McNeal et al. 2004). As noted above, our ASP produced no positive effects on any of the measured outcomes.[16]

We explored whether deviancy training in the context of the All Stars program might have reduced the effectiveness of the curriculum. We noticed in our observations of All Stars that the group leader who appeared to implement the program most faithfully and to connect most effectively with the young people in the

[14] Rorie et al. (in press) reported that the structure level varies by activity type ($p<.01$): life skills instruction 3.54; academic activities 3.09; creative recreation 3.01; active recreation 2.94; other activities (e.g., snack time) 1.99. Structure is measured on a five-point scale, with higher scores indicating greater structure.

[15] Significant associations of structure and other deviant behaviors were also observed, but they were inconsistent across the different levels. For example, "counter-normative" behavior was related to between-observation but not within-observation structure. See Rorie et al. (in press).

[16] A report on the All Stars component of the ASP (Gottfredson et al. 2010b) noted that, although the curriculum was implemented faithfully in the ASP, low attendance due to the voluntary nature of the ASP decreased program exposure. Nevertheless, low exposure did not explain the null effects, as we also observed no effects for those students whose exposure to the program was high.

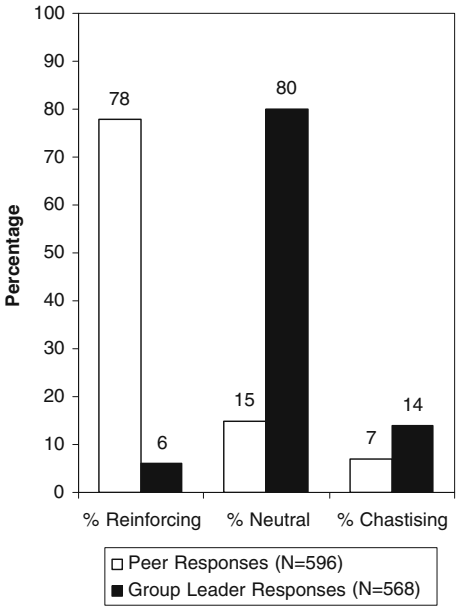


Fig. 1 Responses to violent behavior, peers and group leaders

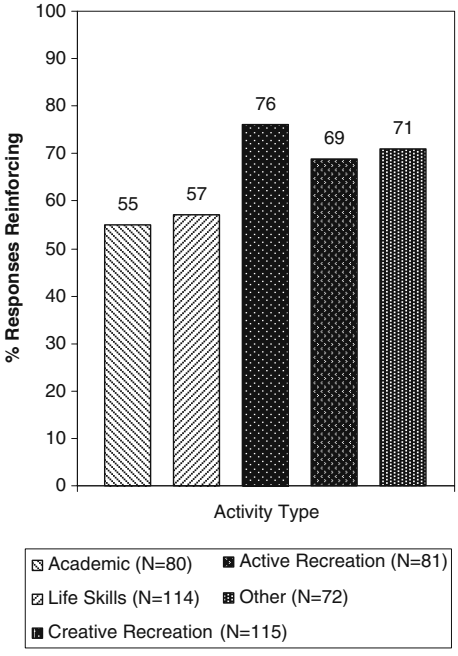


Fig. 2 Peer responses to "counternormative" behavior, by activity

program also tended to subtly approve of their deviant behavior by laughing, smiling, and otherwise responding to students' stories about using drugs and engaging in other delinquent activities. This potentially reinforcing instructor behavior tended to occur more often during the All Stars segment of the program because the curriculum, which requires open discussion of topics including substance use and other illegal behaviors, provided greater opportunity for subtle reinforcement of talk about deviant behaviors. Our systematic observations bore this out. We found that activity type (All Stars vs other activities) moderated the association between program fidelity (measured as described in Gottfredson et al. 2010b) and measured reinforcement of deviant talk to produce more reinforcement of deviant talk in All Stars at the site that implemented the program with the highest quality. Figure 3 shows this result. Group leader responses to deviant talk were most positive (a score of "3" is reinforcing "1" is chastising) during All Stars at the most well-implemented site ($p<.05$). At other sites, group leader responses were similar during All Stars and other activities. Peer responses to deviant talk were also more positive during All Stars than during other activities only at the best site, but the difference for peer responses was not statistically significant.

Some evidence also suggests that the deviancy training mechanism may translate into negative outcomes for youths. The effect size for ASP participation on a scale assessing beliefs that illegal, violent, and risky behaviors are common and acceptable in the peer group (the All Stars mediator we would most expect to respond to deviancy training) is near zero for students attending lower-implementation schools, while it is −.30 (e.g., moderate magnitude favoring the control students) at the site in which we observed the group leader's subtle reinforcement of deviant talk. It is possible that the group leader at this site, who was particularly close with students and related to them easily, created an environment in which students felt more comfortable expressing themselves, but wherein they also learned that drug use and anti-social conduct is commonplace.

This pattern of associations suggests that prevention programs, by encouraging dialogue about deviant behavior among youths and adults, may increase the likelihood that youths will openly express pro-deviancy opinions and that respected adults may unintentionally reinforce these expressions. These programs may inadvertently provide a stage for deviancy training by adults.[17]

Although the research summarized here is suggestive of a deviancy training mechanism that has the capacity to undermine a prevention program that otherwise has the potential to be effective, it is suggestive at best. Stronger evidence would be required to isolate the effects of deviancy training. All the research on deviancy training to date has been correlational and cannot be used to rule out alternative explanations. For example, selection is an equally plausible explanation for the finding that youths in CSYP who attended summer camp more than once were at elevated risk for negative outcomes relative to their equivalent controls. Similarly, in

[17] Interestingly, this same pattern of results has been reported elsewhere. Hallfors et al. (2007) examined mechanisms which might explain negative effects uncovered in an evaluation of a different prevention program, Reconnecting Youth. They found that teacher encouragement of deviant attitudes was positively related to a measure of program fidelity such that teachers who delivered the program with higher fidelity also provided the most encouragement for deviancy. These same high-fidelity teachers were also responsible for producing the most negative student outcomes.

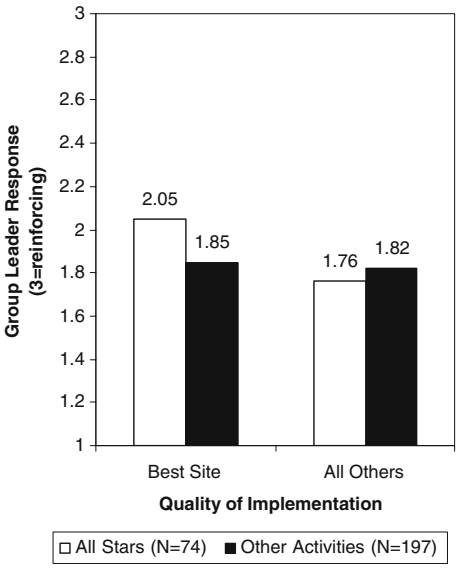


Fig. 3 Group leader responses to deviant talk, All Stars versus other activities, by implementation quality

our ASP study, students may have engaged in more violent behavior during unstructured ASP activity segments because violent behavior tends to destabilize activities rather than because the increased peer reinforcement observed during these activities increases violent behavior. Finally, we recognize that it is not necessarily the reinforcing actions of our best All Stars instructor that led students to endorse negative beliefs at higher rates than their controls. This outcome may have resulted from any number of other characteristics of the teacher or his setting—or it may have been a fluke.

Next steps

The time is ripe for a good experiment. This experiment could manipulate deviancy training in the context of a preventive intervention by randomly assigning youths to one of two conditions. In one condition, teachers would be trained to encourage open discussions of substance use and other deviant acts, and both peers and group leaders would be free to show approval for such expressions if they so chose. In the other condition, teachers would provide the same program content, but would control the message by discouraging free discussions and especially reinforcement for pro-deviancy expressions. Observers would measure the fidelity of implementation of these conditions using direct measures of reinforcement similar to those used in our ASP study. The evaluators would also carefully measure the content of the program delivered in the two conditions as a check on its equivalence. Subsequent misbehavior would be assessed and related to the study conditions. This is the type of evidence that is now required to move our understanding of the possible negative effects of deviancy training to the next level. I hope that this talk

has stimulated the interest of one or more experimental criminologists to pursue this line of research.

In closing, let me say that, all too often, evidence of harmful effects in preventive interventions is downplayed. Joan McCord, reflecting on reasons for our field's reluctance to publicize results from interventions that harm (McCord 2003), suggested that this reluctance is due in part to our fear that "all social programs will be tainted by the ones that are harmful" (p. 17). Researchers who publish negative results may fear that their results will bring harm to the overall prevention cause. I contend, as did McCord, that only by seeking to understand the mechanisms that might lead to both beneficial and harmful effects in prevention programming will we learn how to prevent crime.

References

Academy of Experimental Criminology (n.d.). Retrieved October 21, 2009 from http://www.crim.upenn.edu/aec/jmccord.htm.

Allport, G. (1951). Foreword. In E. Powers & H. Witmer (Eds.), *An experiment in the prevention of delinquency: The Cambridge-Somerville Youth Study*. New York: Columbia University Press.

Cross, A. B., Gottfredson, D. C., Wilson, D. M., Rorie, M., & Connell, N. (2009). The impact of after-school programs on the routine activities of middle school students: results from a randomized, controlled trial. *Criminology & Public Policy, 8*, 391–412.

Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy, 27*(3), 373–390.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: peer groups and problem behavior. *The American Psychologist, 54*(9), 755–764.

Dynarski, M., James-Burdumy, S., Moore, M., Rosenberg, L., Deke, J., & Mansfield, W. (2003). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers Program: Final Report*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington: U.S. Government Printing Office.

Dynarski, M., James-Burdumy, S., Moore, M., Rosenberg, L., Deke, J., & Mansfield, W. (2004). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers Program: New findings*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington: U.S. Government Printing Office.

Gottfredson, D. C. (1986a). An empirical test of school-based environmental and individual interventions to reduce the risk of delinquent behavior. *Criminology, 24*, 705–731.

Gottfredson, D. C. (1986b). *An assessment of a delinquency prevention demonstration with both individual and environmental interventions*. Report No. 361, Center for Social Organization of Schools. Baltimore: The Johns Hopkins University.

Gottfredson, D. C., Cross A., & Soule D. A. (2007). Distinguishing characteristics of effective and ineffective after-school programs to prevent delinquency and victimization. *Criminology & Public Policy, 6*, 289–318.

Gottfredson, D. C., Cross, A. B., Wilson, D. M., Rorie, M., & Connell, N. (2010a). *A randomized trial of the effects of an enhanced after-school program for middle school students*. Final report submitted to the U.S. Department of Education Institute for Educational Sciences.

Gottfredson, D. C., Cross, A. B., Wilson, D. M., Rorie, M., & Connell, N. (2010b). An experimental evaluation of the All Stars prevention curriculum. *Prevention Science, 11*(2), 142–154. doi:10.1007/s11121-009-0156-7.

Gottfredson, D. C., Cross, A. B., Wilson, D. M., Rorie, M., & Connell, N. (in press). Effects of participation in after-school programs for middle school students: A randomized trial. *Journal of Research on Educational Effectiveness*.

Hallfors, D., Cho, H., Dishion, T., Cuffee, J., & Hartman, S. (2007). *The role of teachers in school-based interventions for high risk adolescents: An alternative paradigm for predicting deviancy training*. Unpublished manuscript.

Hansen, W. B., & Dusenbury, L. (2004). All Stars Plus: a competence and motivation enhancement approach to prevention. *Health Education, 104*, 371–381.

Huang, D., Sung Kim, K., Marshall, A., & Perez, P. (2005). *Keeping kids in school: An LA's BEST example*. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center, Center for the Study of Evaluation (CSE), Graduate School of Education and Information Studies, University of California, Los Angeles.

McCord, J. (1978). A thirty-year follow-up of treatment effects. *The American Psychologist, 33*(3), 284–289.

McCord, J. (1981). Consideration of some effects of a counseling program. In S. E. Martin, L. B. Sechrest, & R. Reder (Eds.), *New directions in the rehabilitation of criminal offenders* (pp. 394–405). Washington: The National Academy of Sciences.

McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. In D. Weisburd, A. Petrosino & C. Lum (Eds.), Assessing systematic evidence in crime and justice: methodological concerns and empirical outcomes. *The Annals of the American Academy of Political and Social Science, 587*: 16–30. Thousand Oaks: Sage Publications.

McNeal, R. B., Hansen, W. B., Harrington, N. G., & Giles, S. M. (2004). How All Stars works: an examination of program effects on mediating variables. *Health Education & Behavior, 31*, 165–178.

Patterson, G. R., Dishion, T. J., & Yoerger, K. (2000). Adolescent growth in new forms of problem behavior: macro- and micro-peer dynamics. *Prevention Science, 1*(1), 3–13.

Powers, E., & Witmer, H. (1951). *An experiment in the prevention of delinquency: The Cambridge-Somerville Youth Study*. New York: Columbia University Press.

Rorie, M., Gottfredson, D. C., Cross, A., Wilson, D. M., & Connell, N. (in press). Structure and deviancy training in after-school programs. *Journal of Adolescence*. doi:10.1016/j.adolescence.2010.01.007.

Smith, C., & Kennedy, S. D. (1991). *Final impact evaluation of the Friendly PEERsuasion program for Girls Incorporated*. New York: Girls Incorporated.

Snyder, J., Schrepferman, L., Oeser, J., Patterson, G., Stoolmiller, M., Johnson, K., et al. (2005). Deviancy training and association with deviant peers on young children: occurrence and contribution to early-onset conduct problems. *Development and Psychopathology, 17*, 397–413.

Weisman, S. A., Womer, S. C., Kellstrom, M. A., Bryner, S. L., Kahler, A., Slocum, L., et al. (2002). *Maryland after school community grant program, report on the 2001–2002 school year evaluation of the phase 3 after school programs*. Technical report available from the authors.

Denise C. Gottfredson is a Professor at the University of Maryland Department of Criminal Justice and Criminology. She received a Ph.D. in Social Relations from The Johns Hopkins University, where she specialized in Sociology of Education. D. Gottfredson's research interests include delinquency and delinquency prevention, and particularly the effects of school environments on youth behavior. Gottfredson's most recent research projects include experimental evaluations of the effectiveness of the Baltimore City Drug Treatment Court, the Strengthening Families Program in Washington D.C., and structured after-school programs in Baltimore County, Maryland.