

Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency: A Systematic Review of the Randomized Experimental Evidence

By
ANTHONY PETROSINO,
CAROLYN TURPIN-
PETROSINO,
and
JOHN BUEHLER

Scared Straight and other programs involve organized visits to prison facilities by juvenile delinquents or at-risk kids to deter them from delinquency. Despite several research studies and reviews questioning their effectiveness, they remain in use and have now been tried in at least six nations. The authors report here on the results of a systematic review of randomized experimental tests of this program. Studies that tested any program involving the organized visits of delinquents or at-risk children to penal institutions were included. Each study had to have a no-treatment control condition with at least one outcome measure of "postvisit" criminal behavior. Using extensive search methods, the authors located nine randomized trials meeting eligibility criteria. After describing the studies and appraising their methodological quality, the authors present the narrative findings from each evaluation. A meta-analysis of prevalence rates indicates that the intervention on average is more harmful to juveniles than doing nothing. The authors conclude that governments should institute rigorous programs of research to ensure that well-intentioned treatments do not cause harm to the citizens they pledge to protect.

Keywords: Scared Straight; juvenile awareness; delinquency prevention; meta-analysis; randomized experiments

Background

In the 1970s, inmates serving life sentences at a New Jersey prison began a program to "scare" at-risk or delinquent children from a future life of crime.¹ The program, known as Scared Straight, featured as its main component an aggressive presentation by inmates to juveniles visiting the prison facility. The presentation brutally depicted life in adult prisons and often

Anthony Petrosino is coordinator and Steering Committee member for the Campbell Collaboration Crime and Justice Group (www.aic.gov.au/campbellcj). He is also research consultant for the Study on Decisions in Education at Harvard University. Anthony worked for several years for state justice agencies in New Jersey and Massachusetts, completed his Ph.D. in criminal justice at Rutgers University in 1997, and accepted a Spencer

DOI: 10.1177/0002716203254693

included exaggerated stories of rape and murder (Finckenauer 1982). A television documentary on the program aired in 1979 and provided evidence that sixteen or the seventeen delinquents interviewed in the film remained law-abiding for three months after attending Scared Straight, a 94 percent success rate (Finckenauer 1982). The program received considerable and favorable media attention and was soon replicated in more than thirty jurisdictions nationwide, resulting in special congressional hearings on the program and film by the U.S. House Subcommittee on Human Resources (U.S. House 1979).

The underlying theory of programs like Scared Straight is deterrence. Program advocates and others believe that realistic depictions of life in prison and presentations by inmates will deter juvenile offenders (or children at risk of becoming delinquent) from further involvement with crime. Although the harsh presentation in the earlier New Jersey version is the most famous, inmate presentations are now

Post-Doctoral Fellowship in Evaluation at the Harvard Children's Initiative. His current research includes a systematic review of "what works" to prevent crime against the commercial airline industry.

Carolyn Turpin-Petrosino is associate professor in the Department of Sociology and Criminal Justice at Bridgewater State College in Massachusetts. She received her B.S. from Howard University, a M.S.W. from Rutgers School of Social Work, and a Ph.D. in criminal justice from Rutgers University. Topics of her recent articles include parole decision making, hate crimes, and community policing. She recently won the Fiore World Justice Award to study codification of hate crime legislation in penal codes across the globe.

*John Buehler is a Ph.D. candidate in the Harvard University Department of Statistics. He has a B.A. in mathematics from Williams College and an A.M. in statistics from Harvard University. For 5 years he was a research assistant at the Center for Evaluation at the American Academy of Arts and Sciences. He coauthored a paper ("Comparing Results of Large Clinical Trials to Those of Meta-Analysis," *Statistics in Medicine* 2002) with Professor Emeritus Lincoln Moses of Stanford University and Professor Emeritus Frederick Mosteller of Harvard University. His current research interests involve estimating variance parameters in Gaussian Hierarchical Models.*

NOTE: This review was completed as a pilot test for the newly formed international Campbell Collaboration (C2) using the existing infrastructure of the Cochrane Collaboration Development, Psychosocial, and Learning Disorders Editorial Group. The full text of this review was published in the *Cochrane Library* (beginning in 2002, issue 2) and will also be made available in the *Campbell Collaboration Reviews of Interventions and Policy Effects (C2-RIPE)* database. This review was partially supported by grants from the Smith-Richardson Foundation (to the University of Pennsylvania), the Mellon Foundation (to the Center for Evaluation, Initiatives for Children Program at the American Academy of Arts and Sciences), and the U.K. Home Office (to Cambridge University Institute of Criminology). The Criminal Justice Collection at Rutgers University's Center for Law and Justice and the Gutman Library at the Harvard Graduate School of Education facilitated interlibrary loan requests. We thank Phyllis Schultze and Carla Lillvik for their expertise, patience, and assistance. We appreciate all of the help of the Cochrane Developmental, Psychosocial, and Learning Disorders Group: the overall guidance by Dr. Jane Dennis, coordinator; guidance and editorial criticisms by Professor Geraldine MacDonald, coordinating editor; additional comments by Dr. Stuart Logan and other Editorial Group members; additional searches run by Jo Abbott, trials coordinator, and Celia Almedia, researcher; and the statistical criticisms by Julian Higgins. Comments by Professor Robert Boruch, Sir Iain Chalmers, Dr. Phoebe Cottingham, Professor Lyn Feder, and Professor Joan McCord on earlier drafts of this work have helped.

sometimes designed to be more educational than confrontational but with a similar crime prevention goal (Finckenauer and Gavin 1999; Lundman 1993).² It is not surprising why such programs are popular: they fit with common notions by some on how to prevent or reduce crime (by “getting tough”), they are very inexpensive (a Maryland program was estimated to cost less than U.S.\$1 per participant), and they provide one way for incarcerated offenders to contribute productively to society by preventing youngsters from following down the same path (Finckenauer 1982).

A randomized controlled trial of the New Jersey program in 1982, however, reported no effect on the criminal behavior of participants in comparison with a

The goal of this review is to assess the effects of programs comprising organized visits to prisons of juvenile delinquents (officially adjudicated or convicted by a juvenile court) or predelinquents (children in trouble but not officially adjudicated as delinquents) aimed at deterring them from criminal activity.

no-treatment control group (Finckenauer 1982). In fact, Finckenauer (1982) reported that participants in the experimental program were more likely to be arrested. Yet belief in the program’s efficacy continued. Finckenauer called the process by which policy makers, practitioners, media reporters, and others sometimes latch onto quick, short-term, and inexpensive cures to solve difficult social problems the “Panacea Phenomenon.” Other randomized trials reported in the United States also questioned the effectiveness of Scared Straight–type programs in reducing subsequent criminality (Greater Egypt Regional Planning and Development Commission 1979; Lewis 1983). Consistent with these findings, reviewers of research on the effects of crime prevention programs have not found deterrence-oriented programs like Scared Straight effective (Sherman et al. 1997; Lundman 1993; Lipsey 1992). In fact, the University of Maryland’s well-publicized review of more than five hundred crime prevention evaluations listed Scared Straight as one program that does not work (Sherman et al. 1997).

Despite this seeming convergence of evidence, Scared Straight–type programs remain popular and continue to be used (Finckenauer and Gavin 1999). For example, a program in Carson City, Nevada, brings juvenile delinquents on a tour of an adult Nevada state prison (Scripps 1999). One youngster claimed that the part of the tour that made the most impact on him was “all the inmates calling us for sex and fighting for our belongings” (Scripps 1999). The United Community Action Network (U-CAN) has its own program called Wisetalk in which at-risk youth are locked in a jail cell for more than an hour with four to five parolees. They claim that only ten of three hundred youngsters exposed to this intervention have been re-arrested (U-CAN 2001). In 2001, a group of guards—apparently without the knowledge of administrators—strip-searched Washington, D.C., students during their tours of a local jail under the guise of that they were using “a sound strategy to turn around the lives of wayward kids”—claiming the prior success of Scared Straight (Blum and Woodlee 2001).

Scared Straight and other “kids visit prison” programs have been used in several other nations. For example, it is called the “day in prison” or “day in gaol” in Australia (O’Malley, Coventry, and Walters 1993), “day visits” in the United Kingdom (Lloyd 1995), and the “Ullersmo Project” in Norway (Storvoll and Hovland 1998). Hall (1999) reported positively on a program in Germany designed to scare straight young offenders with ties to neo-Nazi and other organized hate groups. The program has been also tried in Canada (O’Malley, Coventry, and Walters 1993). In a different variant in the United Kingdom, a program was initiated that employed ex-prison guards to re-create a prison atmosphere in public schools, with the goal of deterring any potential lawbreakers (Middleton, Lilford, and Hyde 2001).

In 1999, *Scared Straight: 20 Years Later* was shown on U.S. television and reported similar results to those of the 1979 film (see also Muhammed 1999). The 1999 version reported that ten of the twelve juveniles attending the program have remained offense-free in the three-month follow-up (Muhammed 1999). As in the 1979 television program, no data on a control or comparison group of young people were presented. Positive reports and descriptions of Scared Straight–type programs have also been reported elsewhere (e.g., in Germany [Hall 1999] and in Florida [Rasmussen and Yu 1996]), although it is sometimes imbedded as one component in a multicomponent juvenile intervention package (Trusty 1995; Rasmussen and Yu 1996). In 2000, Petrosino and his colleagues reported on a preliminary analysis of a systematic review, drawing on the raw percentage differences in each study. They found that Scared Straight and like interventions generally *increased* crime between 1 and 28 percent when compared to a no-treatment control group. This article updates that review and utilizes more sophisticated meta-analytic techniques to analyze the data.

Method

The goal of this review is to assess the effects of programs comprising organized visits to prisons of juvenile delinquents (officially adjudicated or convicted by a

juvenile court) or predelinquents (children in trouble but not officially adjudicated as delinquents) aimed at deterring them from criminal activity.

Eligibility criteria

We included only randomized or quasi-randomized (i.e., alternation assignment procedures such as assigning every other case to treatment) controlled trials, provided they had a no-treatment control group. Only studies involving juveniles, that is, children aged seventeen or younger, were included. Participants were delinquents or predelinquents. Studies that contain overlapping samples of juveniles and young adults (e.g., ages thirteen to twenty-one) were also included. The intervention had to feature a visit by program participants to a prison facility as its main component. The interest of citizens, policy and practice decision makers, media, and the research community is in whether Scared Straight and other “kids visit prison” programs have any crime deterrent effect on the kids participating in them. We therefore focused on crime measures: each eligible study reported on at least one outcome measure of subsequent criminality (e.g., arrest, conviction, police contact, self-reported criminality). We also list other measures reported by investigators in case subsequent reviewers focusing on these and other outcomes require them to identify potentially eligible studies.

Search strategy for identification of studies

To minimize publication bias or the possibility journals are more likely to publish findings that reject the null hypothesis (and find programs to be more effective than unpublished literature generally does), we conducted a search strategy designed to identify published and unpublished studies. We also conducted a comprehensive search strategy to minimize discipline bias; that is, evaluations reported in criminological journals or indexed in field-specific abstracting databases might differ from those reported in psychological, sociological, social service, public health, or educational sources.

First, randomized experiments were identified from a larger review of field trials in crime reduction conducted by the first author. Petrosino (1997) used the following methods to find more than 300 randomized experiments (and analyze 150): (1) “handsearch” (i.e., visually inspecting the entire contents) of twenty-nine leading criminology or social science journals; (2) checking the cites reported in the *Registry of Randomized Experiments in Criminal Sanctions* (Weisburd, Sherman, and Petrosino 1990); (3) detailed electronic searches of *Criminal Justice Abstracts*, *Sociological Abstracts* and *Social Development and Planning Abstracts* (Sociofile), Education Resource Information Clearinghouse (ERIC) and *Psychological Abstracts* (PsycInfo); (4) searches by information specialists (before some of these became available online) of eighteen bibliographic databases, including the specialized collection maintained by the National Criminal Justice Reference Service (NCJRS); (5) an extensive mail campaign with more than two hundred researchers and one hundred research centers; (6) published solicitations in association news-

letters; (7) tracking of references in more than fifty relevant systematic reviews and literature syntheses; and (8) tracking of references in relevant bibliographies, books, articles, and other documents. More detail about these search methods can be found in Petrosino (1995, 1997). The citations found in Petrosino (1997) covered literature with a publication date between 1 January 1945 and 31 December 1993. Seven randomized trials meeting the eligibility criteria were identified from this sample.

Second, we augmented this work with searches designed to find experiments possibly overlooked by Petrosino's (1997) search methods and to cover more recent literature (1994-2001). These methods included (1) broad searches of the Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR) developed by the U.K. Cochrane Centre and now supervised by the University of Pennsylvania Graduate School of Education (Petrosino et al. 2000); (2) check of citations from more recent systematic or traditional reviews (e.g., Sherman et al. 1997; Lipsey and Wilson 1998); (3) citation checking of studies and other reports on the program (e.g., Finckenauer and Gavin 1999); (4) e-mail correspondence with selected researchers; and (5) broad searches of the Cochrane Controlled Trials Register [CENTRAL] in the *Cochrane Library*. By broad searches, we mean that we tried to first identify studies relevant to crime or delinquency and then we visually scanned the citations (and abstracts in most cases) to see if any were relevant.

Third, we decided to conduct a more specific search of fourteen available electronic databases relevant to the topic area. Many of these include published and unpublished literature (e.g., dissertations or government reports). Searches were done online using available Harvard University resources or other databases freely searchable via the Internet. Several trips were made to the University of Massachusetts to use *Criminal Justice Abstracts* and other bibliographic databases not accessible at Harvard or via the Internet. Jane Dennis, Jo Abbott, and Celia Almedia of the Cochrane Developmental, Psychosocial, and Learning Disorders Group repeated the searches on most of these databases to cover literature added in 2001. The bibliographic databases and the years searched were *Criminal Justice Abstracts*, 1968 through September 2001; *Current Contents*, 1993 through 2001; *Dissertation Abstracts*, 1981 through August 2001; *Education Full Text*, June 1983 through October 2001; ERIC, 1966 through 2001; *GPO Monthly* (Government Printing Office), 1976 through 2001; MEDLINE (Medical Literature Analysis and Retrieval System Online), 1966 through 2001; National Clearinghouse on Child Abuse and Neglect database, through 2001; NCJRS, through 2001; *Political Sciences Abstracts*, 1975 through March 2001; *PAIS International* (Public Affairs Information Service), 1972 through October 2001; PsychInfo, 1987 through November 2001; *Social Sciences Citation Index*, February 1983 through October 2001; and Sociofile, January 1963 through September 2001.

We anticipated that the amount of literature on Scared Straight would be of moderate size and that our best course of action would be to identify all citations relevant to the program and screen them for potential leads to eligible studies. This removed the need to include keywords for identifying randomized trials (e.g., ran-

dom assignment) in our searches. After several trial runs, we found that nearly all documents used phrases like *Scared Straight* or *juvenile awareness* in the title or abstract of the citation. Therefore, the following searches were run in each relevant database to identify relevant citations: *Scared Straight; (prison or jail or reformatory or institution) and (orientation or visit or tour); prisoner run or offender run or inmate run; prison awareness or prison aversion or juvenile awareness; and (rap session or speak out or confrontation) and (prisoner or lifer or inmate or offender)*.

Finally, we conducted searches of the Internet and World Wide Web using the above terms in two popular search engines: Hotbot and Altavista. We later updated this with another search using Google.

Selection of trials

The search methods above generated more than five hundred citations (most had abstracts). Anthony Petrosino (AP) screened these citations, determining that thirty were to evaluation reports. AP and Carolyn Turpin-Petrosino (CTP) independently examined these citations and were in agreement that eleven were leads to potential randomized trials. Seven had already been retrieved in AP's earlier review (Petrosino 1997). We determined that the full-text reports for four should be pursued. These were obtained either through interlibrary loan or by visits to area libraries. Upon inspection of the full-text reports, we determined that two studies should be excluded. Dean's (1982) randomized trial studying the impact of Project Aware in a Wisconsin state prison did not include any postprogram measure of offending, and our attempts to find the author or retrieve this data from any other reports by the Wisconsin Department of Corrections have been unsuccessful. Chesney-Lind's (1981) study of Stay Straight in Hawaii was also excluded, as matching rather than random assignment to conditions was used. After all exclusions, we were left with nine randomized trials for analysis.³ We did not find any reports of ongoing trials.

Data management and extraction

AP extracted data from each of the nine main study reports using a specially designed instrument. The data collection instrument was adapted from his earlier study (Petrosino 1997) and included such items as year published, type of document, location of study, evidence that randomization successful, randomization problems, type of prison setting (e.g., adult, reformatory), number in treatment group, number in control group, average age of participants in study, percentage of white participants in study, severity of prior offence record of participants, study attrition problems, number of follow-up measurements of outcome variable, type of outcomes collected: noncriminal (e.g., educational, health) and criminal, and specific crime outcome data for experimental and control groups at each interval.

In cases in which outcome information was missing from the original reports, we made attempts via e-mail and regular mail correspondence to retrieve the data for the analysis from the original investigators. We were unsuccessful in obtaining any

additional data from investigators, although in two cases, we retrieved unpublished (fugitive or “grey” literature) master’s theses from university libraries to see if they contained more detailed reporting (Cook 1990; Locke 1984). They did not. We ran statistical analyses using Cochrane Collaboration’s MetaView statistical software, a component of Review Manager Version 4.1 (RevMan). These were repeated, and additional analyses run, using Meta Analyst software created by Dr. Joseph Lau of the New England Cochrane Center. One of us (John Buehler) also created meta-analytic formulae in Excel to double-check three of the analyses. Results were identical.

Description of Studies

Collectively, the nine studies were conducted in eight different states, with Michigan the site for two studies (Yarborough 1979; Michigan Department of Corrections 1967). No set of researchers conducted more than one experiment. The studies span the years from 1967 to 1992. Five studies were unpublished and were disseminated in government documents or dissertations; the remaining four were found in academic journal or book publications. The average age of the juvenile participants in each study ranged from fifteen to seventeen. Only the New Jersey study included girls (Finckenauer 1982). Racial composition across the nine experiments was diverse, ranging from 36 to 84 percent white. Most of the studies dealt with delinquent youth already in contact with the juvenile justice system. Nearly 1,000 (946) juveniles or young adults participated in the nine randomized studies.

All of the experiments were simple two-group experiments except Vreeland’s (1981) evaluation of the Texas Face-to-Face program. Only one study used quasi-random alternation techniques to assign participants (Cook and Spirrisson 1992); the remaining studies claimed to use randomization, although not all were explicit about how such assignment was conducted. Only the Texas study (Vreeland 1981) included data on self-report measures. In two studies (Cook and Spirrisson 1992; Locke et al. 1986), no prevalence rates were reported. Some of the studies that did include average or mean rates did not include standard deviations to make it possible to compute the weighted mean effect sizes. Also, the follow-up periods were diverse and included measurements at three, six, nine, twelve, and twenty-four months. We summarize each below, in chronological order by publication date.

In an internal, unpublished government document, the Michigan Department of Corrections (1967) reported on a trial testing a program that involved taking adjudicated juvenile boys on a tour of a state reformatory. Unfortunately, the report is remarkably brief. Sixty juvenile delinquent boys were randomly assigned to attend two tours of a state reformatory or to a no-treatment control group. Tours included fifteen juveniles at a time. No other part of the program is described. Recidivism was measured as a petition in juvenile court for either a new offense or a violation of existing probation order.

The Scared Straight program at the Menard Correctional Facility in Illinois started in 1978 and is described as a frank and realistic portrayal of adult prison life

(Greater Egypt Regional Planning and Development Commission 1979). The researchers randomly assigned 161 youth aged thirteen to eighteen to attend the program or a no-treatment control. The participants were a mix of delinquents or children at risk of becoming delinquent.

In the JOLT (Juvenile Offenders Learn Truth) program, juvenile delinquents in contact with one of four Michigan county courts participated (Yarborough 1979). Each juvenile spent five total hours (half of that time in the rap session) in the facility. After a tour of the facility, they were escorted to the cell, subjected to interaction with inmates (e.g., taunting), and then taken to a confrontational rap session with inmates. In the evaluation, 227 youngsters were randomly assigned to JOLT or to a no-treatment control. Participants were compared on a variety of crime outcomes collected from participating courts at three- and six-month follow-ups.

*Whether relying on the actual data reported
or measures of statistical significance, the nine
trials do not yield evidence for a positive
effect for Scared Straight and other
juvenile awareness programs on
subsequent delinquency.*

The Insiders Program in Virginia was described as an inmate-run, confrontational intervention with verbal intimidation and graphic descriptions of adult prison life (Orchowsky and Taylor 1981). Juveniles were locked in a cell fifteen at a time and told about the daily routine by a guard. They then participated in a two-hour confrontational rap session with inmates. Juvenile delinquents from three court service units in Virginia participated in the study. The investigators randomly assigned eighty juveniles aged thirteen to twenty with two or more prior adjudications for delinquency to the Insiders Program or a no-treatment control group. Orchowsky and Taylor (1981) reported on a variety of crime outcome measures at six-, nine-, and twelve-month intervals.

The Face-to-Face program in Texas included a thirteen-hour orientation session in which the juvenile lived as an inmate. Counseling followed. Participants were fifteen to seventeen years of age and on probation from Dallas County Juvenile Court; most averaged two to three offenses before the study. One hundred sixty boys were assigned to four conditions: prison orientation and counseling, ori-

entation only, counseling only, or a no-treatment control group. Vreeland (1981) examined official court records and self-reported delinquency at six months.

The New Jersey Lifers' Program began in 1975 and stressed confrontation with groups of juveniles, ages eleven to eighteen, who participated in a rap session. Finckenaue (1982) randomly assigned eighty-two juveniles, some of whom were not delinquents, to the program or to a no-treatment control group. He then followed them for six months in the community, using official court records to assess their behavior.

The California SQUIRES (San Quentin Utilization of Inmate Resources, Experience and Studies) program, which began in 1964, was the oldest in the United States. The SQUIRES program included male juvenile delinquents from two California counties between the ages of fourteen and eighteen, most with multiple prior arrests. The intervention included confrontational rap sessions with rough language, guided tours of prison with personal interaction with prisoners, and a review of pictures depicting prison violence. The intervention took place one day per week over three weeks. The rap session was three hours long and normally included twenty youngsters at a time. In the study, 108 participants were randomly assigned to treatment or to a no-treatment control group. Lewis compares them on seven crime outcomes at twelve months.

The Kansas Juvenile Education Program (JEP) was an intervention to educate children about the law and the consequences of violating it. The program also tried to roughly match juveniles with inmates based on personality types. Fifty-two juvenile delinquents aged fourteen to nineteen from three Kansas counties were randomly assigned while on probation to JEP or a no-treatment control. The investigators examined official (from police and court sources) and self-report crime outcomes at six months for program attendees and a no-treatment control group.

The Mississippi Project Aware was a nonconfrontational educational program comprising one five-hour session run by prisoners. The intervention was delivered to juveniles in groups numbering from six to thirty. In the study, 176 juveniles between the ages of twelve and sixteen under the jurisdiction of the county youth court were randomly assigned to the program or to a no-treatment control. The experimental and control groups were compared on a variety of crime outcomes retrieved from court records at twelve and twenty-four months.

Methodological Quality of Included Studies

There are many factors on which to grade the quality of studies. Complicating any assessment of methods is that review teams, by and large, must rely on written reports by investigators. In some cases, methodology sections may be briskly written (sometimes because of journal space requirements), and key features of design and analysis may be deleted or considerably condensed. We determined that four were most critical to criminological experiments and practical to extract from the experimental reports. These were (1) randomization integrity, (2) attrition from

initial sample, (3) blinding of outcome assessors, and (4) fidelity of program implementation. We discuss each of these below.

Randomization integrity

Did the investigators report that participants to experimental or control conditions experienced any violation or subversion of random assignment procedures? If so, did they report how many cases were incorrectly assigned to conditions? Only one study reported problems with randomization, and they were dramatic (Finckenauer 1982). Only eight of the eleven participating agencies that referred troubled or delinquent boys to the program correctly assigned their cases. Finckenauer (1982) did conduct additional analyses in an attempt to compensate for violation of randomization. We agreed that a sensitivity analysis should be done to determine the influence of this evaluation on the pooled analysis.

Attrition

Did the investigators report major attrition or loss of participants from the sample initially randomized? If so, did they report how much attrition from the initial sample occurred? The Virginia Insiders study reported a major loss of participants from the initial randomization sample (Orchowsky and Taylor 1981). This was reported, however, at the second and third follow-up intervals (not the first, at six months). Because there was a paucity of data beyond the first follow-up interval across studies, we only conducted a pooled analysis using the this “first effect.” Therefore, a sensitivity analysis of the impact of this later attrition was not performed. The Michigan JOLT study did report a large number of no-shows, but they were deleted from the analysis. The problem is that we do not know how many participants were initially assigned, and we have no assurances from investigators that the remaining sample was similar to the initial sample. We determined to also conduct a sensitivity analysis to determine the influence of this study on the pooled analysis.

Outcome assessor blinding

It is nearly impossible to “blind” either the practitioners or the kids to experimental conditions, particularly when evaluating stark contrasts such as Scared Straight versus nothing. But it is possible to blind outcome assessors—who are collecting data on success and failure—to treatment assignment. Only one study reported blinding outcome assessors (Michigan Department of Corrections 1967). Nonetheless, we believe the likelihood of corruption by any bias from outcome data assessors is very low. The reason is that in several of the experiments reported here, a law enforcement agency (a state criminal records agency or the Federal Bureau of Investigation) far removed from the treatment setting did the criminal records check. We cannot imagine that these agencies would be given anything

more than the participants' names and other identifying information—not what group they were assigned to.

Program implementation

Did the investigators report that the program was so poorly implemented that the evaluation was not an accurate assessment of the effectiveness of the intervention? These programs appear to be relatively simple and short-term and therefore pose few problems for implementation. Not one investigator reported any implementation problems. By and large, the kids received what they were designed to receive.

Results

Narrative findings

Whether relying on the actual data reported or measures of statistical significance, the nine trials do not yield evidence for a positive effect for Scared Straight and other juvenile awareness programs on subsequent delinquency. For example, the Michigan Department of Corrections (1967) study found that 43 percent of the experimental group recidivated, compared to only 17 percent of the control group. No statistical test is reported. Curiously, more attention is not provided to this large negative result in the original document.

In Illinois, the outcomes are statistically insignificant but negative in direction, with 17 percent of the experimental participants being recontacted by police in contrast to 12 percent of the controls (Greater Egypt Regional Planning and Development Commission 1979). The authors concluded, "Based on all available findings one would be ill advised to recommend continuation or expansion of the juvenile prison tours. All empirical findings indicate little positive outcome, indeed, they may actually indicate negative effects" (*ibid.*, 19). Researchers report no effect for the program on attitude measures (Jesness Inventory, Piers-Harris Self-Concept Scale). In contrast, interview and mail surveys of participants and their parents and teachers indicated unanimous support for the program (*ibid.*, 12). Researchers also note how positive and enthusiastic the adult inmates were about their efforts.

The second Michigan study also reported very little difference between the intervention and control group (Yarborough 1979). The average offense rate for program participants, however, was .69 compared to .47 for the control group. Yarborough (1979, 14) concluded that "the inescapable conclusion was that youngsters who participated in the program, undergoing the JOLT experience, did no better than their control counterparts."

The only positive findings, though statistically insignificant, were reported in Virginia (Orchowsky and Taylor 1981). At six months, the results slightly favored

the control group (39 percent of controls had new court intakes versus 41 percent of experimental participants), but they favored the experimental participants at nine and twelve months. The investigators noted, however, that the attrition rates in their experiment were dramatic. At nine months, 42 percent of the original sample dropped out, and at twelve months, 55 percent dropped out. The investigators conducted analyses that seemed to indicate that the constituted groups were still comparable on selected factors.

A study of the Face-to-Face Program in Texas also reported little effect for these interventions. Vreeland (1981) reported that the control participants outperformed the three treatment groups on official delinquency (28 percent delinquent versus 39 percent for the prison orientation plus counseling, 36 percent for the

Despite the variability in the type of intervention used, on average these programs result in an increase in criminality in the experimental group when compared to a no-treatment control. According to these experiments, doing nothing would have been better than exposing juvenile to the program.

prison only, and 39 percent for the counseling only). The self-report measure, however, showed the reverse. None of these findings were statistically significant. There were discrepancies between the self-report and official data; some who were officially charged did not self-report the offense and vice versa. Viewing all the data, Vreeland concluded that there was no evidence that Face-to-Face was an effective delinquency prevention program. He found no effect for Face-to-Face on several attitudinal measures, including the "Attitudes toward Obeying Law Scale."

Finckenauer (1982) reported that 41 percent of the kids who attended the Scared Straight program in New Jersey committed new offenses, while only 11 percent of controls did, a difference that was statistically significant. He also reported that the program participants committed more serious offenses. He also reported no

impact of the program on nine attitude measures except one: experimental participants do much worse on a measure called "attitudes toward crime." His concerns about randomization integrity are treated in a later sensitivity analysis.

Additional evidence of a possible harmful effect can be found in the evaluation of the SQUIRES program. Lewis (1983) reported that 81 percent of the program participants were arrested compared to 67 percent of the controls. He also found that the program did worse with seriously delinquent youths, leading him to conclude that such kids could not be "turned around by short-term programs such as SQUIRES . . . a pattern for higher risk youth suggested that the SQUIRES program may have been detrimental" (p. 222). The only data supporting a deterrent effect for the program was the average length of time it took to be rearrested: 4.1 months for experimental participants and 3.3 months for controls. Data were reported on eight attitudinal measures, and Lewis reported that the program favored the experimental group on all of them.

Locke and his colleagues (1986) reported little effect of JEP in the Kansas state prison. Both groups improved from pretest to posttest, but the investigators concluded that there were no differences between experimental and control groups on any of the crime outcomes measured. Investigators also reported no effect for the program on the Jesness and Cerkevich psychological tests.

Finally, little difference was found between experimental and control participants in the Mississippi Project Aware study (Cook and Spurrison 1992). For example, the mean offending rate for controls at twelve months was 1.25 for control cases versus 1.32 for Project Aware participants. Both groups improved from twelve to twenty-four months, but the control mean offending rate was slightly lower than the experimental group. The investigators concluded that "attending the treatment program had no significant effect on the frequency or severity of subsequent offenses" (*ibid.*, 97). The investigators also reported on two educational measures: school attendance and dropouts. Curiously, they reported that Project Aware reduced school dropout rates but noted that "it is not clear how the program succeeded in reducing dropout rates" (*ibid.*, 97).

Meta-analysis

For each study, we extracted all of the relevant crime outcome data. Our initial plan (the Cochrane/Campbell protocol) included an organization of analyses by examining official reports (from government administrative records) distinct from self-reported criminality (obtained from investigator-administered survey questionnaires). Given that we expected a diverse number of measures of crime to be reported, we thought it would be best to organize it into four indexes that would be most relevant to policy and practice: prevalence rates (i.e., what percentage of each group failed or succeeded?), average incidence rates (i.e., what was the average number of offenses or other incidents per individual in each group?), offense severity rates (i.e., what was the average severity of offenses per individual in each group?), and latency (i.e., how long was the average return to crime or failure delayed per individual in each group?).

Where possible, we converted the outcome measures to Cohen's effect size d , which is also known as the standardized mean difference (between the experimental and control groups divided by the pooled standard deviation). Using *d-Stat* meta-analytic software created by Blair Johnson of the University of Connecticut, d can be approximated using prevalence data, test statistics, and other information when the means and standard deviations are not available. Table 1 provides the distribution of these effect sizes, which indicate the size and direction of effect for the program. Bold type indicates positive effect sizes (where the intervention had a positive impact on criminal behavior). As Table 1 demonstrates, nearly all of the positive effects for Scared Straight are generated by a single report by Orchowsky and Taylor (1981).

Given that few outcome measures of crime were reported across the studies, we were limited to a single meta-analysis. We report the crime outcomes for official measures at "first effect" (and usually the only effect reported). Each of the analyses focuses on the prevalence data, as the outcomes reporting means or averages are sparse and often do not include standard deviations. Thus, because the data rely on dichotomous outcomes, both analyses report odds ratios (ORs) for each study and their 95 percent confidence intervals (CIs). Because there is some disagreement in the literature about this, we tested the data assuming both random and fixed effects models for weighting the treatment effects across the studies.

The forest graph in Figure 1 plots the ORs for the seven studies reporting prevalence rates. We assume a random effects model (that the studies do not come from some single underlying population). Figure 1 shows that intervention increases the crime or delinquency outcomes at the first follow-up period. The mean OR is 1.72 (95 percent CI = 1.13 to 2.62) and is statistically significant.⁴ The intervention increases the odds of offending about 1.7:1 (1.7 treatment kids offend for every control participant who offends).

There is always a question about whether or not the results are being driven by experiments that reported methodological problems. To test for this, we conducted a sensitivity analysis (Figure 2). Specifically, we excluded the two studies identified in our methodological assessment as having potentially threatening flaws: the Finckenaue (1982) experiment because of concerns about randomization breakdown, and the Yarborough (1979) study because of the deletion of no-shows (which could indicate a potential for large attrition from the initial study sample). We again ran analyses assuming both random and fixed effects models (which did not differ). The deletion of these studies did not alter the results: these programs, according to the evidence reported here, have an overall negative (and statistically significant) impact on subsequent offending.

Discussion

These randomized trials, conducted over a quarter century in eight different jurisdictions and involving nearly one thousand participants, provide evidence that Scared Straight and other "juvenile awareness" programs are not effective as a

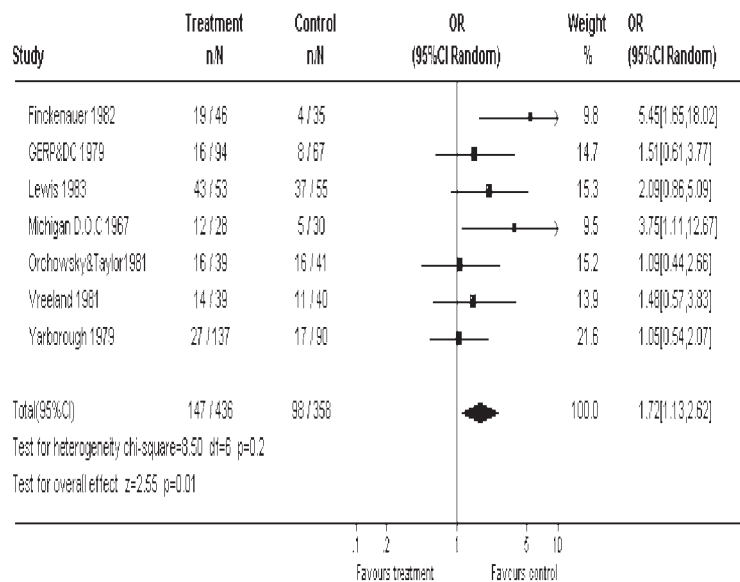
TABLE 1
 FULL ARRAY OF OUTCOME MEASURES TRANSFORMED TO COHEN'S EFFECT SIZE (*d*)

Study	Type of Data	Three Months				Six Months			
		P	I	S	L	P	I	S	L
MI DOC 1967	Court petition or probation violated					-.60			
GERP&DC 1979	Police contacts					-.14			
Yarborough 1979	New offenses	-.03	-.17	-.14	-.13	-.04	-.18	-.09	.10
	Petitions	.03				.04			
	Mean days in detention		.52				-.02		
Orchowsky and Taylor 1981	New intakes					-.04	.15	.29	
Vreeland 1981	Officially recorded delinquency					-.17			
	Self-reported delinquency					.21			
Finckenauer 1982	Complaints, arrests, and adjudications					-.71		-.60	
	Correction for randomization failure					-.33			
Lewis 1983	Arrests								
Locke et al. 1986	Juvenile court records								
	Self-reported offending								
							<i>F</i> = .75, unknown direction		
							E does worse, but no data provided		
Cook and Spurrison 1992	Reoffending (<i>F</i> test of E vs. C over time)								

		Nine Months				Twelve Months				Twenty-Four Months
		P	I	S	L	P	I	S	L	
MI DOC 1967	Court petition or probation violated									
GERP&DC 1979	Police contacts									
Yarborough 1979	New offenses									
	Petitions									
	Mean days in detention									
Orchowsky and Taylor 1981	New intakes	.50	.68	.82		.44	.83	.83		
Vreeland 1981	Officially recorded delinquency									
	Self-reported delinquency									
Finckenauer 1982	Complaints, arrests, and adjudications									
	Correction for randomization failure									
Lewis 1983	Arrests					-.32	E2.1, C2.2	<i>ns</i>	E4.1m, C3.3m (s)	
Locke et al. 1986	Juvenile court records									
	Self-reported offending									
Cook and Spurrison 1992	Reoffending (<i>F</i> test of E vs. C over time)									-.09 I

NOTE: Bold type indicates positive effect sizes. P = prevalence (percentage of each group failing); I = incidence (average offending per person); S = severity (average severity score or percentage of group committing "serious" offenses); L = latency (average amount of time to first offense per offender); E = experimental group; C = control group; m = months; s = statistically significant. MI DOC = Michigan Department of Corrections; GERP&DC = Greater Egypt Regional Planning and Development Commission.

FIGURE 1
FIRST EFFECT OF INTERVENTION, OFFICIAL
CRIME MEASURES, RANDOM-EFFECTS MODEL

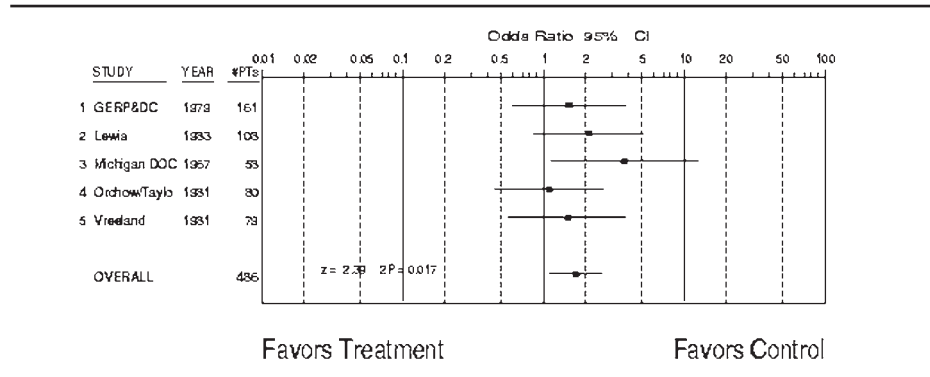


NOTE: n = number of participants reoffending; N = number assigned to group; OR = odds ratio; CI = confidence interval; weight = amount of weight given to study in analysis; GERP&DC = Greater Egypt Regional Planning and Development Commission; D.O.C. = Department of Corrections.

stand-alone crime prevention strategy. More important, they provide empirical evidence—under experimental conditions—that these programs likely increase the odds that children exposed to them will commit another delinquent offense. Despite the variability in the type of intervention used, on average these programs result in an increase in criminality in the experimental group when compared to a no-treatment control. According to these experiments, doing nothing would have been better than exposing juveniles to the program.

We note that the other two trials that did not report prevalence data for the meta-analysis also reported no effect for the intervention (Cook and Spurrison 1992; Locke et al. 1986). Indeed, the mean data from the Mississippi study is also negative in direction, and the Kansas investigators reported that the self-report data showed a negative impact. Our findings are also supported by a meta-analysis of juvenile prevention and treatment programs by Lipsey (1992), who indicated that the effect size for eleven “shock incarceration and Scared Straight programs” was $-.14$ (or produced about 7 percent higher recidivism rates in experimental participants than controls assuming a 50 percent baseline).

FIGURE 2
SENSITIVITY ANALYSIS (EXCLUDING FINCKENAUER 1982
AND YARBOROUGH 1979), RANDOM-EFFECTS MODEL



NOTE: PTs = number of participants; CI = confidence interval; GERP&DC = Greater Egypt Regional Planning and Development Commission; DOC = Department of Corrections. The DerSimonian and Laird procedure (1986) was used to create the random effects model.

Given the strong suggestion here that these programs have a harmful effect, they raise a dilemma for policy makers. Criminological interventions, when they cause harm, are not just toxic to the participants. They result in increased misery to ordinary citizens that come from the “extra” criminal victimization they create when compared to just doing nothing at all. Policy makers should take steps to build the kind of research infrastructure within their jurisdiction that could rigorously evaluate criminological interventions to ensure they are not harmful to the very citizens they aim to help.

Implications for practice

We note the following irony. Despite the gloomy findings reported here and elsewhere, Scared Straight and its derivatives continue in use, although a randomized trial has not been reported since 1992. As Finckenaue and Gavin (1999) noted, when the negative results from the California SQUIRES study came out, the response was to end evaluation—not the program. Today the SQUIRES program continues, evaluated by the testimonials of prisoners and participants alike. Some may argue that these trials, with the most recent reported in 1992, do not apply to the “newer” Scared Straight–type programs. We believe that our review places the onus on every jurisdiction to show how their current or proposed program is different from the ones studies here. Given that, they should then put in place rigorous evaluation to ensure that no harm is caused by the intervention.

Some literature indicates the program can have a positive effect on the inmate providers, and that argument is sometimes used to legitimize use of the program. Others claim that the program by itself is of little value but could be effective as part

of an overall multicomponent package of interventions delivered to youth. This raises ethical issues of whether such programs should still be permitted if they do increase criminality in juvenile participants.

Despite these findings here and our earlier report (Petrosino, Turpin-Petrosino, and Finckenauer 2000), we still get inquiries about how to get someone's son, daughter, or friend into a "Scared Straight" program. Many of these people are understandably looking for any program than can help "turn around" a wayward or antisocial youth. Unfortunately, we found no evidence that would support using this program for a particular type of kid with a special constellation of personality or other characteristics.

Implications for research

One of the critical questions raised by this review is why the program has a criminogenic effect. Some investigators presented theoretical or post hoc rationale for such results, but we did not find any of the theory-driven or theory-based evaluations that would have allowed investigators to test causal mechanisms or mediators that could provide clues as to why Scared Straight fails. Future research studies, including experimental trials, ought to formulate a causal model diagramming how the program is theorized to work—and then test critical variables that can be operationalized, measured, and tested (Petrosino 2000).

The type of broad search we undertook also should allow us to do a more in-depth study of the effects of using different evaluative designs to test criminological interventions. We now have a good collection of empirical studies on Scared Straight and plan to examine these for experimental, quasi-experimental, and nonexperimental evaluative design types.

Notes

1. Programs featuring inmates as speakers who describe their life experiences and the current reality of prison life have a rather long history, in the United States at least (Michigan Department of Corrections 1967).

2. Some of these programs featured interactive discussions between the inmates and juveniles, also referred to as "rap sessions."

3. The more detailed version of this review was published in the *Cochrane Library* (issue 2, 2002) and is forthcoming in *Campbell Collaboration Reviews of Interventions and Policy Effects (C2-RIPE)*. Readers are invited to consult these publications for more details on the nine included studies, including methodological features. It includes a full list of studies excluded from this synthesis and the rationale for such exclusions.

4. Assuming a fixed-effects model (that the studies come from one underlying population) did not change these findings. The odds ratio was 1.68, and this was also statistically significant.

References

- Blum J., and Y. Woodlee. 2001. Trying to give kids a good scare. *Washington Post*, 3 June, p. C01.
- Chesney-Lind, M. 1981. *The Na Pa'ahao: The experience of the prisoners. A juvenile awareness program*. Report no. 258. Manoa: University of Hawaii at Manoa, Youth Development and Research Center.

- Cook, D. D. 1990. Effects of a non-confrontational prisoner-run juvenile delinquency deterrence program. Unpublished master's thesis, Department of Psychology, Mississippi State University.
- Cook, D. D., and C. L. Spirrison. 1992. Effects of a prisoner-operated delinquency deterrence program: Mississippi's Project Aware. *Journal of Offender Rehabilitation* 17:89-99.
- Dean, D. G. 1982. The impact of a juvenile awareness program on select personality traits of male clients. *Journal of Offender Counseling, Services and Rehabilitation* 6 (3):73-85.
- DerSimonian, R., and N. Laird. 1986. Meta-analysis in clinical trials. *Controlled Clinical Trials* 7:177-88.
- Finckenaue, J. O. 1982. *Scared Straight and the Panacea Phenomenon*. Englewood Cliffs, NJ: Prentice Hall.
- Finckenaue, J. O., and P. W. Gavin. 1999. *Scared Straight: The Panacea Phenomenon revisited*. Prospect Heights, IL: Waveland Press.
- Greater Egypt Regional Planning and Development Commission. 1979. *Menard Correctional Center: Juvenile tours impact study*. Carbondale, IL: Greater Egypt Regional Planning and Development Commission.
- Hall, A. 1999. Jailhouse shock aims to scare youths straight. *The Scotsman*, 26 October, p. 12.
- Lewis, R. V. 1983. Scared Straight—California style: Evaluation of the San Quentin SQUIRES program. *Criminal Justice and Behavior* 10 (2): 209-26.
- Lipsey, M. W. 1992. Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In *Meta-analysis for explanation*, edited by T. C. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. L. Light, T. A. Louis, and F. M. Mosteller, 83-127. New York: Russell Sage Foundation.
- Lipsey, M. W., and D. B. Wilson. 1998. Effective intervention for serious juvenile offenders: A synthesis of research. In *Serious and violent juvenile offenders: Risk factors and successful interventions*, edited by R. Loeber and D. P. Farrington, 313-45. Thousand Oaks, CA: Sage.
- Lloyd, C. 1995. *To scare straight or educate? The British experience of day visits to prison for young people*. Home Office Research Study no. 149. London: Home Office.
- Locke, T. P. 1984. An analysis of the Kansas State Penitentiary Juvenile Education Program. Unpublished master's thesis, Department of Psychology, University of Kansas.
- Locke, T. P., G. M. Johnson, K. Kirigin-Ram, J. D. Atwater, and M. Gerrard. 1986. An evaluation of a juvenile education program in a state penitentiary. *Evaluation Review* 10:281-98.
- Lundman, R. J. 1993. *Prevention and control of juvenile delinquency*. 2d ed. New York: Oxford University Press.
- Michigan Department of Corrections. 1967. *A six month follow-up of juvenile delinquents visiting the Ionia Reformatory*. Research Report no. 4. Lansing: Michigan Department of Corrections.
- Middleton, John, Richard Lilford, and Chris Hyde. 2001. Applying Campbell Principles. Letter to the editor. *British Medical Journal* 323:1252.
- Muhammed, L. 1999. Kids and crooks revisited: Some were "Scared Straight!" *USA Today*, 12 April, p. 4D.
- O'Malley P., G. Coventry, and R. Walters. 1993. Victoria's Day in Prison Program: An evaluation and critique. *Australian and New Zealand Journal of Criminology* 26 (2): 171-83.
- Orchowsky, S., and K. Taylor. 1981. *The Insiders juvenile crime prevention program*. Richmond: Virginia Department of Corrections.
- Petrosino, A. J. 1995. The hunt for experimental reports: Document search and efforts for a "what works?" meta-analysis. *Journal of Crime and Justice* 18:63-80.
- . 1997. "What works?" revisited again: A meta-analysis of randomized experiments in rehabilitation, deterrence and delinquency prevention. Dissertation, Rutgers University, Ann Arbor, MI.
- . 2000. Answering the why question in evaluation: The causal-model approach. *Canadian Journal of Program Evaluation* 15 (1): 1-24.
- Petrosino, Anthony J., Robert F. Boruch, Cath Rounding, Steve McDonald, and Iain Chalmers. 2000. The Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR) to facilitate the preparation and maintenance of systematic reviews of social and educational interventions. *Evaluation Research in Education* 14 (3/4): 293-307.
- Petrosino A., C. Turpin-Petrosino, and J. O. Finckenaue. 2000. Well-meaning programs can have harmful effects! Lessons from experiments in Scared Straight and other like programs. *Crime & Delinquency* 46:354-79.
- Rasmussen, D. W., and Y. Yu. 1996. *An evaluation of juvenile justice innovations in Duval County, Florida*. Tallahassee: Florida State University.

- Scared Straight! 20 years later*. 1999. Broadcast 8 P.M., 15 April, on the UPN television network, host Danny Glover. Available from <http://www.contrib.andrew.cmu.edu/~aaron2/upn/upn-specials.html>.
- Scripps, J. 1999. Prison tour serves as a wake-up call. *The Forum*, 27 October, p. 1.
- Sherman, L. W., D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, and S. Bushway. 1997. *Preventing crime: What works, what doesn't, what's promising. A report to the United States Congress*. College Park: University of Maryland, Department of Criminology and Criminal Justice.
- Storvall, A. E., and A. Hovland. 1998. Ullersmoprojektet: "Scared Straight" I Norge 1992-1996 (The Ullersmo Project: Scared Straight in Norway 1992-1996). *Nordisk Tidsskrift for Kriminalvidenskab* 85 (2): 122-35.
- Trusty, G. 1995. Sheriff's office, juvenile court resources used for restarts. *Louisiana Youth Care Magazine* 18 (5): 3-5.
- United Community Action Network (U-CAN). 2001. Services: Wisetalk (Scared Straight). Retrieved 26 April 2001 from www.ucan.av.org/services.htm.
- U.S. House. 1979. Committee on Education and Labor. *Oversight on Scared Straight: Hearings before the House Subcommittee on Human Resources*. 96th Congress, 1st sess. 4 June. Washington, DC: Government Printing Office.
- Vreeland, A. D. 1981. Evaluation of Face-to-Face: A juvenile aversion program. Unpublished doctoral dissertation, University of Texas, Dallas.
- Weisburd, D. L., L. W. Sherman, and A. J. Petrosino. 1990. *Registry of randomized experiments in criminal sanctions, 1950-1983*. Los Altos, CA: Sociometrics Corporation, Data Holdings of the National Institute of Justice.
- Yarborough, J. C. 1979. *Evaluation of JOLT as a deterrence program*. Lansing: Michigan Department of Corrections.