

*DRAFT II: May 8, 2001*

*The effects of “Scared Straight“ and other juvenile awareness programs  
on subsequent offending*

**A PILOT TEST FOR THE CAMPBELL COLLABORATION  
CAMPBELL CRIME & JUSTICE GROUP**

USING THE EXISTING INFRASTRUCTURE OF THE  
COCHRANE DEVELOPMENTAL, PSYCHOSOCIAL & LEARNING DISORDERS GROUP  
DR. JANE DENNIS, COORDINATOR AND PROFESSOR GERALDINE MACDONALD, CHAIRPERSON

Anthony Petrosino  
American Academy of Arts & Sciences and Harvard University

Carolyn Turpin-Petrosino  
Bridgewater State College

John Buehler  
Harvard University

Acknowledgements

This work was principally supported by \$5,000 from a grant from the Smith-Richardson Foundation to the University of Pennsylvania Graduate School of Education (Robert Boruch, Principal Investigator). It received partial support from a Mellon Foundation grant to the Center for Evaluation, Initiatives for Children Program at the American Academy of Arts & Sciences (Frederick Mosteller, Principal Investigator) and a grant from the U.K. Home Office to Cambridge University Institute of Criminology (David Farrington, Principal Investigator). The latter two sources supported Anthony Petrosino's time during work and use of his office and computer at the American Academy of Arts & Sciences.

We appreciate the support, cajoling and comments by Dr. Jane Dennis, Coordinator and Dr. Geraldine MacDonald, Chairperson of the Cochrane Developmental, Psychosocial and Learning Disorders Group. Comments by Robert Boruch, Sir Iain Chalmers, Lynette Feder and Stuart Logan helped.

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.

## **ABSTRACT**

### **Background**

Scared Straight and other programs involve organized visits to prison by juvenile delinquents or kids at-risk for criminal behavior. The goal of such programs is that through first-hand observation of prison life and interaction - sometimes realistically brutal - with adult inmates, these juveniles will be deterred from future lawbreaking behavior. Despite research studies and reviews of varying quality questioning their effectiveness, they remain in use and have been tried in at least six nations (USA, UK, Canada, Australia, Germany, Norway). A 1999 film entitled "Scared Straight! 20 Years Later" aired in the USA and offered evidence of a strong, long-term effect of a New Jersey Scared Straight program on subsequent offending (Aims Multimedia 1999).

### **Objectives**

To assess the effects of programs comprising organized visits to prisons of juvenile delinquents (officially adjudicated or convicted by a juvenile court) or pre-delinquents (children in trouble but not officially adjudicated as delinquents) aimed at deterring them from criminal activity.

### **Search strategy**

Extensive searching using multiple methods was conducted. Earlier work by the first author in identifying randomized field trials 1945-1993 relevant to criminology formed an initial pool of studies. This was updated with specific searches of 14 electronic data bases using keywords particular to the intervention (e.g. "Scared Straight"), broad searches of the Campbell Collaboration Social, Psychological, Educational & Criminological Trials Register and the Cochrane Controlled Trials Register, citation chasing, and correspondence with investigators.

### **Selection criteria**

Studies that tested the effects of any program involving the organized visits of juvenile delinquents or children at-risk for delinquency to penal institutions were included. Studies that included overlapping samples of juvenile and young adults (e.g. ages 14-20) were included. We only considered studies that randomly or quasi-randomly (i.e. alternation) assigned participants to conditions. Each study had to have a no-treatment control condition. The study had to report at least one outcome measure of "post-visit" criminal behavior.

### **Data collection & analysis**

We conducted three main analyses of prevalence rates using official data, as information from other sources (self-report) or indexes (means) was either sporadically reported or missing critical information (standard deviations). Three analytic strategies were adopted. We examined (1) first-effects, (2) mean effects; and (3) best effects. We computed Odds Ratios (OR) for prevalence data, and assumed both fixed and random effects models in our analyses. We also computed the full array of Cohen's effect size (d) for all available outcome statistics.

### **Main results**

All three analytic strategies using OR show intervention to be more harmful than doing nothing. This is true whether we examined the first effect reported, the average effect (across all prevalence measures) in each study, or the strongest effect for the program. The full array of outcome measures (converted to d) show intervention had either no impact or harmful effects on measures of prevalence, incidence, severity and latency.

### **Reviewers' conclusions**

We conclude that programs like Scared Straight and other juvenile awareness programs likely have a harmful effect and increase delinquency relative to doing nothing at all to the same youths. Given these results, agencies that permit such programs must rigorously evaluate them not only to ensure that they are doing what they purport to do (prevent crime) - but at the very least they do not cause more harm than good.

### **Potential conflict of interest**

The authors did publish an earlier version of this review in the journal *Crime & Delinquency* that concluded that these programs likely had harmful effects (Petrosino, et al. 2000). We also published that article with Professor James Finckenauer (Finckenauer 1982; Finckenauer and Gavin 1999), who conducted the most famous experiment finding negative results. The authors, however, feel this is balanced by their desire to see prisoners do something constructive and effective to help delinquent youth!

## BACKGROUND

In the 1970s, inmates serving life sentences at a New Jersey (USA) prison began a program to "scare" or deter at-risk or delinquent children from a future life of crime. The program, known as "Scared Straight," featured as its main component an aggressive presentation by these inmates to juveniles visiting the prison facility. The presentation brutally depicted life in adult prisons, and often included exaggerated stories of rape and murder (Finckenauer 1982). An award-winning television documentary on the program aired in 1979 and provided evidence that 16 of the 17 delinquents remained law abiding for three months after attending "Scared Straight," a 94% success rate (Finckenauer 1982). Other data provided in the film indicated success rates that varied between 80 and 90% (Finckenauer 1982). The program received considerable and favourable media attention and was soon replicated in over 30 jurisdictions in the United States, resulting in special Congressional hearings on the program and film by the House Subcommittee on Human Resources (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 for becoming delinquent from further involvement with crime. Although the harsh and sometimes vulgar presentation in the earlier New Jersey version is the most famous, inmate presentations are now sometimes designed to be more educational than confrontational but with a similar crime prevention goal (Finckenauer and Gavin 1999; Lundman 1995). In fact, 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 (Brodsky 1970; Michigan Department of Corrections 1967). It is not surprising why such programs are popular. First, they fit with common notions by some on how to prevent or reduce crime (by "getting tough"). Second, they are very inexpensive; a Maryland program was estimated to cost approximately 80 cents per participant! Finally, 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 trial testing the New Jersey program in 1982, however, reported no effect on the criminal behavior of participants in comparison with a no treatment control group (Finckenauer 1982). In fact, Finckenauer reported that participants in the experimental program were more likely to be arrested. Yet, beliefs in the program's efficacy continued. Finckenauer called this the "Panacea Phenomenon," to explain why policymakers, practitioners, media and others often latch onto quick, short-term and inexpensive cures to solve difficult social problems. Other randomized trials reported in the USA soon questioned the effectiveness of Scared Straight type programs in reducing subsequent criminality (GERP&DC 1979, Lewis 1983). Parallel 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, Lipsey 1992; Lundman 1993). In fact, the University of Maryland Report to the United States Congress had many problems with the evaluation evidence across a wide variety of programs but did not hesitate in listing Scared Straight as one that doesn't work (Sherman et al 1997).

Despite the convergence of evidence from studies and reviews, Scared Straight type programs remain popular and continue to be used in the United States (Finckenauer and Gavin 1999). For example, a program in Carson City, Nevada (USA) 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 1999:1). The United Community Action Network has its own program called "Wisetalk" in which at-risk youth are locked in a jail cell for over an hour with 4-5 parolees. They claim that only 10 of 300 youngsters exposed to this intervention have been rearrested ([www.ucan.av.org](http://www.ucan.av.org)).

Scared Straight and other "kids visit prison" programs are also used in several other nations. For example, it is called the "day in prison" or "day in gaol" in Australia (O'Malley et al 1993), "day visits" in the United Kingdom (Lloyd 1995), and the "Ullersmo Project" in Norway (Storvall and Hovland 1998). Hall (1999) reports positively on a program in Germany designed to deter young offenders with ties to Neo-Nazi and other organized hate groups. "Scared Straight" has been also tried in Canada (O'Malley et al 1993).

In 1999, "Scared Straight: 20 Years Later" was shown on United States television and claimed similar results to the 1979 film (UPN 1999; "Kids and Crooks," 1999). In this version, the film reports that 10 of the 12 juveniles attending the program have remained offense free in the three months 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). Given that no prior review has exclusively examined Scared Straight type programs, a systematic review is required to assess the overall weight of the evidence from all relevant randomised trials.

## **METHODS OF THE REVIEW**

### **Objectives**

To assess the effects of programs comprising organized visits to prisons of juvenile delinquents (officially adjudicated or convicted by a juvenile court) or pre-delinquents (children in trouble but not officially adjudicated as delinquents), aimed at deterring them from criminal activity.

### **Criteria for considering studies for this review**

#### *Types of studies*

We included only randomized or quasi-randomized (i.e. alternation assignment procedures) controlled studies with or without blinding in our review. Each study must have included a no-treatment control group. We report the bibliography of excluded studies in Appendix 1.

#### *Types of participants*

Only studies involving juveniles, i.e. children 17 years of age or younger, were included. Participants were delinquents or pre-delinquents. Studies that contain overlapping samples of juveniles and young adults (e.g. ages 13-21) were also included.

#### *Types of interventions*

The intervention featured a visit by program participants to a prison facility as its main component. Most of the programs included a presentation by the inmates, ranging from graphic (Finckenauer 1982) to educational (Cook and Spirison 1992). Programs sometimes featured an orientation session (living as a prisoner for 8 hours, etc.) or a tour of the facility.

### *Types of outcome measures*

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 will therefore focus on crime measures. We do list the other measures reported by investigators in case reviewers in the Campbell Collaboration require them to identify potentially eligible studies.

## **Searching for Studies**

In order to reduce potential of publication bias, we conducted a search strategy designed to identify published and unpublished studies. We also conducted a comprehensive search strategy to reduce potential for discipline bias, i.e. that evaluations reported in criminological journals or indexed in field-specific abstracting data bases might differ from those reported in psychological, sociological, social service, public health or educational sources. We agree with those who argue that multiple search methods should be used rather than solely relying on electronic searches of bibliographic and abstract data bases in order to reduce potential for bias (Hunter and Schmidt 1992).

First, randomized experiments were identified from a larger review of field trials in crime reduction conducted by the first author (Petrosino 1997). Petrosino used the following methods to find more than 300 randomized experiments (and analyze 150): (1) handsearch (i.e. visually inspecting the entire contents) of 29 leading criminology or social science journals; (2) checking the cites reported in the Registry of Randomized Experiments in Criminal Sanctions (Weisburd et al. 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 of 18 bibliographic databases, including the National Criminal Justice Reference Service (NCJRS); (5) an extensive mail campaign with over 200 researchers and 100 research centers; (6) published solicitations in association newsletters; (7) tracking of references in over 50 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 published or available through 1993.

Second, we augmented this work with searches designed to uncover experiments missed by Petrosino (1997) and to cover more recent literature (1994-2001). These methods included: (1) broad searches of the Campbell Collaboration Social, Psychological, Educational & 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 to provide coverage of more recent

studies (e.g. Sherman et al. 1997; Lipsey and Wilson 1998); (3) citation checking of documents relevant to Scared Straight and like programs (e.g. Finckenaue and Gavin 1999); (4) email correspondence with investigators; and (5) broad searches of the Cochrane Controlled Trials Register in the Cochrane Library (Cochrane Collaboration 2001). By broad searches, we mean that we tried to first identify studies relevant to crime or delinquency and then visually scan the citations or abstracts to see if any were relevant to this intervention.

Third, we decided to conduct a more specific search of 14 available electronic data bases 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, Lowell to use Criminal Justice Abstracts and other data bases not accessible at Harvard or via the Internet. The bibliographic data bases and the years searched were:

Criminal Justice Abstracts, 1968- September 2000

Current Contents, 1993-2000

Dissertation Abstracts, 1981-August 2000

Education Full Text, June 1983-October 2000

ERIC (Education Resource Information Clearinghouse) 1966-2000

GPO Monthly (Government Printing Office Monthly), 1976-2000

Medline 1966-2001

National Clearinghouse on Child Abuse and Neglect (through 2001)

NCJRS (National Criminal Justice Reference Service) -2001

Political Sciences Abstracts, 1975-March 2000

PAIS International (Public Affairs Information Service), 1972-October 2000

PsychInfo (Psychological Abstracts) 1987-November 2000

Social Sciences Citation Index, February 1983-October 2000

Sociofile (Sociological Abstracts and Social Planning And Development Abstracts) January 1963-September 2000

The following specific searches were run in each relevant database to identify potentially eligible studies for the review:

"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"

("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.

## **Selection of Studies**

The search methods above generated over 500 citations (most with abstracts). We screened out most of this literature because it was comprised of opinion pieces, advocacy, or other non-empirical literature. We then screened citations and/or abstracts to over 30 potentially eligible evaluation studies and agreed that eleven full text reports should be pursued. These were obtained either through interlibrary loan or by our own visits to area libraries.

Upon inspection of the full reports, we agreed that two of the studies should be excluded. Dean's (1982) randomized trial studying the impact of Project Aware in a Wisconsin State Prison did not include any post-program 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. We also agreed to exclude Chesney-Lind's (1981) study of Stay Straight in Hawaii, in which she compared two groups of youngsters who attended the program with a comparison group that did not. Random assignment to conditions was not used. After all exclusions, we were left with nine randomized trials for analysis. We did not learn of any ongoing trials.

### **Data Management and Extraction**

Data were extracted from each of the nine main study reports, using a specially designed instrument. In cases in which outcome information was missing from the original reports, we made attempts via email 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 obscure Masters' theses from University Libraries to see if they contained more detailed reporting (Cook 1990; Locke 1984). They did not.

## **RESULTS**

### **Description of studies**

Collectively, the nine experiments 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 1967-1992. The first five studies located were unpublished and were disseminated in government documents or dissertations; the remaining four were found in academic journal or book publications. Our searches, therefore, were able to identify and retrieve some documents from the fugitive literature that are generally more difficult for reviewers to take account of. Few of the prior syntheses of crime prevention programs included the nine Scared Straight experiments we review here. For example, the University of Maryland Report to the US Congress by Sherman, et al. 1997 concludes that Scared Straight does not work based on negative results reported in three evaluations (Finckenauer 1982; Buckner and Chesney-Lind 1983; Lewis 1983) and the meta-analysis reported by Lipsey (1992).

The average age of the juvenile participants in each study ranged from 15 to 17. Only the New Jersey study included girls (Finckenauer 1982). Racial composition across the nine experiments was diverse, ranging from 36% to 84% white. Most of the studies dealt with delinquent youths already in contact with the juvenile justice system. All of the experiments included a no-treatment control group and all - but one - were simple two-group experiments. Vreeland (1981) is the exception. He used a factorial design in which juveniles were randomly assigned to four conditions: (1) prison

orientation and counseling; (2) prison orientation only; (3) counseling only; and (4) no-treatment control. Appendix 2 provides full detail on each of the included studies. We describe each study in more detail below (Petrosino, et al. 2000). Appendix 2 summarizes information on each of the nine studies.

MICHIGAN DEPARTMENT OF CORRECTIONS, USA, 1967. In an internal, unpublished government document, the Michigan Department of Corrections reported testing a program that took juvenile delinquents on two tours of a state reformatory. Sixty juvenile delinquents were randomly assigned to attend two tours of a state reformatory or to a no-treatment control group. Tours included 15 juveniles at a time. No other part of the program is described. Recidivism was measured as either a petition in juvenile court for either a new offense or a violation of an existing probation order. At six months, 43% of the experimental group recidivated, compared to only 17% of the control group. Curiously, more attention is not provided to this large negative result in the original document.

THE GREATER EGYPT PLANNING AND DEVELOPMENT COMMISSION, ILLINOIS, USA, 1979. The Greater Egypt Regional Planning and Development Commission (1979) examined the effects of a Scared Straight program in Illinois with a no treatment control group. This Menard Correctional Facility program started in 1978 and is described as a frank and realistic portray of adult prison life. The researchers randomly assigned 161 youths between the ages of 13-18 to attend the program or a no-treatment control. The participants were a mix of delinquents or children at-risk of becoming delinquent. Again, the outcomes are negative in direction, with 17% of the experimental participants being recontacted by police in contrast to 12% of the controls. The authors concluded that "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" (p. 19). Researchers report no effect for program on Jesness Inventory or Piers Harris Self-Concept Scale. In contrast, interview and mail surveys of participants and their parents and teachers indicated unanimous support for the program (p.12). Researchers also note how positive and enthusiastic inmates were about their efforts.

MICHIGAN JOLT STUDY, USA, 1979. Yarborough (1979) reported the second experimental study conducted in Michigan, this time of the Juvenile Offenders Learn Truth or JOLT program, which began in 1978. In the JOLT program, juvenile delinquents in contact with one of four Michigan county courts participated. Inmates administered the program twice per week. 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. Yarborough compared the study participants on a variety of crime outcomes collected from participating courts at three and six month follow-ups. Although the differences were small and varied across these outcomes, most results were again in the direction of favoring the control group. For example, at six months, Yarborough reported that 31% of the experimental group committed new criminal offenses compared to 29% of the controls. The average offense rate for program participants was .69 compared to .47 for the control group. Yarborough (p. 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."

VIRGINIA INSIDERS PROGRAM, USA, 1981. Orchowsky and Taylor (1981) presented the only positive results from the Scared Straight experiments. The Insiders Program was described as an inmate-run, confrontational intervention with verbal intimidation and graphic descriptions of adult prison life. Juveniles were locked in a cell 15 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 80 juveniles ages 13-20 with two or more prior adjudications for delinquency to the Insiders program or a no treatment control group. Orchowsky and Taylor report on a variety of crime outcome measures at six, nine and twelve month intervals. The percentage of juveniles in each group who fail favored the control group at six months (39% of controls had new court intakes versus 41% of experimental participants). The results, however, favored the experimental participants at nine and twelve months. The investigators noted, however, that the attrition rates in their experiment were dramatic: at both nine months (42% of the original sample dropped out) and at twelve months (55% had dropped out). The investigators conducted analyses that seemed to indicate that the constituted groups were still comparable on selected factors, although no analyses were done to indicate if the attendees differed from drop-outs.

TEXAS FACE-TO-FACE PROGRAM, USA, 1981. Vreeland (1981) reported on a factorial experiment to determine the effects of different components of the Texas Face-to-Face juvenile aversion program. The Face-to-Face program began in 1979 and included a thirteen-hour orientation session in which the juvenile lived as an inmate, followed by counseling. Participants were 15-17 years of age, on probation from Dallas County Juvenile Court; most averaged 2-3 offenses before the study. They were randomly assigned 160 boys to four conditions: prison orientation and counseling, orientation only, counseling only, or a no-treatment control group. He examined official court records and self-reported delinquency at six months, finding that the control participants outperformed the three treatment groups on official delinquency (28% delinquent versus 39% for the prison orientation plus counseling, 36% for the prison only, and 39% for the counseling only). The self-report measure, however, showed the reverse. All three treatment groups had similar proportions of participants who self-reported offenses (59%), whereas 69% of the control group self-reported offenses. Vreeland found that there were discrepancies between the self-report and official data; some who were officially charged did not self-report the offense and vice-versa. He seems to place more confidence that the official data captures more harmful offenses by participants in the study, stating that "official records have been shown to be reasonably accurate with respect to the more serious crimes of persistent delinquents" (p. 24). Viewing all the data, Vreeland concluded that there was no evidence that Face-to-Face was an effective delinquency prevention program. He finds no effect on several attitudinal measures, including the "Attitudes Toward Obeying Law Scale."

NEW JERSEY SCARED STRAIGHT PROGRAM, USA, 1982. Finckenauer (1982) conducted the most famous experiment testing the Scared Straight program. The New Jersey Lifers' Program began in 1975 and stressed confrontation with groups of juveniles ages 11-18 who participated in a rap session. Finckenauer randomly assigned 82 juvenile delinquents and non-delinquents (but considered at-risk) to the program or to a no treatment control group. He then followed them for six months in the community. Using official court records, he reported that 41% of the kids who attended Scared Straight committed new offenses, while only 11% of controls did. He also found that the program participants committed more serious offenses. He also reports no impact of the program on nine attitude measures except one: experimentals do much worse on a measure called "attitudes toward crime."

CALIFORNIA SQUIRES PROGRAM, USA, 1983. Lewis (1983) provided some more evidence of a possible harmful effect in his evaluation of the San Quentin Utilization of Inmate Resources, Experience and Studies (SQUIRES) program. This is supposedly the oldest such program in the USA, beginning in 1964. The SQUIRES program included male juvenile delinquents from two California counties between the ages of 14-18, 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 20 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. Though a number of different measures were used, Lewis reported that 81% of the program participants were arrested compared to 67% 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 experimentals and 3.3 months for controls. Data were reported on 8 attitudinal measures, and Lewis reported that the program favored the experimental group on all of them.

KANSAS JUVENILE EDUCATION PROGRAM, USA, 1986. Locke and his colleagues (1986) reported little effect of the Juvenile Education Program in the Kansas State Prison. This intervention was designed to educate children about the law and the consequences of violating it. The program also tried to match juveniles with inmates. Fifty-two juvenile delinquents age 14-19 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. Both groups improved from pretest to posttest but 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 Cerkovich psychological tests.

MISSISSIPPI PROJECT AWARE, USA, 1992. Finally, Cook and Spirrison (1992) reported on an experimental evaluation of Mississippi's Project Aware program. Project Aware was a non-confrontational, educational program lasting only one five-hour session run by prisoners. The intervention was delivered to juveniles in aggregate groups numbering from 6-30. In the study, 176 juveniles between 12-16 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 12 and 24 months. Most of the findings favored the control participants, but again the differences between the groups were small. For example, the mean offending rate for controls at 12 months was 1.25 for control cases versus 1.32 for Project Aware participants. Both groups improved from 12 to 24 months, but the control mean offending rate was still 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" (p. 97). The investigators also reported on two educational measures: school attendance and drop-out. Curiously, they report an effect for the program on school dropout data, but note that "...it is not clear how the program succeeded in reducing dropout rates..." (p.97).

## **Assessment of Methodological Quality of the Original Studies**

Table 1 provides our summary of the methodological quality of the original studies. Although there are many important methodological issues and different types of validity, from a reviewers' perspective, internal validity should be accorded great weight. This is because the reviewer must have some confidence that the program is responsible for the observed results. Given that we are only considering randomized studies, four factors seem most important in our confidence about the causal connection between the program and the result. These characteristics are considered the most critical in determining if the results from the studies can be believed.

(1) Randomization Integrity. Did the investigators report that randomization of participants to experimental or control conditions experience any violation or subversion of random assignment procedures? If so, did they report how many cases were incorrectly assigned to conditions? Research has shown that subversion of random assignment protocol could lead to groups that are not equivalent at the start of the experiment and bias one group toward a better performance than the other (Dennis 1988).

(2) 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? Attrition can result in the final analysis sample being comprised of two groups that are no longer equivalent on sociodemographic factors (Boruch 1997).

(3) Outcome Bias. Did the investigators report that any steps were taken to "blind" either the practitioners enrolling the kids in the program or the police or other agencies responsible for collecting the outcome data to treatment assignment? In his reanalysis of the seminal Community Treatment Project (CTP) in California, Lerman (1975) reported that CTP's impact on parole failure rates was not due to the program's impact on offending. Instead, Lerman reported that parole officers were technically violating control participants and overlooking the same behavior in the CTP group. Lerman urges us to be skeptical of efforts that involve the program designers and planners also collecting the outcome data.

(4) Program Implementation. Did the investigators report that the program was implemented? Although programs are never implemented perfectly, some interventions are so poorly designed, planned and delivered that they comprise no treatment at all. In addition, the implementation may have been so poor that the treatment rivals the control group in duration, intensity and content of intervention.

As indicated in Table 2, juvenile awareness programs appear relatively straightforward to implement. Not a single report indicated any problem with establishing and running the program during the study period. These evaluations, however, were far from perfect. Some of the problems are a result of inadequate reporting of study details. Although they employed random assignment, far too little information was given about how randomization was done, whether the groups were equivalent on major factors after randomization (and after attrition), and why the groups ended up with unequal sample sizes. Ironically, only the earliest study – the 1967 Michigan Department of Corrections – blinded those collecting outcome data to ensure that the results were not tampered with. Only Finckenaue's 1982 study experienced dramatic violation of randomization protocol and analyzed the participants as assigned. Cook and Spurrison (1992) removed such violations from their sample.

Several of the studies experienced a significant loss of participants from the original randomly assigned sample.

**Table 2. Methodological Characteristics and Problems Reported in the Studies**

<i>Study (Total N)</i>	<i>Randomization</i>	<i>Attrition</i>	<i>Outcome Bias</i>	<i>Implementation</i>	<i>Methodological Summary</i>
<b>Michigan Department of Corrections 1967 (60)</b>	Random numbers tables used to allocate, no test for equivalence reported.	Only two participants lost.	Juvenile home records used in follow-up; data investigators were blind to group allocation	No problems reported.	The one troubling aspect is the failure to conduct a test for equivalence, particularly with only 60 total persons assigned. Nonetheless, there is nothing else to question the observed findings.
<b>GEP&amp;DC 1979 (161)</b>	Random assignment, no further information	None reported	Subsequent police reports, no problems reported.	No problems reported.	Nothing in the report seems to indicate that the findings should be questioned.
<b>Yarborough 1979 (227)</b>	Research unit handled random assignment, good protocol in place, test for equivalence satisfactory	The study has many no-shows whom are dropped from analysis	Researchers collected data from court files but unknown if blind to conditions. Government agency still reported a rather negative result for its own program.	No problems reported.	The no-shows and its lack of attention in the report trouble us. Again, nothing in the report suggests anything other than a null or slightly negative effect for JOLT.
<b>Orchowsky &amp; Taylor 1981 (80)</b>	Random assignment used, test for equivalence satisfactory	The study drops 41% at 9 months and 55% at 12 months, PIs report tests for equivalence at 9 and 12 months are satisfactory	Juvenile court intake data is the primary source but no description on how collected.	No problems reported.	The massive attrition at 9 and 12 months also corresponds with positive results reported for the program after negative impact at 6 months. The tests for equivalence, however, do seem to indicate the groups were still comparable.
<b>Vreeland 1981 (79)</b>	Random assignment used, test for equivalence satisfactory	No attrition for the two groups (of the 4 in the experiment) reported	Used court data and self-report, no other information provided	No problems reported.	There is nothing in the report to lead us to question the findings.

<b>Finckenaue 1982 (81)</b>	Randomization broke down, 6 of the 11 referral agencies violated assignment protocol, test for equivalence showed 59% of E had a prior record, only 40% of C	None reported.	Researchers collected the data from court files, not program staff.	No problems reported.	Randomization breakdown is cause for concern. PI does report additional analyses for agencies that followed protocol: 31% of E recidivated compared to 17% of C.
<b>Lewis 1983 (108)</b>	Test for equivalence is satisfactory but age slightly favors experimental group	Only one participant lost during follow-up	Although the CA Youth Authority ran the program and collected the data, they report negative effects for the program.	No problems reported.	There is nothing in the study report to support any lack of confidence in the observed findings.
<b>Locke et al. 1986 (53)</b>	Randomization used, test for equivalence satisfactory (though not stated if done after attrition)	40% of an already small sample lost in follow-up, leaving 32 in the study	Two researchers collected court data.	No problems reported.	The study appears to have severe attrition, limiting our confidence. The PIs report no effect for treatment but do not provide enough data for computation of odds ratios or weighted mean differences.
<b>Cook &amp; Spirison 1992 (176)</b>	Quasi-random allocation using odd-even assignment of case files (with initial numbering quasi-random – all cases numbered consecutively). Some breakdown is reported but actual percentage is unknown; cases were dropped.  No test for equivalence reported before or after attrition.	24% lost in follow-up, no analysis to ensure groups still equivalent	Data retrieved from court system. No other information provided.	No problems reported	The attrition gives us cause for concern, particularly with no tests for equivalence. But the major problem with the study is the failure of the investigators to report the necessary standard deviations for the meta-analysis. All available data seem to indicate a slightly negative impact for the program on crime measures.

## Meta-Analysis

For each study, we extracted all of the relevant crime outcome data. Given the limitation of the data, we conducted a meta-analysis following two assumptions. Our first is to focus on official measures of criminality. Although Locke et al. do not present the raw data, both they and Vreeland (1981) found wide discrepancies between self-report measures and official data. Juveniles were more likely to self-report less serious offenses but less likely to report serious offenses. In other words, the more serious the offending, the likelier that the official measures were more accurate.

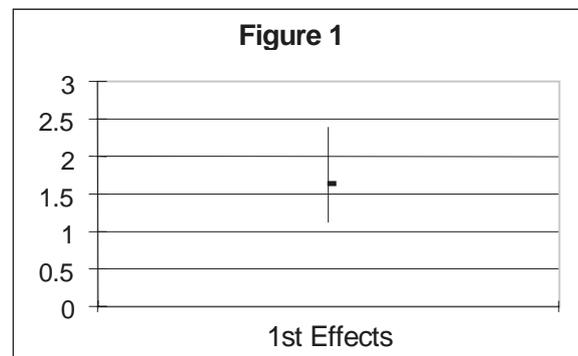
The “Sellin rule“ was the basis for our second assumption. Thorsten Sellin was a famous criminologist who wrote that data from earlier points in the criminal justice system more accurately reflect offending behavior than those captured later in the system (Petrosino 1997). This is because data from later points in the system likely reflect administrative and legal decisions rather than offending conduct. Thus, police data are preferred to court data, and court data are preferred over parole or incarceration data. We followed this in the analyses below, using police contact or arrest data when available. We then computed three analyses:

- (1) first effects using official measures of crime
- (2) the average effect for each study (averaged across different intervals)
- (3) the best effect reported within each study

Because most data were dichotomous and comprised of failure rates, our first approach involved computing Odds Ratios for each study, and their 95% Confidence Intervals (CI). Because there is debate about the ideal assumption to hold about the population of meta-analytic samples, we assumed both random and fixed effects models for weighting the treatment effects across the studies, for each of the three analyses above. We used Excel software to compute the analyses. Appendix 2 provides a more detailed summary of our statistical approach.

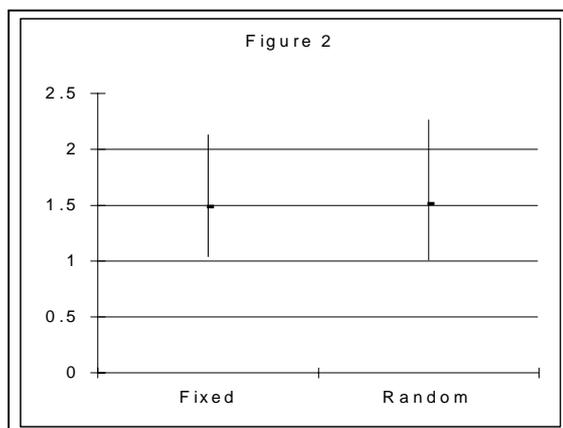
### Analysis 1. First effects for Prevalence Rates

As Figure 1 shows, the seven studies reporting prevalence rates using official data show that intervention increases the crime at the first follow-up period. Interpretation is straightforward: “1“ is a zero impact. Anything above “1“ is negative, i.e. it increases the odds of committing crime. Anything below “1“ decreases those odds. The results in Figure 1 were identical whether we assumed a fixed or random effects model. The mean Odds Ratio (OR) was 1.64, and the 95% confidence intervals were 1.13-2.39. Although the studies were comprised of small samples, overall this negative result is statistically significant.



### Analysis 2. The mean effect across prevalence rates at different intervals

Although the “first effect“ is usually the one that should represent the greatest opportunity for treatment to show a positive effect – particularly deterrence-oriented programs like those reported here – it is also true that studies sometimes report prevalence data at later follow-up intervals. It is also true that these data vary. Some argue that a conservative approach is to average these effects to create one “mean“ effect measure for each study. In this analysis, we again used the prevalence data for the seven studies,

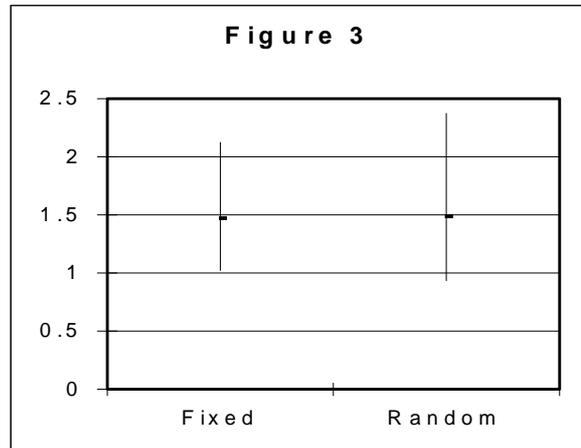


but averaged it for studies with multiple follow-up intervals. For example, Orchowsky and Taylor (1981) report three distinct outcome intervals (6, 9, and 12 months). We averaged these to create one mean or average effect for that study.

As Figure 2 shows, the mean effect for the seven studies reporting official prevalence data is still negative for intervention, whether we assume a fixed or random effects model. For the fixed effects model, the mean OR is 1.49 (95% CI 1.04-2.13); the mean OR assuming a random effects model is 1.51 (CI 1.01-1.51). Again, this negative impact for intervention is statistically significant when the seven studies are combined.

Analysis 3. The best effect using official prevalence rates

Although both the first effect and mean effect analyses shows intervention to be harmful, our last analysis using OR gives the program the benefit of the doubt. We report the best effect from each study. As Figure 3 shows, the mean OR remains negative (i.e. above 1) whether a fixed (OR 1.47, CI 1.02-2.12) or random (OR 1.49, CI .93-2.97) effects model is assumed. Only when the random effects model is assumed does the lower CI dip below 1 (at .93) but the spread across the seven studies is much greater. The meta-analysis assuming the random effects model is not significant.



Additional analysis. In Table 2, we computed Cohen’s effect size measure (*d*) for each outcome in which the computation was possible. This was done using d-Stat meta-analytic software (Johnson 1989). We then organized these data according to our original protocol: by prevalence (p), incidence (I), severity (s) and latency (l). We then organized these data by follow-up intervals: 3, 6, 9, 12 and 24 months. We also planned to report separately on self-report and official data. When the report simply stated an effect was found or provided raw data without necessary information to compute Cohen’s *d*, we include it in Table 2. The reason for using Cohen’s effect size is that there are formula available to approximate *d* from test statistics, probabilities and other data even when standard deviations or dichotomous proportions are not available. It is also a relatively straightforward interpretation: negative effect sizes indicate the intervention group does worse than the control group; positive effect sizes indicate the intervention group does better than the control. The larger the effect size, the larger the program effect whether positive or negative in direction. As a baseline for comparison, Lipsey (1992) found that the nearly 400 experimental and quasi-experimental evaluations of juvenile delinquency prevention and treatment programs averaged an effect size of .10 (approximately 10% reduction in recidivism if a baseline of 50% for the control is assumed).

Table 2 also reports positive effects by using gray in the background to help them to stand out from negative effects. Other than the Orchowsky and Taylor follow-ups at 9 and 12 months, almost all the effect sizes are negative in direction. Contrast this again with Lipsey’s (1992) study: 66% of the effect sizes at first follow-up were positive indicating treatment reduced subsequent offending.

Again, there is little to convince us that these programs have anything other than a null or negative impact on crime, regardless of index or follow-up period.

Table 2 clearly shows that few studies reported more than one or two of these important indicators. For example, only Vreeland (1981) included data on self-reported crime. In two studies (Cook and Spirrisson 1992; Locke et al. 1986), no prevalence rates were reported. Some of the studies that did include incidence data (e.g. the average or mean rates of offending) did not include standard deviations to make it possible to compute effect sizes. The follow-up periods were diverse.

## DISCUSSION

These randomized trials, conducted over a 25-year period in eight different jurisdictions, provide evidence that Scared Straight and other "juvenile awareness" programs are not effective as a stand-alone crime prevention strategy. More importantly, they provide some 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, ranging from harsh, confrontational interactions to tours of the facility converge on the same result: an increase in criminality in the experimental group when compared to a no-treatment control. It appears that doing nothing would have been better than exposing juveniles to these programs.

Doing harm has a different implication in this case than in medicine. In the case of the Cochrane review of human albumin solution (Alderson et al 2001), for example, the reviewers reported that the intervention increased mortality. Although indirect or psychic harms were probably bestowed on the families of patients, the patients under treatment experienced the direct harm. In the case of criminal justice, increasing harm not only means that the persons exposed to an intervention do worse - it also means that more crimes are committed in the community against others who were not exposed to the treatment and did not consent to it. Doing more harm than good in criminal justice puts more persons at risk for victimization or the direct negative impacts of criminal behavior.

### Reviewers' conclusions

#### *Implications for practice*

As Finckenauer and Gavin (1999) have reported, these programs remain in use and will continue. Every jurisdiction believes its program is the one that runs counter to this research - that they're "doing it better." At the very least, agencies that allow this program have a duty to rigorously test it. This is not only to ensure that their program is doing what it claims, i.e. preventing juvenile delinquency, but also to determine that the intervention is not doing more harm than good.

#### *Implications for research*

Some of the literature indicates the program has rehabilitative effects on inmate providers and that argument is used to legitimize the use of the program. Others claim that the program by itself is of little value but could be instrumental if imbedded in an overall multi-component package of interventions delivered to youths. These assertions must be exposed to rigorous tests, including

randomization of participants to conditions. Such an evaluation ethically must include a short-term follow-up so that the program can be terminated quickly should early results indicate that the program is harmful.

## **REFERENCES TO STUDIES**

### **References to included studies**

Cook DD, Spirrison CL. Effects of a prisoner-operated delinquency deterrence program: Mississippi's Project Aware. *Journal of Offender Rehabilitation* 1992;17:89-99.

Cook DD. Effects of a non-confrontational prisoner-run juvenile delinquency deterrence program. Unpublished Masters thesis, Department of Psychology, Mississippi State University 1990.

Finckenauer JO. *Scared Straight and the Panacea Phenomenon*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

Greater Egypt Regional Planning & Development Commission. *Menard Correctional Center: Juvenile tours impact study*. Carbondale, IL: author 1979.

Lewis RV. Scared straight--California style: evaluation of the San Quentin SQUIRES program. *Criminal Justice & Behavior* 1983;10(2):209-226.

Locke TP, Johnson GM, Kirigin-Ram K, Atwater JD, Gerrard M. An evaluation of a juvenile education program in a state penitentiary. *Evaluation Review* 1986;10:281-298.

Locke TP. *An Analysis of the Kansas State Penitentiary Juvenile Education Program*. Unpublished Masters thesis, Department of Psychology, University of Kansas 1982.

Michigan Department of Corrections. *A six month follow-up of juvenile delinquents visiting the Ionia Reormatory*. Research Report No. 4. Lansing: Michigan Department of Corrections 1967.

Orchowsky S, Taylor K. *The Insiders juvenile crime prevention program*. Richmond: Virginia Department of Corrections 1981.

Vreeland AD. *Evaluation of Face-to-Face: A juvenile aversion program*. Unpublished doctoral dissertation, University of Texas, Dallas 1981.

Yarborough JC. *Evaluation of JOLT as a deterrence program*. Lansing, Michigan: Michigan Department of Corrections 1979.

### **References to excluded studies**

Aims Multimedia. *Scared Straight! 20 years later* [film]. Clatsworth, California, Author 1999.

Ashcraft JL. *Closeout report on the Orleans Parish (LA) Juvenile Awareness Program*. New Orleans, LA: Mayor's Criminal Justice Coordinating Council 1979.

Berry RL. Shape up: the effects of a prison aversion program on recidivism and family dynamics. PhD dissertation, Department of Psychology, University of Northern Colorado 1985.

Brodsky SL. The prisoner as an agent of attitude change: a study of prison profiles' effects. *British Journal of Criminology* 1970;10:280-285.

Buckner JC, Chesney-Lind M. Dramatic cures for juvenile crime: An evaluation of a prisoner-run delinquency prevention program. *Criminal Justice & Behavior* 1983;10(2):227-247.

Chesney-Lind M. 'Ike Na Pa'ahao: the experience of the prisoners. A juvenile awareness program. Report No. 258. Manoa, Hawaii: University of Hawaii at Manoa, Youth Development & Research Center 1981.

Dean DG. The impact of a juvenile awareness program on select personality traits of male clients. *Journal of Offender Counseling, Services and Rehabilitation* 1982;6(3):73-85.

Gilman L, Milin RK. An evaluation of the Lifers' Group Juvenile Awareness Program. New Jersey State Prison, Rahway, New Jersey. Trenton, NJ: Department of Corrections 1977.

Langer S. Fear in the deterrence of delinquency: A critical analysis of the Rahway State Prison Lifer's Program. Phd Dissertation, Criminal Justice, Rutgers University, Newark, New Jersey 1980.

Lloyd C. To scare straight or educate? The British experience of day visits to prison for young people. Home Office Research Study No. 149. London, UK: Home Office 1995.

Mitchell JJ, Williams SA. SOS: Reducing juvenile recidivism. *Corrections Today* 1986;48(3):70-71.

Nelson Z. The Day in Prison Program. The First Twelve Months.. Victoria, Australia: Community based Corrections Division, Victoria Office of Corrections 1991.

New South Wales Bureau of Crime Statistics and Research. Day in Gaol Programme: Research Report No. 8. Author 1980.

Nygaard CH. Synopsis of the findings, conclusions and recommendations from the evaluational study of the Department's Youth Assistance Programs. New York State Department of Correctional Services, Division of Program Planning, Research and Evaluation 1980.

O'Malley P, Coventry G, Walters R. Victoria's Day in Prison Program: An evaluation and critique. *Australian and New Zealand Journal of Criminology* 1993;26(2):171-183.

Portnoy RN. Ego defenses as a predictor of deterrence. Ph.D. dissertation, Department of Psychology, University of Nebraska 1986.

Rasmussen DW, Yi Y. An evaluation of juvenile justice innovations in Duval County, Florida. Tallahassee, FL: Florida State University 1996.

Shapiro A. Scared Straight [film]. Santa Monica, CA: Pyramid Films 1978.

Storvall AE, Hovland A. Ullersmoprojektet: "Scared Straight" I Norge 1992-1996 [The Ullersmo Project: Scared Straight in Norway 1992-1996]. *Nordisk Tidsskrift for Kriminalvidenskab* 1998;85(2):122-135.

Trotti TC. Save the Children Program - An Analysis of Client Reaction. Columbia, SC: South Carolina Department of Youth Services, Research and Evaluation Unit 1980.

### **Additional references**

Andrews DA, Zinger I, Hoge RD, Bonta J, Gendreau P, and Cullen FT. Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis.. *Criminology* 369-404;28(3).

Finckenauer JO. Scared Straight and the Panacea Phenomenon. Englewood Cliffs, NJ: Prentice-Hall, 1982.

Finckenauer JO, Gavin PW, Hovland A, Storvoll E. Scared Straight: The Panacea Phenomenon Revisited. Prospect Heights, Ill. USA: Waveland Press, 1999.

Greater Egypt Regional Planning and Development Commission. Evaluation Report: Menard Correctional Center Juvenile Tours Impact Study. Carbondale, IL: Greater Egypt Regional Planning and Development Commission, 1979.

Hedges L, Olkin I. *Statistical Methods for Meta-Analysis*. 1985.

Lewis RV. Scared straight--California style: Evaluation of the San Quentin Squires program. *Criminal Justice and Behavior* 1983;10(2):209-226.

Lipsey MW. *Design Sensitivity*. Newbury Park, CA: Sage, 1990.

Lipsey MW. Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In: Cook TC, Cooper H, Cordray DS, Hartmann H, Hedges LV, Light RL, Louis TA, and Mosteller FM, editor(s). *Meta-Analysis for Explanation*. New York: Russell Sage, 1992:83-127.

Lipton D, Martinson R, Wilks J. *The Effectiveness of Correctional Treatment*. New York, NY: Praeger, 1975.

Petrosino AJ. 'What Works?' Revisited Again: A Meta-Analysis of Randomized Experiments in Rehabilitation, Deterrence and Delinquency Prevention [dissertation]. Ann Arbor (MI): Rutgers University, 1997.

Petrosino A, Boruch RF, Rounding C, McDonald S, Chalmers I. Assembling a social, psychological, educational and criminological trials register (SPECTR). *Evaluation Research in Education* 2000. forthcoming..

Sherman LW, Gottfredson D, MacKenzie D, Eck J, Reuter P, Bushway S. Preventing Crime: What Works, What Doesn't, What's Promising. A Report to the United States Congress. College Pk, MD: University of Maryland, Department of Criminology and Criminal Justice, 1997.

UPN 1999. Scared Straight! 20 Years Later. Television programme on US television: host Danny Glover (<http://www.contrib.andrew.cmu.edu/~aaron2/upn/upn-specials.html>) 8pm 15 April 1999.

Weisburd DL, Sherman LW, Petrosino AJ. Registry of Randomized Experiments in Criminal Sanctions, 1950-1983. Los Altos, CA: Sociometrics Corporation, Data Holdings of the National Institute of Justice, 1990.

## Appendix 1. Characteristics of excluded studies

Study	Reason for exclusion
Aims Multimedia	Used post-test only data without a control group.
Ashcraft	Used a pre-post test without a control group.
Berry	Used a matched comparison group without randomization.
Brodsky	Used a pre-post design without a control group.
Buckner and Lind	Used a matched comparison group without randomization.
Chesney-Lind	Used a non-equivalent comparison group design without randomization.
Dean	Used randomization but did not include any measures of criminal behavior.
Gilman and Milin	Used archival data from three sources for post-test only follow-ups without a control group.
Langer	Used a matched comparison group without randomization.
Lloyd	Case studies of three day visit programs in the U.K. No control group is included.
Mitchell and William	Used pre-post data without a control group.
Nelson	Used post-test only data without a control group.
NSW Bureau of Statis	Used post-test only data without a control group.
Nygaard	Report on process and implementation data only. No follow-up or control group reported.
O'Malley et al	Process and implementation data on Australia's Victoria prison program. No control group.
Portnoy	This study randomly assigned juveniles from high school to watch the Scared Straight video or a more neutral film. It did not involve the actual program. No follow-up data on criminal offenses were reported.
Rasmussen and Yi	Used multivariate regression on county crime rates to estimate prevention impact of program, no control group or randomization employed.
Shapiro	Used post-test only data without a control group.
Storvoll and Hovland	Process and implementation data are reported on Norway's Scared Straight program. No follow-up or control group included.
Trotti	Used post-test data of reactions of participants, without a control group.

## Appendix 2. Characteristics of included studies

Study	Methods	Participants	Interventions	Outcomes	Notes
Cook and Spirrison	Quasi-random assignment - researchers numbered court files and assigned all odd numbered ones to intervention group	176 juvenile delinquents ages 12-16 under jurisdiction of one Mississippi county youth court, 36% white, 100% male	Educational, prisoner-run 5 hour session, designed to be non-confrontational	12 and 24 month follow-ups of official court record data, average offending rates, and severity of offense. School attendance and school drop-out.	No standard deviations reported with any mean data, no group percentages, attempts to retrieve these data from author and other primary documents failed.
FINCKENAUER	Random assignment	81 delinquent or children ages 11-18 at-risk for delinquency, 50% had prior record of offending, 40% were white, 80% male	One visit, a confrontational rap session lasting approximately 3 hours with inmates serving life sentence	Six month follow-up of official complaints, arrests or adjudications. Severity of offense. Attitudes: -toward criminals -toward crime -toward law -toward justice -toward police -toward prison -toward punishment -self-image	
GERP&DC	Random assignment	161 delinquent or children at-risk for delinquency, 100% male, 84% white, ages 13-18	Confrontational rap session with inmates	5-15 months follow-up of contacts with police Piers Harris Children's Self-Concept Scale Jesness Inventory	
LEWIS	Random assignment	108 juvenile delinquents from two California counties, most with extensive prior record, ages 14-18, 100% male, mostly non-white	Three total visits (one per week) including confrontational rap sessions, guided tours of prison and interaction with prisoners, review of pictures of prison violence	Twelve month follow-up of percentage arrested, average number of arrests, percentage charged, average number of charges, charges by type of offense, offense severity, time to first arrest Attitudes: -toward police -toward school -toward crime -toward prison -toward work camp Semantic Differential Test	Over 100 moderating analyses performed on the data.
LOCKE et al	Random assignment	53 juvenile delinquents ages 14-19 on probation from three Kansas counties, 65% white, 100% male	Non-confrontational, educational interaction, tried to match juvenile with inmate	Minimum six month follow-up of self-reported crime and juvenile court and police records of official offending	No standard deviations reported with any mean data, no group percentages, attempts to retrieve these data from author and other primary documents failed.
MICHIGAN D.O.C.	Assignment using random numbers table, data collectors were blind to assignment	60 juvenile delinquents from one Michigan county	Two tours of a Michigan reformatory	Six month follow-up of official petition for delinquency or probation violation	Brief internal report that does not fully describe nature of intervention.
Orchowsky and Taylor	Random assignment	80 juvenile delinquents (with minimum two offenses), ages 13-	Confrontational, inmate-run program, locked in cell,	Six, nine and twelve month follow-ups of official measures of offending including new	

Study	Methods	Participants	Interventions	Outcomes	Notes
		20, 100% male	introduction by guard, two hour session with inmates	court intakes, average number of court intakes, severity of offense	
Vreeland	Randomly assigned to four groups	160 juvenile delinquents given probation by Dallas County Court, 100% male, 40% white, ages 15-17, averaged 2-3 prior offenses	one-day orientation lasting 13 hours, including haircut and physical labor	Six month follow-up of official (using court records) and self-reported data to establish percentage offending Attitude toward Law Friend Survey Deterrence questionnaire Self-image Jesness Checklist	To remain consistent with other interventions in this review, we took the orientation group comparison with the no-treatment control group. The orientation plus counseling group, however, was almost identical to the orientation only group in final results.
Yarborough	Researchers randomly assigned participants according to random numbers table	227 juvenile delinquents under jurisdiction of courts in four Michigan counties	Tour of facility, separated and take to cell for interaction with inmates, confrontational session with inmates, one visit five hours long	Three and six month follow-ups of official juvenile crime as measured by subsequent court petitions, new offenses, average offense rate, weeks to new offense, type of offense charged, average days in detention	Extensive moderating analyses done.



**Table 2. Full Array of Outcome Measures Transformed To Cohen's Effect Size (d)**

Study	Type of Data	3 MONTHS				6 MONTHS				9 MONTHS				12 MONTHS				24 Months	
		P	I	S	L	P	I	S	L	P	I	S	L	P	I	S	L		
MI DOC 67	Court petition or probation violated					-.60													
GEP&DC 79	Police contacts					-.14													
Yarborough 79	New offenses	-.03	-.17	-.14	-.13	-.04	-.18	-.09	.10										
	Petitions	.03				.04													
	Mean days in detention		.52				-.02												
Orchowsky & Taylor 81	New intakes					-.04	.15	.29		.50	.68	.82		44	.83	.83			
Vreeland 81	Officially recorded delinquency					-.17													
	Self-reported delinquency					.21													
Finckenaer 82	Complaints, arrests and adjudications					-.71		-.60											
	Correction for randomization failure					-.33													
Lewis 83	Arrests													-.32	E2.1, C2.2	NS	E4.1m, C3.3m (S)		
Locke et al 86	Juvenile court records						F=.75, unknown direction												
	Self-reported offending						E does worse, but no data provided												
Cook & Spirrison 92	Reoffending (F test of E v C over time)																		-.09 I

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)