



Bridgewater State University

Virtual Commons - Bridgewater State University

Criminal Justice Faculty Publications

Criminal Justice Department

2013

'Scared Straight' and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency

Anthony Petrosino

Carolyn Turpin-Petrosino
Bridgewater State University, cpetrosino@bridgew.edu

Meghan E. Hollis-Peel
Northeastern University

Julia G. Lavenberg
University of Pennsylvania

Follow this and additional works at: https://vc.bridgew.edu/crim_fac



Part of the [Criminology Commons](#)

Virtual Commons Citation

Petrosino, Anthony; Turpin-Petrosino, Carolyn; Hollis-Peel, Meghan E.; and Lavenberg, Julia G. (2013). 'Scared Straight' and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency. In *Criminal Justice Faculty Publications*. Paper 32.
Available at: https://vc.bridgew.edu/crim_fac/32

This item is available as part of Virtual Commons, the open-access institutional repository of Bridgewater State University, Bridgewater, Massachusetts.



Cochrane
Library

Cochrane Database of Systematic Reviews

'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency (Review)

Petrosino A, Turpin-Petrosino C, Hollis-Peel ME, Lavenberg JG

Petrosino A, Turpin-Petrosino C, Hollis-Peel ME, Lavenberg JG.

'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency.

Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD002796.

DOI: 10.1002/14651858.CD002796.pub2.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	4
METHODS	4
RESULTS	6
Figure 1.	11
DISCUSSION	14
AUTHORS' CONCLUSIONS	15
ACKNOWLEDGEMENTS	16
REFERENCES	17
CHARACTERISTICS OF STUDIES	20
DATA AND ANALYSES	32
Analysis 1.1. Comparison 1 Intervention versus control, crime outcome, Outcome 1 Postintervention - group recidivism rates - official measures only (fixed-effect).	33
Analysis 1.2. Comparison 1 Intervention versus control, crime outcome, Outcome 2 Postintervention - group recidivism rates - official measures only (random-effects).	34
Analysis 1.3. Comparison 1 Intervention versus control, crime outcome, Outcome 3 Sensitivity analysis - excluding Finckenauer study.	35
Analysis 1.4. Comparison 1 Intervention versus control, crime outcome, Outcome 4 Sensitivity analysis - excluding Yarborough study.	36
Analysis 1.5. Comparison 1 Intervention versus control, crime outcome, Outcome 5 Sensitivity analysis - excluding both Finckenauer and Yarborough studies.	37
ADDITIONAL TABLES	37
APPENDICES	39
FEEDBACK	41
WHAT'S NEW	42
HISTORY	42
CONTRIBUTIONS OF AUTHORS	43
DECLARATIONS OF INTEREST	43
SOURCES OF SUPPORT	43
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	44
NOTES	44
INDEX TERMS	44

[Intervention Review]

'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency

Anthony Petrosino¹, Carolyn Turpin-Petrosino², Meghan E Hollis-Peel³, Julia G Lavenberg⁴

¹Learning Innovations, WestEd, Woburn, Massachusetts, USA. ²Criminal Justice, Bridgewater State College, Bridgewater, Massachusetts, USA. ³School of Criminology and Criminal Justice, Northeastern University, Boston, MA, USA. ⁴Penn Medicine Center for Evidence-based Practice, University of Pennsylvania, Philadelphia, USA

Contact address: Anthony Petrosino, Learning Innovations, WestEd, 300 Unicorn Drive, 5th Floor, Woburn, Massachusetts, 01801-3324, USA. anthony_petrosino@hotmail.com. apetros@wested.org.

Editorial group: Cochrane Developmental, Psychosocial and Learning Problems Group.

Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 4, 2013.

Review content assessed as up-to-date: 30 June 2012.

Citation: Petrosino A, Turpin-Petrosino C, Hollis-Peel ME, Lavenberg JG. 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews* 2013, Issue 4. Art. No.: CD002796. DOI: 10.1002/14651858.CD002796.pub2.

Copyright © 2013 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

'Scared Straight' and other similar programs involve organized visits to prison by juvenile delinquents or children at risk for criminal behavior. Programs are designed to deter participants from future offending through firsthand observation of prison life and interaction with adult inmates. These programs remain in use despite research questioning their effectiveness. This is an update of a 2002 review.

Objectives

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

Search methods

To update this review, we searched 22 electronic databases, including CENTRAL, MEDLINE, PsycINFO, and Criminal Justice Abstracts, in December 2011. In addition, we searched clinical trials registries, consulted experts, conducted Google Scholar searches, and followed up on all relevant citations.

Selection criteria

We included studies that tested programs involving the organized visits of delinquents or children at risk for delinquency to penal institutions such as prisons or reformatories. Studies that had overlapping samples of juvenile and young adults (for example, ages 14 to 20 years) were included. We only considered studies that assigned participants to conditions randomly or quasi-randomly (that is, by odd/even assignment to conditions). Each study had to have a no-treatment control condition and at least one outcome measure of 'post-visit' criminal behavior.

Data collection and analysis

The search methods for the original review generated 487 citations, most of which had abstracts. The lead review author screened these citations, determining that 30 were evaluation reports. Two review authors independently examined these citations and agreed that 11 were potential randomized trials. All reports were obtained. Upon inspection of the full-text reports, two review authors independently agreed to exclude two studies, resulting in nine randomized trials. The lead review author extracted data from each of the nine study reports using a specially designed instrument. In cases in which outcome information was missing from the original reports, we made attempts via correspondence to retrieve the data for the analysis from the original investigators. Outcome data were independently checked by a second review author (CTP).

In this review, we report the results of each of the nine trials narratively. We conducted two meta-analyses of seven studies that provided postintervention offending rates using official data. Information from other sources (for example, self-report) was either missing from some studies or critical information was omitted (for example, standard deviations). We examined the immediate post-treatment effects (that is, 'first-effects') by computing odds ratios (OR) for data on proportions of each group reoffending, and assumed both fixed-effect and random-effects models in our analyses.

Main results

We have included nine studies in this review. All were part of the original systematic review; no new trials meeting eligibility criteria were identified through our updated searches. The studies were conducted in eight different states of the USA, during the years 1967 to 1992. Nearly 1000 (946) juveniles or young adults of different races participated, almost all males. The average age of the participants in each study ranged from 15 to 17 years.

Meta-analyses of seven studies show the intervention to be more harmful than doing nothing. The OR (fixed-effect) for effects on first post-treatment effect on officially measured criminal behavior indicated a negative program effect (OR 1.68, 95% confidence interval (CI) 1.20 to 2.36) and nearly identical regardless of the meta-analytic strategy (random-effects OR 1.72, 95% CI 1.13 to 2.62). Sensitivity analyses (random-effects) showed the findings were robust even when removing one study with an inadequate randomization strategy (OR 1.47, 95% CI 1.03 to 2.11), or when removing one study with high attrition (OR 1.96, 95% CI 1.25 to 3.08), or both (OR 1.68, 95% CI 1.10 to 2.58).

Authors' conclusions

We conclude that programs such as 'Scared Straight' increase delinquency relative to doing nothing at all to similar youths. Given these results, we cannot recommend this program as a crime prevention strategy. Agencies that permit such programs, therefore, must rigorously evaluate them, to ensure that they do not cause more harm than good to the very citizens they pledge to protect.

PLAIN LANGUAGE SUMMARY

'Scared straight' and other juvenile awareness programs for preventing juvenile delinquency

Programs such as 'Scared Straight' involve organized visits to prison facilities by juvenile delinquents or children at risk for becoming delinquent. The programs are designed to deter participants from future offending by providing firsthand observations of prison life and interaction with adult inmates. This review, which is an update of one published in 2002, includes nine studies that involved 946 teenagers, almost all males. The studies were conducted in different parts of the USA and involved young people of different races whose average age ranged from 15 to 17 years. Results indicate that not only do these programs fail to deter crime, but they actually lead to more offending behavior. The intervention increases the odds of offending by between 1.6 to 1 and 1.7 to 1. Government officials permitting this program need to adopt rigorous evaluation efforts to ensure that they are not causing more harm to the very citizens they pledge to protect.

BACKGROUND

Description of the condition

Juvenile delinquency, also known as juvenile offending or youth crime, is illegal behavior committed by someone before becoming an adult. The second United Nations Congress on the Prevention of Crime and Treatment of the Offender recommended that the meaning of the term juvenile delinquency should be restricted as far as possible to violations of the criminal law (Kvaraceus 1964). Juveniles are considered to be those persons who have yet to reach age 18 years. Although laws vary across nations, juvenile delinquents, therefore, would be those who have been found guilty (adjudicated) of committing a law violation before they are 18 years of age. A significant percentage of violent and nonviolent offenses are committed by juveniles. For example, in the USA, 15% of all persons arrested by the police for illegal behavior in 2008 were juveniles (US Census 2012). Besides the problem of youth crime, offending as a juvenile is a risk factor for later involvement with the criminal justice system as an adult (McCord 2001). Thus, governments everywhere are looking for effective interventions to address juvenile delinquency. 'Scared Straight' and similar type programs have been used in various places in the world, and offer a low-cost and easy to implement strategy to prevent juvenile delinquency.

Description of the intervention

The basic component of programs such as Scared Straight is organized visits to prison facilities by juvenile delinquents or children at risk for becoming delinquent. Nearly all of these interventions have the juveniles interact with inmates confined in the facility. The most famous of these, 'Scared Straight' in New Jersey (USA), included confrontational 'rap' sessions in which adult inmates shared graphic stories about prison life with the juveniles. Other programs have included less confrontational and more educational sessions, in which inmates shared their life stories and described the choices they made that ultimately led to imprisonment. In the Texas Face-to-Face program, juveniles spent one day living as an adult prisoner and the intervention also included a counseling component.

The most well-known version of the Scared Straight type programs was initiated in the 1970s, as inmates serving life sentences at a New Jersey prison began a program to 'scare' or deter at-risk or delinquent children from a future life of crime. It featured as its main component an aggressive presentation by inmates to juveniles visiting the prison facility. The presentation depicted life in adult prisons, and often included exaggerated stories of rape and murder (Finckenaue 1982). A television documentary on the program aired in 1979 provided evidence that 16 of the 17 delinquents remained law-abiding for three months after attending Scared Straight, and claimed a 94% success rate (Finckenaue

1982). Other data provided in the film indicated success rates that varied between 80% and 90% (Finckenaue 1982). The program received considerable and favorable media attention and was soon replicated in over 30 states nationwide, resulting in special Congressional hearings on the program and the film by the US House Subcommittee on Human Resources (US HCEL 1979).

Scared Straight and other 'kids visit prison' programs are also used in other nations. For example, the 'day in prison' or 'day in gaol' in Australia (O'Malley 1993), 'day visits' in the UK (Lloyd 1995) and the 'Ullersmo Project' in Norway (Storvoll 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 1993). In 1999, 'Scared Straight: 20 Years Later' (UPN 1999; 'Kids and Crooks') was shown on US television and claimed similar results to the 1979 film. In this version, the film reports that 10 of the 12 juveniles attending the program 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 in Germany (Hall 1999) and in Florida (USA) (Rasmussen 1996). Sometimes the program is embedded as one component in a multicomponent juvenile intervention package (Trusty 1995; Rasmussen 1996).

How the intervention might work

The underlying theory of programs such as 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 well known, inmate presentations are now sometimes designed to be more educational than confrontational but with a similar crime prevention goal (Lundman 1993; Finckenaue 1999). Some of these programs feature discussions in which the adult inmates confront and challenge the juveniles about their behavior, also referred to as 'rap sessions'. Programs featuring inmates as speakers who describe their life experiences and the current reality of prison life have a rather long history, in the USA at least (Michigan D.O.C. 1967; Brodsky 1970).

Why it is important to do this review

In 1982, a randomized controlled trial testing the New Jersey program was published, reporting no effect on the criminal behavior of participants in comparison with a no-treatment control group (Finckenaue 1982). In fact, Finckenaue reported that participants in the experimental program were more likely to be arrested.

Other randomized trials reported in the USA also questioned the effectiveness of Scared Straight-type programs in reducing subsequent criminality (GERP&DC 1979; Lewis 1983).

Despite the convergence of evidence from these studies, Scared Straight-type programs remained popular and continued to be used in the USA through the 1990s (Finckenaue 1999). For example, a program in Carson City, Nevada (USA) took 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 has its own program called 'Wisetalk' in which at-risk youth are locked in a jail cell for over one hour with four or five parolees. They claim that only 10 of 300 youngsters exposed to this intervention were re-arrested (U-CAN 2001). In 2001, a group of guards - apparently without the knowledge of administrators - strip-searched Washington DC students during their tours of a local jail under the guise that they were using "a sound strategy to turn around the lives of wayward kids" - claiming the prior success of Scared Straight (Blum 2001). It is not surprising that such programs are popular: they fit with some commonly held notions about how to prevent or reduce crime (by 'getting tough'); they are very inexpensive (a Maryland program was estimated to cost less than USD1 US per participant); and they provide one way for incarcerated offenders to contribute productively to society by preventing youngsters from following the same path (Finckenaue 1982).

In 2000, Petrosino and his colleagues reported on a preliminary systematic review of nine randomized field trials, drawing on the raw percentage differences in each study (Petrosino 2000). They found that programs such as Scared Straight generally increased crime between 1% and 28% in the experimental group when compared to a no-treatment control group. In 2002, our formal Cochrane review was published (Petrosino 2002) (simultaneously as a pilot Campbell Collaboration review), which updated the 2000 work and used more sophisticated meta-analytic techniques. We reported similarly negative findings for Scared Straight and juvenile-awareness programs.

Still, Scared Straight type programs continue. In 2003, then-Governor of Illinois, Rod Blagojevich, signed a bill into law that mandated the Chicago Public School system set up a program called 'Choices' (Swanson 2003). The program would identify students at risk for committing future crime and set up a program to give them 'tours of state prison' to discourage any future criminal conduct (Swanson 2003). More recently, the Arts and Entertainment (A&E) station has been running a weekly series entitled 'Beyond Scared Straight'. Created by the producer of the original Scared Straight program (Arnold Shapiro), the program is now the highest rated in A&E's history (Denhart 2011). The success of the television show has renewed interest in Scared Straight and similar programs as a crime prevention strategy (for example, Denhart 2011), but has also resulted in criticism that it ignores a long his-

tory of scientific evidence (for example, Robinson 2011).

The question about whether Scared Straight and similar programs have a crime deterrent effect is best answered by continued examination of the existing scientific evidence. The current review updates the version published in 2002 and includes new and extended searches to December 2011, as well as additional analyses.

OBJECTIVES

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.

METHODS

Criteria for considering studies for this review

Types of studies

Only studies that used randomization or quasi-random procedures (that is, alternate assignment such as all odd numbered cases to treatment and even numbered cases to control) to assign participants, with or without blinding, were included, provided they had a no-treatment control group.

Types of participants

Only studies involving juveniles, that is children 17 years of age or younger, were included. Participants were delinquents or predelinquents. Studies that contain overlapping samples of juveniles and young adults (for example, ages 13 to 21 years) were also included.

Types of interventions

Only studies that featured as their main component a visit by program participants to a prison facility were included. Programs may include a presentation by the inmates, ranging from graphic (Finckenaue 1982) to educational (Cook 1992). Additionally, programs may feature an orientation session (for example, living as a prisoner for eight hours) or a tour of the facility.

Types of outcome measures

Primary outcomes

The interest of citizens, policy and practice decision-makers, media, and the research community is in whether Scared Straight

and its variations have any crime deterrent effect, therefore crime measures are our primary outcomes. Studies had to report at least one outcome of subsequent offending behavior, as measured by such indices as arrests, convictions, contacts with police or self-reported offenses.

Secondary outcomes

We had no secondary outcomes in our analysis, although 'non-crime' measures (for example, attitudinal, educational) reported by the primary investigators are included in [Table 1](#) to enable review authors in the Cochrane and Campbell Collaborations to identify potentially eligible studies for their systematic reviews.

Search methods for identification of studies

To minimize publication bias, 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, that 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. The search methods for the original review are described in detail in [Appendix 1](#).

In December 2011 we searched 11 of the 16 previously searched databases, and expanded our searches to include an additional nine bibliographic sources. We searched all available years of the additional sources, and limited the search of the databases used previously to 2001 onwards. The five databases not searched for this update included one that was no longer accessible (C2-Spectr), and four that produced zero yield in the previous searches (Current Contents, GPO Monthly, National Clearinghouse of Child Abuse and Neglect (NCCAN) abstracts, and Political Science Abstracts). In November 2012 we also searched two trials registers. The 22 databases searched during the update were:

- Cochrane Central Register of Controlled Trials (CENTRAL), 2011(4), searched December 2011
- Academic Search Premier, all available dates to December 2011
- Ovid MEDLINE, 2001 to December 2011
- Clinical Trials.Gov, all available dates, searched November 2012
- Criminal Justice Abstracts, 2001 to December 2011
- Directory of Open Access Journals, all available dates to December 2011
- Dissertations and Theses (ProQuest), which covers Dissertation Abstracts, 2001 to December 2011
- Education FullText, 2001 to December 2011
- ERIC (Proquest), 2001 to December 2011
- Google Scholar, all available dates, searched December 2011
- HeinOnline, all dates to December 2011

- Illinois Researcher Information Service (IRIS), all dates to December 2011
- International Bibliography of the Social Sciences, 2001 to December 2011
- National Criminal Justice Reference Service Abstracts Database (NCJRS), 2001 to December 2011
- Public Affairs Information Service (PAIS), 2001 to December 2011
- PsycArticles, all dates to December 2011
- PsycINFO, 2001 to December 2011
- SCOPUS Science Direct, all dates to December 2011
- Scandinavian Research Council for Criminology, all dates to December 2011
- Sociofile, including Sociological Abstracts and Social Planning and Development Abstracts, 2001 to December 2011
- SSCI (Web of Science), which includes the Social Science Citation Index (SSCI), 2001 to December 2011
- World Health Organization International Clinical Trials Registry Platform (ICTRP), searched November 2012

Our keywords were similar to those used in the previous two searches. A list of search terms is provided in [Appendix 1](#).

We also contacted an informal list of researchers in the field, and examined citations in relevant literature, including previous systematic and narrative reviews. We did not limit our results to English language journals, and did retrieve some abstracts in Spanish (but none to empirical studies), but one limitation is that our search terms were entered in English. Our next update will include a wider range of terms and translation of these terms into Spanish and French languages.

Data collection and analysis

Selection of studies

AP screened citations generated for the original review. AP and CTP independently examined these citations. Full reports were obtained for 11 potential randomized trials. Both review authors agreed that two of these should be excluded. Arbitration was not required as the two review authors agreed. For this update, two review authors (MHP and JL) scanned each citation and determined that there were no trials suitable for inclusion in this review. Details of six new 'excluded studies' with reasons for exclusion are provided in [Excluded studies](#).

Data extraction and management

AP extracted data from each of the nine main study reports using a specially designed instrument adapted from his earlier study ([Petrosino 1997](#)), and included items are listed in the 'Characteristics of included studies'. Where 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. Investigators were helpful but unable to locate additional data. In two cases we retrieved unpublished Masters' theses from university libraries to see if they contained this information (Locke 1984; Cook 1990). They did not. Another review author (CTP) double checked all extracted data on outcomes to ensure they were correct.

Assessment of risk of bias in included studies

For each study, we assessed methodological quality using the Cochrane 'Risk of bias' tool. The study reports generally lacked explicit details about randomization and concealment, and the 'Risk of bias' ratings reflect the uncertainty stemming from this lack of description. The Cochrane 'Risk of bias' tool asks review authors to rate each of the following areas of risk:

1. random sequence generation;
2. allocation concealment;
3. blinding of participants and personnel;
4. blinding of outcome assessment;
5. incomplete outcome data (attrition);
6. selective reporting;
7. other sources of bias. Here we rated whether the

implementation of the program rendered a fair test. This is a very low cost and easy to implement program, and no reports included details of program implementation problems.

Measures of treatment effect

Studies had to include at least one outcome of subsequent offending behavior, as measured by such indices as arrests, convictions, contacts with police or self-reported offences. 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 effect on these measures. Although we do not analyze them, we list other 'noncrime measures' and their effects (for example, attitudinal, educational) reported by evaluators in case subsequent review authors in the Cochrane or Campbell Collaborations require them.

Unit of analysis issues

All of the included studies involved randomization of individuals to conditions. No cluster-randomized trials were located. Most studies involved a single treatment and a single control group; in one instance in which multiple groups were involved (Vreeland 1981), we only included data from the strongest contrast (the most intensive treatment versus control).

Dealing with missing data

As mentioned earlier, we made unsuccessful attempts to acquire missing outcome data for two studies (Locke 1986; Cook 1992). Due to the lack of subsequent follow-up intervals for outcome

measurement in the included studies, we focused exclusively on first treatment effects. This likely limited missing outcome data problems as only one study experienced postrandomization attrition (Yarborough 1979). We examined the impact of excluding this study in a sensitivity analysis, discussed below.

Assessment of heterogeneity

The included studies represent some variation in geographic locations, specific types of interventions implemented, and juvenile treatment populations. Thus, heterogeneity should be examined, although the small number of included studies makes interpretation risky. The Chi² and I² statistics for heterogeneity are reported.

Assessment of reporting biases

Seven studies were included in the meta-analyses, and just two were published in academic peer-reviewed publications (Finckenaue 1982; Lewis 1983). Therefore, we do not believe publication bias is a threat to the results. In the future, if additional studies are located, we will include Egger's regression test for funnel plot asymmetry (Egger 1997).

Data synthesis

Using Review Manager software (RevMan 2011), we expressed dichotomous outcome measures of crime as odds ratios (OR). We reported the 95% confidence intervals (CI). Both fixed-effect and random-effects models were assumed across the randomized trials and compared to assess the impact of statistical heterogeneity, and both were reported. We examined OR at first follow-up interval, that is, first post-treatment effect.

Subgroup analysis and investigation of heterogeneity

No subgroup analyses were determined a priori at the protocol stage. We did not change our plans, given that only seven studies that included outcome data for analysis. Thus, we did not explore heterogeneity by conducting analyses of subgroups or moderators.

Sensitivity analysis

We conducted two sensitivity analyses that examined the impact on the results of excluding studies with significant methodological issues. The first analysis involved dropping a study that experienced randomization problems (Finckenaue 1982). The second sensitivity analysis involved dropping a study that involved substantial postrandomization attrition (Yarborough 1979).

RESULTS

Description of studies

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

Michigan Department of Corrections (1967)

In an internal, unpublished government document, the Michigan Department of Corrections reported a trial testing a program that involved taking adjudicated juvenile boys on a tour of a state reformatory (Michigan D.O.C. 1967). 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 15 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 Michigan Department of Corrections found that 43% of the experimental group reoffended, compared to only 17% of the control group. This large negative result curiously receives little attention in the original document.

The Greater Egypt Planning and Development Commission, Illinois, USA (1979)

This program at the Menard Correctional Facility started in 1978 and is described as a frank and realistic portrayal of adult prison life. The researchers randomly assigned 161 youths aged 13 to 18 years to attend the program or a no-treatment control. The participants were a mix of delinquents or children at risk of becoming delinquent. Participants were compared on their subsequent contact with police, on two personality inventories (Piers-Berne and Jesness) and used surveys of parents, teachers, inmates and young people. The outcomes are also negative in direction but not statistically significant, with 17% of the experimental participants being recontacted by police in contrast to 12% of the controls (GERP&DC 1979). 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 the program on two attitude tests (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 (p. 12). Researchers also note how positive and enthusiastic inmates were about their efforts.

Michigan JOLT Study, USA (Yarborough 1979)

In the Juvenile Offenders Learn Truth (JOLT) program, juvenile delinquents in contact with one of four Michigan county courts participated. Each juvenile spent five total hours in the facility. Half of this time was spent in a confrontational 'rap' session. This

followed a tour of the facility, during which participants were escorted to a cell and exposed to interaction with inmates (for example, taunting). 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 months' follow-up. This second Michigan study reported very little difference between the intervention and control group (Yarborough 1979). The average offense rate for program participants, however, was 0.69 compared to 0.47 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 (Orchowsky and Taylor 1981)

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 to 20 years 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 12-month intervals. The only positive findings, though not statistically significant, were reported in Virginia (Orchowsky 1981). Although the difference at six months was not statistically significant (39% of controls had new court intakes versus 41% of experimental participants), they favor the experimental participants at nine and 12 months. The investigators noted, however, that the attrition rates in their experiment were dramatic. At nine months, 42% of the original sample dropped out, and at 12 months, 55% dropped out. The investigators conducted analyses that seemed to indicate that the constituted groups were still comparable on selected factors.

Texas Face-to-Face Program, USA (Vreeland 1981)

The Face-to-Face program included a 13-hour orientation session in which the juvenile lived as an inmate followed by counseling. Participants were 15 to 17 years of age and on probation from Dallas County Juvenile Court; most averaged two or three offenses before the study. A total of 160 boys were randomly assigned to four conditions: prison orientation and counseling, orientation only, counseling only or a no-treatment control group. Vreeland examined official court records and self-reported delinquency at six months. This evaluation also reported little effect for the intervention (Vreeland 1981). Vreeland reported that the control participants outperformed the three treatment groups on official delinquency (28% delinquent for control versus 39% for prison

orientation plus counseling versus 36% for prison only versus 39% for counseling only). This more robust measure contradicts data from the self-report measures used, which suggest that all three treatment groups did better than the no-treatment controls. None of these findings reached a level of statistical significance. 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 for Face-to-Face on several attitudinal measures, including the 'Attitudes Toward Obeying Law Scale.'

New Jersey 'Scared Straight' Program, USA (Finckenauer 1982)

The New Jersey Lifers' Program began in 1975 and stressed confrontation with groups of juveniles ages 11 to 18 years who participated in a rap session. Finckenauer randomly assigned 82 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. Finckenauer reported that 41% of the children and young people who attended the 'Scared Straight' program in New Jersey committed new offenses, while only 11% of the controls did, a difference that was statistically significant (Finckenauer 1982). He also reported that the program participants committed more serious offenses and that the program had no impact on nine attitude measures with the exception of a measure called 'attitudes toward crime.' On this measure experimental participants did much worse than controls. We deal with Finckenauer's own concerns about randomization integrity in a sensitivity analysis that is reported later.

California SQUIRES Program, USA (Lewis 1983)

This is supposedly the oldest such program in the USA beginning in 1964 (Lewis 1983). The San Quentin Utilization of Inmate Resources, Experience and Studies (SQUIRES) program included male juvenile delinquents from two California counties between the ages of 14 and 18 years, 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 compared participants on seven crime outcomes at 12 months. 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 children and young people 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 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, again underscoring the difficulty of achieving behavioral change even when positively affecting the attitudes of juvenile delinquents.

Kansas Juvenile Education Program, USA (Locke et al. 1986)

Kansas Juvenile Education Program (KEP) was designed to educate children about the law and the consequences of violating it (Locke 1986). The program also tried to match juveniles with inmates based on personality types. Fifty-two juvenile delinquents aged 14 to 19 years from three Kansas counties were randomly assigned while on probation to KEP or a no-treatment control. The investigators examined official (from police and court sources) and self-report crime outcomes at six months. Locke and his colleagues reported little effect of the KEP program. Both groups improved from pretest to post-test 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 Cerkovich attitude tests.

Mississippi Project Aware, USA (Cooke and Spurrison 1992)

Project Aware was a nonconfrontational, educational program comprising one five-hour session run by prisoners (Cook 1992). The intervention was delivered to juveniles in groups of six to 30. In the study, 176 juveniles (ages 12 to 16 years) 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. Little difference was found between experimental and control participants in the study. 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 dropout. Curiously, they report an effect for the program on school dropout data, but not that "...it is not clear how the program succeeded in reducing dropout rates..." (p. 97).

Results of the search

The search methods for the original review generated 487 citations, most of which had abstracts. AP screened these citations,

determining that 30 were evaluation reports. AP and CTP independently examined these citations and agreed that 11 were potential randomized trials. All reports were obtained. Upon inspection of the full-text reports, we excluded two studies. One study was excluded because it did not include any post program measure of offending. This was 'Project Aware', which had been conducted in a Wisconsin prison (Dean 1982). Attempts to contact the study author or retrieve these data from any other reports by the Wisconsin Department of Corrections have been unsuccessful. A second study of 'Stay Straight', conducted in Hawaii, was also excluded, due to the absence of random assignment (Chesney-Lind 1981). After the two exclusions, we were left with nine randomized trials. Our updated searches yielded no new eligible studies or reports of any ongoing trials. Two review authors (MHP and JL) scanned each citation and identified five potentially relevant reports. One, an evaluation of a Scared Straight program for truants, was excluded because it did not involve randomization (Bazemore 2004). Another study was excluded because it did not include eligible outcome measures; it measured change in attitudes toward jail or prison (Feinstein 2005). Two articles discussed a related 'experiment' (Blunkett 2008; Wilson 2010), but upon further examination we discovered these studies did not use experimental methods or eligible outcomes. Another positive descriptive report was identified of a juvenile awareness program involving 'fear appeal messages' (Windell 2005), but no evaluative data were provided. A systematic review (Klenowski 2010) was identified that included narrative descriptions of 10 studies, but it contained no new studies eligible for inclusion in our review. Thus, information contained in this update is based on studies located for the previous review.

Included studies

Collectively, the nine studies were conducted in eight different states of the USA, with Michigan the site for two studies (Michigan D.O.C. 1967; Yarborough 1979). No set of researchers conducted more than one experiment. The studies span the years 1967 to 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. The average age of the juvenile participants in each study ranged from 15 to 17 years. Only the New Jersey study included girls (Finckenaer 1982). Racial composition across the nine studies was diverse, ranging from 36% to 84% white people. Most of the studies dealt with delinquent youths already in contact with the juvenile justice system. All of the experiments were simple two-group experiments except Vreeland's evaluation of the Texas Face-to-Face program (Vreeland 1981). Only one study used quasi-random alternation techniques to assign participants (Cook 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 from self-report measures. In two studies (Locke 1986; Cook 1992), no postintervention offending rates were reported. Some of the studies that in-

cluded 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, 12 and 24 months.

Excluded studies

There were six studies that were excluded during this update. These are often included in other review authors' samples. We describe these in more detail below, along with their reason for exclusion. Bazemore 2004 evaluated a 'Scared Straight' program for truants, however their study did not involve randomization. This program involved a collaborative intervention administered by a local sheriff's department. It followed 550 youth (350 'treatment' and 200 'control'). Three outcome measures were used: (1) whether or not youths returned to school the next day or were stopped by an officer (different measures for treatment and control youths), (2) comparison of the number of unexcused absences 30 days pre-intervention and postintervention and (3) total number of days of school missed following the intervention. Delinquent involvement was also measured. This study provided mixed results regarding program effectiveness.

Berry 1985 evaluated the 'Shape Up' program carried out in Colorado. The experimental group consisted of 30 males ages 14 to 18 years, and the control group consisted of 27 males of the same age. The study used a matched comparison group design and did not use randomization. The study assessed perception of certainty, severity and seriousness of punishment; delinquency proneness; intelligence quotient (IQ); family dynamics as measured by Family Adaptability and Cohesion Evaluation Scale (FACES) II, and recidivism rates. No difference was found between the two groups on attitude change, re-arrest, conviction, and weighted seriousness of crime after program involvement.

Buckner 1983 evaluated a program, 'Stay Straight', which was carried out in Hawaii. This study did not randomize participants and instead used a matched comparison group design. They assessed rearrest rates, finding that there was no effect on female participants. Male participants had higher rearrest rates than nonparticipants following the intervention.

Chesney-Lind 1981 evaluated a program, 'Stay Straight', which was carried out in Hawaii. This study was excluded due to a lack of random assignment of participants. An after-the-fact matched group design was used in this study. The frequency and severity of police arrests in the year following program exposure was used as an outcome measure.

Dean 1982 evaluated a two-session juvenile awareness program in Wisconsin. This study used a small sample of boys who were involved in a residential treatment program for delinquents. The study assessed 13 traits thought to be associated with a delinquent personality finding internal locus of control had increased significantly, while chance expectation and social self concept had decreased significantly. A pretest-post-test design with randomiza-

tion was used, but no data on delinquency outcomes were collected.

Langer 1980 evaluated the Juvenile Awareness Program of the Lifers' Group at the Rahway State Prison in New Jersey. This study used a matched comparison group design. The study assessed delinquent involvement, finding that at the 10-month follow-up there was no significant difference between treatment and control groups. At long-term (average of 22 months) follow-up, the control group had significantly higher delinquency rates than the

treatment group.

Risk of bias in included studies

Review authors AP and MHP rated quality of included studies using The Cochrane Collaboration's 'Risk of bias' tool (Higgins 2011). Unfortunately, clear data on all seven items in the 'Risk of bias' tool was not included in study reports. Figure 1 provides summary results, and we discuss each of the rating areas below.

Figure 1. Risk of bias summary: review authors' judgments about each risk of bias item for each included study
 Green circle: low risk of bias Question mark: unclear risk of bias Red circle: high risk of bias

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding (performance bias and detection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Cook 1992	⊖	?	?	⊖	?	⊖	+	+
Finckenaer 1982	⊖	?	?	⊖	+	+	+	+
GERP&DC 1979	?	?	?	⊖	?	+	+	+
Lewis 1983	+	?	?	⊖	?	+	+	+
Locke 1986	+	?	?	⊖	?	⊖	+	+
Michigan D.O.C. 1967	?	?	?	⊖	+	+	+	+
Orchowsky 1981	+	?	?	⊖	?	+	+	+
Vreeland 1981	+	?	?	⊖	?	+	+	+
Yarborough 1979	+	?	?	⊖	?	⊖	+	+

Allocation

Random sequence generation

Finckenaue reported violations of randomization (Finckenaue 1982). Only eight of the 11 participating agencies that referred troubled or delinquent boys to the program correctly assigned their cases. Finckenaue 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 (Analysis 1.3). Another study was rated as at high risk of bias because alternation was used (Cook 1992). This latter study was not included in the meta-analysis because it did not include data on postintervention offending. Two other studies did not provide any further information on randomization and their risk of bias was rated as 'unclear' (GERP& DC 1979; Michigan D.O.C. 1967).

Allocation concealment

All of the studies are rated as presenting 'unclear' risk as there is no information on how randomization was performed.

Blinding

Blinding of participants and personnel (performance bias)

Blinding was not possible in these studies, and all are rated as presenting 'high risk'.

Blinding of outcome assessment (detection bias)

We should note that only one study author reported that steps were taken to 'blind' those responsible for collecting the outcome data to treatment assignment (Michigan D.O.C. 1967) and is rated as presenting 'low risk'. All others are rated as presenting 'unclear risk'.

Incomplete outcome data

Six studies experienced little or no attrition and are rated as presenting 'low risk' of bias. Two studies appeared to report significant attrition (defined as 10% or more from the originally randomized sample). The Virginia Insiders study reported a major loss of participants from the initial randomization sample (Orchowsky 1981). They reported this, however, at the second and third follow-up intervals (not the first, at six months). Because there was a paucity of data beyond the immediate follow-up interval across studies, we only conducted a pooled analysis using data at that

time interval. Therefore a sensitivity analysis of the impact of this later attrition was not performed. The Cook study is also rated as presenting a 'high risk' due to attrition, but the study did not include data for the first follow-up and was not included in any meta-analyses (Cook 1992).

The Michigan JOLT study reported a large number of no-shows but they were deleted from the analysis (Yarborough 1979). The problem is that we do not know how many participants were initially assigned and no data were reported that the remaining sample was similar to the initial sample. We also conducted a sensitivity analysis to determine the influence of this study on the pooled analysis.

Selective reporting

We rated this as presenting a 'low risk' of bias across the studies. In several cases, the program was a government intervention and the researchers were employed by the same agency; nonetheless, the negative or null findings were clearly presented (Michigan D.O.C. 1967; GERP&DC 1979; Yarborough 1979; Orchowsky 1981; Lewis 1983). In three instances, the authors were students and a number of outcomes were presented (Vreeland 1981; Locke 1986; Cook 1992). In another instance, the author was an academic researcher who presented a number of findings in an academic book (Finckenaue 1982).

Other potential sources of bias

In terms of 'other bias' as rated on the tool, a major threat to study results is if the program is so poorly implemented that it does not represent a true test of the treatment. Scared Straight programs appear to be relatively simple and short-term and pose few problems for implementation. No investigator reported implementation problems, and we rated these as 'low risk' of bias. We should note that not one of the nine included studies provided data on monitoring of the control group to determine if compensation was an issue. It is probably very unlikely that control group participants received anything like Scared Straight but it was not specifically addressed by authors of the reports.

Effects of interventions

Findings from the individual studies

Whether relying on the actual data reported or measures of statistical significance, the nine trials do not yield evidence for the effectiveness of Scared Straight and other juvenile awareness programs on subsequent delinquency. In the first such study, the Michigan Department of Corrections found that 43% of the experimental

group reoffended, compared to only 17% of the control group (Michigan D.O.C. 1967). No test of statistical significance was reported by the trialists. We performed a Chi² test, which indicated no statistical significance for this outcome, likely due to the low statistical power of the sample. The original document does not comment on this large percentage difference.

In Illinois, the outcomes were also negative in direction but not statistically significant, with 17% of the experimental participants being recontacted by police in contrast to 12% of the controls (GERP&DC 1979). 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 reported no effect for the program on two attitude tests (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 (p. 12). Researchers also note how positive and enthusiastic 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 0.69 compared to 0.47 for the control group. As Yarborough (p. 14) pointed out, “...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 not statistically significant, were reported in Virginia (Orchowsky 1981). Although the difference at six months was not statistically significant (39% of controls had new court intakes versus 41% of experimental participants), they favor the experimental participants at nine and 12 months. The investigators noted, however, that the attrition rates in their experiment were dramatic. At nine months, 42% of the original sample dropped out, and at 12 months, 55% dropped out. The investigators conducted analyses that seemed to indicate that the constituted groups were still comparable on selected factors such as race and age.

A study of the Face-to-Face program in Texas also reported little effect for these interventions (Vreeland 1981). Vreeland 1981 reported that the control participants outperformed the three treatment groups on official delinquency (28% delinquent for control versus 39% for prison orientation plus counseling versus 36% for prison only versus 39% for counseling only). This more robust measure contradicts data from the self-report measures used, which suggest that all three treatment groups did better than the no-treatment controls. None of these findings reached a level of statistical significance. Viewing all the data, Vreeland 1981 concluded that there was no evidence that Face-to-Face was an effective delinquency prevention program. He finds no effect for Face-to-Face on several attitudinal measures, including the Attitudes Toward Obeying Law Scale.

Finckenaue 1982 reported that 41% of the children and young

people who attended the Scared Straight program in New Jersey committed new offenses, while only 11% of controls did, a difference that was statistically significant. He also reported that the program participants committed more serious offenses and that the program had no impact on nine attitude measures with the exception of a measure called ‘attitudes toward crime.’ On this measure experimental participants did much worse than control participants. We deal with Finckenaue’s own concerns about randomization integrity in this study in a sensitivity analysis.

Additional evidence of a possible harmful effect can be found in the evaluation of the California SQUIRES program (Lewis 1983). Lewis 1983 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 children and young people 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 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 control participants. Data were reported on eight attitudinal measures, and Lewis reported that the program favored the experimental group on all of them, again underscoring the difficulty of achieving behavioral change even when positively affecting the attitudes of juvenile delinquents.

Locke and his colleagues reported little effect of the Juvenile Education Program in the Kansas State Prison (Locke 1986). Both groups improved from pretest to post-test 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 Cerkovich attitude tests.

Finally, little difference was found between experimental and control participants in the Mississippi Project Aware study (Cook 1992). For example, the mean offending rate for control participants at 12 months was 1.25 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 dropout. 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).

Meta-analysis

For each study, we extracted all of the relevant crime outcome data. Our protocol included an organization of analyses by examining official reports (from government administrative records) distinct from self-reported criminality (obtained from investigator-admin-

istered survey questionnaires). Given that we expected a diverse number of measures of crime to be reported, the protocol called for us to organize it into four indexes that would be most relevant to policy and practice. These included prevalence rates (what percentage of each group reoffended or did not?), average incidence rates (what was the average number of offenses or other incidents per individual in each group?), offense severity rates (what was the average severity of offenses per individual in each group?) and latency (how long was the average return to crime or failure delayed per individual in each group?). As Table 1 shows, however, few measures except for prevalence were reported.

Given the limitation of the data, we conducted one meta-analysis. We report the crime outcomes for official measures at the first-effect or first (and usually the only) follow-up interval period reported. Each analysis focused on proportion data (that is, the proportion of each group reoffending), as the outcomes reporting means or averages were sparse and often did not include the standard deviations. Thus, because the data relied on dichotomous outcomes, both analyses report ORs and 95% CIs for each study. As a sensitivity analysis, we assume both random-effects and fixed-effect models for treatment effects across the studies.

Immediate post-treatment effects for reoffending rates: official measures

The analysis of the data in comparison Table 1 from the seven studies reporting reoffending rates shows that intervention increases the crime or delinquency outcomes at the first follow-up period. Assuming either a fixed-effect or random-effects model does not change its overall negative impact. Using a fixed-effect model, the OR was 1.68 (95% CI 1.20 to 2.36). Heterogeneity statistics should be interpreted with caution given that only seven studies were included in the meta-analysis ($\text{Chi}^2 = 8.49$, P value = 0.20, $I^2 = 29\%$) (Analysis 1.1). The mean OR assuming a random-effects model was similar at 1.72 (95% CI 1.13 to 2.62); heterogeneity statistics were nearly identical ($\text{Chi}^2 = 8.50$, P value = 0.20, $I^2 = 29\%$) (Analysis 1.2). Both fixed-effect OR and random-effects OR are statistically significant; the intervention increases the odds of offending by between 1.6 to 1 and 1.7 to 1.

Sensitivity analysis 1. Excluding Finckenauer study

We excluded the Finckenauer study from the analysis because of its randomization problems. Finckenauer reported that only eight of the 11 referring agencies correctly followed the randomization procedures. His reanalyses taking these randomization problems into account still indicated a negative impact. Nonetheless, we determined to examine the impact of this study on the meta-analytic findings. Given the little difference in OR whether assuming a fixed-effect or random-effects model, we conducted a meta-analysis assuming a random-effects model. Given that the Finckenauer study reported the largest negative effects for the program, it is

not surprising that the OR decreased. However, it is still negative in direction at 1.47, and statistically significant (95% CI 1.03 to 2.11). Heterogeneity statistics should be interpreted with caution given the small number of studies ($\text{Tau}^2 = 0.00$; $\text{Chi}^2 = 4.25$, degrees of freedom (df) = 5, P value = 0.51; $I^2 = 0\%$) (Analysis 1.3).

Sensitivity analysis 2. Excluding Yarborough study

We excluded the Yarborough study because of its deletion of no-shows postrandomization from analysis of the results, indicating a potential for high attrition bias. Yarborough did not report any analyses to indicate how this affected the remaining sample. We again assumed a random-effects model. The deletion of this study did not alter the overall negative impact of these programs, as the OR was 1.96. This is statistically significant (95% CI 1.25 to 3.08). Heterogeneity statistics should be interpreted with caution given the small number of studies ($\text{Tau}^2 = 0.06$; $\text{Chi}^2 = 6.25$, $df = 5$, P value = 0.28; $I^2 = 20\%$) (Analysis 1.4).

Although the methodological limitations of the studies warrant our sensitivity analyses, their exclusion did not alter the main conclusion of the meta-analyses: a significant negative impact of the program.

Sensitivity analysis 3. Excluding both Finckenauer and Yarborough studies

We excluded both the Finckenauer and Yarborough studies to see how this affected the overall meta-analysis. As Analysis 1.5 shows, even with two studies removed for sensitivity analysis, the overall effect of the intervention in the five remaining studies shows a 'criminogenic' effect that is statistically significant, that is, favors the control group not Scared Straight (OR 1.69, 95% CI 1.10 to 2.58). Heterogeneity statistics should be interpreted with caution given only five studies are in the analysis ($\text{Tau}^2 = 0.00$; $\text{Chi}^2 = 2.94$, $df = 4$, P value = 0.57, $I^2 = 0\%$).

DISCUSSION

Summary of main results

These randomized trials, conducted over a 25-year period in eight different US states, provide evidence that Scared Straight and other 'juvenile awareness' programs are not effective as a stand-alone crime prevention strategy. More importantly, they provide empirical evidence - under experimental conditions - that these programs likely increase the odds that children exposed to them will commit offenses in future. Despite the variability in the type of intervention used, ranging from harsh, confrontational interactions to tours of the facility, they converge on the same result: an increase in criminality in the experimental group when compared to a no-

treatment control. Doing nothing would have been better than exposing juveniles to the program.

We noted that the other two trials that did not report prevalence data for the meta-analysis also reported no effect for the intervention (Locke 1986; Cook 1992). Indeed, the mean data from the Mississippi study was also negative in direction, and the Kansas investigators reported that the self-reported data showed a negative impact.

Overall completeness and applicability of evidence

Given that the seven trials used in the meta-analysis were conducted in six states using different conceptions of the intervention underscore the high external validity of these findings. However, note that all trials were of US programs, and no trial was reported after 1992. Indeed, no trial included in the meta-analysis was reported since 1983.

Quality of the evidence

Nine randomized trials were included in the review; only randomized trials, if implemented with good fidelity, produce statistically unbiased effects. However, the nine studies were not exemplars of trial quality. These were small studies, with very few providing convincing evidence that they reduced bias threats as measured by the Cochrane 'Risk of bias' tool (Figure 1). In fact, for some of the bias threats, the trials were rated with a great deal of uncertainty due to the lack of descriptive data in the report. However, three sensitivity analyses were conducted, the first dropping the study that experienced the greatest threat of bias due to randomization compromise (Finckenaer 1982), the second study that lost a considerable number of participants postrandomization (Yarborough 1979), and the third dropping them both. The effect sizes remained stable in all three analyses, indicating that the negative effect for Scared Straight and other juvenile awareness findings is robust.

Potential biases in the review process

Although we believe we have identified all relevant RCTs, it is possible that studies in languages other than English and not indexed in English language databases could have been missed. In addition, it is possible that the sensitivity of our search could have been increased; for example, by using additional indexing terms specific to the databases we searched and using truncation to ensure we searched for word variations. A revised search strategy will be developed for the next update.

Agreements and disagreements with other studies or reviews

The results of this review converge with the findings from many other narrative or quantitative reviews. This is expected as the reviews generally consider the same studies. For example, reviewers of research on the effects of crime prevention programs have not found deterrence-oriented programs, such as Scared Straight, effective (Lipsey 1992; Lundman 1993; Sherman 1997). In fact, the University of Maryland's well-publicised review of over 500 crime prevention evaluations listed Scared Straight as one program that 'doesn't work' (Sherman 1997). These findings also mirror a meta-analysis of juvenile prevention and treatment programs by Lipsey 1992, who indicated that the effect size for 11 "shock incarceration and 'Scared Straight' programs" was -0.14 (or produced about 7% higher recidivism rates in experimental participants than control participants assuming a 50% baseline).

The one disagreement, in terms of syntheses of evidence, is with the US Department of Justice's CrimeSolutions.Gov registry of effects on crime policies and programs (US Department of Justice 2012). The Crime Solutions project has rated the evidence as inconclusive. There are two reasons for the discrepancy. The first is that the Crime Solutions rating scheme relies on statistical significance to determine whether there is evidence of effect; indeed, some of the program evaluations included here were underpowered due to small sample size and did not report a statistically significant finding. Second, the Crime Solutions project is defining Scared Straight narrowly, as its initial iteration in New Jersey defined it, in contrast with the broader definition of Scared Straight and similar "kids visit prison" programs used here. Thus, while Crime Solutions is only considering a small set of studies that examined a narrowly defined intervention (known as Scared Straight), this review includes nine program evaluations that would fall under a broader heading of juvenile awareness programs.

AUTHORS' CONCLUSIONS

Implications for practice

The strong indication here is that these programs have a harmful effect. This raises a dilemma for policymakers. Criminological interventions, when they cause harm, are not just toxic to the participants. They cause more harm to citizens who were not part of the experiment because of the increase in criminal victimization. Policymakers 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. We believe that our updated review places the onus on every jurisdiction to show how their current or proposed program is different than the ones studied 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 inmates involved in the prison visits and that argument is sometimes used to legitimize use of the program. These arguments are undoubtedly used under the assumption that the program does no harm. In light of the findings of this review, assertions that Scared Straight and similar programs ought to be used because they have other positive effects raises ethical questions about potentially harming children (and others in the community who may be victimized) in order to accomplish other important, but latent, goals.

The authors have received communications from different prison facilities that are using a juvenile awareness program. One argument used to sustain such programs is that the research reported here does not apply to their particular program. Our recommendation is that correctional research units, either at the facility or at a regional or national government level, collaborate with program staff to conduct a rigorous evaluation. If such units do not exist or cannot conduct their own study, we suggest they collaborate with a local university, college or research firm that could undertake this work to ensure that the program is working as planned and not unintentionally causing more harm than good.

Correctional administrators sometimes ask whether our results are relevant to their particular program. For example, inmates running the program may go outside the prison to speak at schools about their life experiences. Our review only looked at programs involving visits of young people to prisons, and, as far as we know, no review has examined juvenile awareness interventions that involve offenders leaving prison grounds to speak to children at school. We are not aware of any controlled studies testing it.

We receive periodic correspondence from concerned citizens about how to get a juvenile who is in trouble with the law into a Scared Straight program. We cannot, in good conscience, recommend this program. Our response to these well-meaning citizens is to refer them to national, regional or local centers that specialize in youth crime prevention services.

Implications for research

One question that continues to arise about these findings is why Scared Straight and similar programs seem to lead to more crime rather than less in participants. What is the critical mechanism? Although there were many good post-hoc theories about this, none of the evaluations were structured to provide the kind of mediating variables necessary to respond to this in the context of a systematic review (Petrosino 2000). One explanation may be 'peer contagion' (Dishion 1999). According to this theory, any positive impact by an intervention for youth might be offset by processes of peer influence that occur when deviant youths are allowed to interact with each other in groups, such as what occurs in Scared Straight and similar programs. This would need to be explicitly tested in careful evaluation studies to confirm as a potential mechanism for

harmful effects.

We plan to update this review again within 36 months to incorporate any new studies or respond to cogent criticisms. Given that we found only nine studies (and only seven were used in the meta-analysis), we were cautious not to propose the use of moderating variables in subsequent analyses. Initially we wondered if one program factor might have particular salience, which was the degree of harshness in the inmate presentations. It may be that the more brutal and vulgar the presentation, the more that it causes a type of 'backfire' effect, producing in the juveniles the very behavior it seeks to deter. However, when looking at this more closely, we discovered that one trial involving a tour of a reformatory *with no presentation* reported one of the largest negative effects (Michigan D.O.C. 1967).

This review has led us to consider two others, contingent on future funding. 'Shock value'-type interventions are tried across many fields. For example, high school students are sometimes shown horrific footage of car accidents in order to deter them from drinking and driving. In industrial arts classes, students are shown films of what occurs when safety glasses are not worn; this is often graphic and is designed to increase compliance with such regulations. There are many other examples across fields. But is there any evidence that any of these 'shock value' interventions work? Or do they produce disappointing, or even toxic, results as we have reported here? The early evidence is not promising, as fear appeals in reducing drug and alcohol among young people have been described in at least one review as 'disappointing' (Prevention First 2008).

It may be true that Scared Straight and similar programs do not work because they only convey a threat that juveniles do not think will be carried out. What about the evidence for deterrence if it is not a third-party threat but actual involvement in the juvenile justice system? There has been a wide range of randomized trials that test for the effects of official processing in juvenile courts with some other intervention (such as diverting the child from such processing). Is there evidence that the delivery of a threat - official system processing - deters future criminal behavior? Petrosino 2010 examined 29 randomized trials that evaluated the effects of some diversionary alternative (services or outright release) and compared it to official processing or progression deeper into the juvenile justice system. That review, published by the Campbell Collaboration, also indicated that formal system processing or progression had no crime deterrent effect, and, in some instances, increased crime in contrast to diversionary alternatives.

ACKNOWLEDGEMENTS

The original review in 2002 was principally supported by USD5000 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 UK 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.

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 appreciated the guidance and editorial comments of Dr. Jane

Dennis, Professor Geraldine Macdonald (Co-ordinating Editor), Dr. Julian Higgins, Dr. Stuart Logan and other members of the Cochrane Developmental, Psychosocial and Learning Disorders Group. Criticisms of the original review by Professor Robert Boruch, Sir Iain Chalmers, Dr. Phoebe Cottingham, Professor Lyn Feder, Professor Hiroshi Tsutomi and Professor Joan McCord also helped.

We acknowledge the contributions of John Buehler to the original review. John passed away in 2003.

The original review and the review update were produced within the Cochrane Developmental, Psychosocial and Learning Problems Group and the Campbell Crime and Justice Group. The 2012 update was financially supported by the Campbell Collaboration, Oslo, Norway. We appreciate the assistance of Eammon Noonan, Chief Executive Officer, in facilitating this funding.

REFERENCES

References to studies included in this review

Cook 1992 *{published data only}*

Cook DD. *Effects of a Non-Confrontational Prisoner-Run Juvenile Delinquency Deterrence Program [Masters Thesis]*. Mississippi, US: Mississippi State University, 1990.

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

Finckenauer 1982 *{published data only}*

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

GERP&DC 1979 *{published data only}*

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

Lewis 1983 *{published data only}*

Lewis RV. Scared Straight - California style: evaluation of the San Quentin SQUIRES program. *Criminal Justice & Behavior* 1983;**10**(2):209–26.

Locke 1986 *{published data only}*

Locke TP. *An Analysis of the Kansas State Penitentiary Juvenile Education Program [Masters thesis]*. Lawrence, KS: University of Kansas, 1982.

* 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**:281–98.

Michigan D.O.C. 1967 *{published data only}*

Michigan Department of Corrections. A Six Month Follow-up of Juvenile Delinquents Visiting the Ionia Reformatory. Research Report No. 4. Lansing: Michigan Department of Corrections 1967.

Orchowsky 1981 *{published data only}*

Orchowsky S, Taylor K. The Insiders Juvenile Crime Prevention Program. Richmond: Virginia Department of Corrections 1981.

Vreeland 1981 *{published data only}*

Vreeland AD. *Evaluation of Face-to-Face: A Juvenile Aversion Program [thesis]*. Dallas, TX: University of Texas, 1981.

Yarborough 1979 *{published data only}*

Yarborough JC. Evaluation of JOLT as a Deterrence Program. Michigan Department of Corrections 1979.

References to studies excluded from this review

Ashcraft 1970 *{unpublished data only}*

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

Bazemore 2004 *{published data only}*

Bazemore G, Stinchcomb JB, Leip LA. Scared smart or bored straight? Testing deterrence logic in an evaluation of police-led truancy intervention. *Justice Quarterly* 2004;**21**(2):269–99.

Berry 1985 *{unpublished data only}*

Berry RL. *Shape Up: The Effects of a Prison Aversion Program on Recidivism and Family Dynamics [PhD thesis]*. Greeley, CO: University of Northern Colorado, 1985.

Blunkett 2008 *{published data only}*

Blunkett D. Scared straight?. *Public Policy Research* 2008;**15**(2):77–9.

Brodsky 1970 *{published data only}*

Brodsky SL. The prisoner as an agent of attitude change: a study of prison profiles' effects. *British Journal of Criminology* 1970;**10**:280–5.

- Buckner 1983** *{published data only}*
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):227–47.
- Chesney-Lind 1981** *{unpublished data only}*
Lind MC. Ike Na Pa'ahao: the Experience of the Prisoners (formerly Stay Straight) - Evaluation. A Juvenile Awareness Program. Report No. 258 University of Hawaii at Manoa, Youth Development & Research Center 1981.
- Dean 1982** *{published data only}*
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.
- Feinstein 2005** *{published data only}*
Feinstein S. Another look at scared straight. *The Journal of Correctional Education* 2005;**56**(1):40–4.
- Gilman 1977** *{unpublished data only}*
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 1980** *{unpublished data only}*
Langer S. *Fear in the Deterrence of Delinquency: a Critical Analysis of the Rahway State Prison Lifers' Program* [PhD thesis]. Newark, NJ: Rutgers University, 1980.
- Lloyd 1995** *{published data only}*
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 1995.
- Mitchell 1986** *{published data only}*
Mitchell JJ, Williams SA. SOS: reducing juvenile recidivism. *Corrections Today* 1986;**48**(3):70–1.
- Muhammed 1999** *{published data only}*
Muhammed L. Kids and crooks revisited: some were 'Scared Straight!'. USA Today 1999:4D.
- Nelson 1991** *{unpublished data only}*
Nelson Z. The Day in Prison Program. The First Twelve Months. Community based Corrections Division, Victoria Office of Corrections, Australia 1991.
- NSW BoS 1980** *{unpublished data only}*
New South Wales Bureau of Crime Statistics and Research. Day in Gaol Programme: Research Report No. 8. New South Wales Bureau of Crime Statistics and Research 1980.
- Nygaard 1980** *{published data only}*
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 1993** *{published data only}*
O'Malley B, Coventry G, Walters R. Victoria's Day in Prison Program: an evaluation and critique. *Australian and New Zealand Journal of Criminology* 1993;**26**(2):171–83.
- Portnoy 1986** *{published data only}*
Portnoy RN. *Ego Defenses as a Predictor of Deterrence* [PhD thesis]. Lincoln, NE: University of Nebraska, 1986.
- Rasmussen 1996** *{unpublished data only}*
Rasmussen DW, Yu Y. An Evaluation of Juvenile Justice Innovations in Duval County, Florida. Tallahassee, FL: Florida State University 1996.
- Shapiro 1978** *{unpublished data only}*
Shapiro A. Scared Straight [film]. Santa Monica, CA: Pyramid Films 1978.
- Storvoll 1998** *{published data only}*
Storvall AE, Hovland A. The Ullersmo Project: Scared Straight in Norway 1992-1996 [Ullersmoprojektet: "Scared Straight" I Norge 1992–1996]. *Nordisk Trdskrift for Kriminalvidenskab* 1998;**85**(2):122–35.
- Trotti 1980** *{unpublished data only}*
Trotti TC. Save the Children Program - An Analysis of Client Reaction. South Carolina Department of Youth Services, Research and Evaluation Unit 1980.
- Wilson 2010** *{published data only}*
Wilson D, Groombridge N. I'm making a TV programme here: reality TV's *Banged Up* and public criminology. *Howard Journal of Criminology* 2010;**49**(1):1–17.
- Windell 2005** *{published data only}*
Windell JO, Allen Jr JS. An application of fear appeal messages to enhance the benefits of a jail encounter program for youthful offenders. *Youth Violence and Juvenile Justice* 2005;**3**(4):388–94.

Additional references

- Blum 2001**
Blum J, Woodlee Y. Trying to give kids a good scare. The Washington Post 2001 Jun 3:C01.
- Cook 1990**
Cook DD. *Effects of A Non-Confrontational Prisoner-Run Juvenile Delinquency Deterrence Program* [Master's thesis]. Mississippi, US: Mississippi State University, 1990.
- Denhart 2011**
Denhart A. Beyond Scared Straight's real-life controversy, 2011. www.thedailybeast.com/articles/2011/02/23/beyond-scared-strights-real-life-controversy.html (accessed 9 February 2013).
- Dishion 1999**
Dishion TJ, McCord J, Poulin F. When interventions harm. Peer groups and problem behavior. *American Psychologist* 1999;**54**:755–64.
- Egger 1997**
Egger M, Davey-Smith, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**7**:629–34.
- Finckenauer 1999**
Finckenauer JO, Gavin PW, Hovland A, Storvoll E. [SCARED STRAIGHT! : THE PANACEA PHENOMENON REVISITED / JAMES O.

- FINCKENAUER, PATRICIA W. GAVIN ; WITH ARILD HOVLAND, ELISABET STORVOLL. Prospect Heights, Ill. : Waveland Press, c1999.]. *Scared Straight: The Panacea Phenomenon Revisited*. Prospect Heights, IL: Waveland Press, 1999.
- Hall 1999**
Hall A. Jailhouse shock aims to scare youths straight. *The Scotsman* 1999 26 Oct:12.
- Higgins 2011**
Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Klenowski 2010**
Klenowski PM, Bell KJ, Dodson KD. An empirical evaluation of juvenile awareness programs in the United States: can juveniles be 'Scared Straight?'. *Journal of Offender Rehabilitation* 2010;**49**:254–72.
- Kvaraceus 1964**
Kvaraceus W. *Juvenile Delinquency: A Problem for the Modern World*. New York, NY, USA: UNESCO, 1964.
- Lipsey 1992**
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, et al. editor(s). *Meta-Analysis for Explanation*. New York: Russell Sage, 1992:83–127.
- Locke 1984**
Locke TP. *An Analysis of the Kansas State Penitentiary Juvenile Education Program [Master's thesis]*. Lawrence, KS: University of Kansas, 1984.
- Lundman 1993**
Lundman RJ. *Prevention and Control of Juvenile Delinquency*. 2nd Edition. New York: Oxford University Press, 1993.
- McCord 2001**
McCord J, Widom C, Crowell N. *Juvenile Crime, Juvenile Justice*. Washington, DC: National Academies Press, 2001.
- Petrosino 1997**
Petrosino AJ. *'What Works?' Revisited Again: A Meta-Analysis of Randomized Experiments in Rehabilitation, Deterrence and Delinquency Prevention [thesis]*. Ann Arbor, MI: Rutgers University, 1997.
- Petrosino 2000**
Petrosino A, Turpin-Petrosino C, Finckenauer JO. Well-meaning programs can have harmful effects! Lessons from experiments in Scared Straight and other like programs. *Crime & Delinquency* 2000;**46**:354–79.
- Petrosino 2010**
Petrosino A, Turpin-Petrosino C, Guckenburg S. Formal system processing of juveniles: effects on delinquency. *Campbell Systematic Reviews* 2010; Vol. 1. [DOI: 10.4073/csr.2010.1]
- Prevention First 2008**
Prevention First. Ineffectiveness of Fear Appeals in Youth Alcohol, Tobacco and Other Drug (ATOD) Prevention. Springfield, IL 2008.
- RevMan 2011 [Computer program]**
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.
- Robinson 2011**
Robinson LO, Slowikowski J. Scary - and ineffective. Traumatizing at-risk kids is not the way to lead them away from crime and drugs, 2011. articles.baltimoresun.com/2011-01-31/news/bs-ed-scared-straight-20110131 '1' straight-type-programs-straight-program-youths (accessed 9 February 2013).
- Scripps 1999**
Scripps J. Prison tour serves as a wake-up call. *The Forum* 1999 Oct 27:1.
- Sherman 1997**
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, 1997.
- Swanson 2003**
Swanson, United Press International. Commentary: Scared Straight, 2003. www.upi.com/Business/News/Security-Industry/2003/08/21/Commentary-Scared-Straight/UPI-54131061493455/3w3635 (accessed 9 February 2013).
- Trusty 1995**
Trusty G. Sheriff's office, juvenile court resources used for restarts. *Louisiana Youth Care Magazine* 1995; Vol. 18, issue 5:3–5.
- U-CAN 2001**
United Community Action Network. Services: Wisetalk (Scared Straight). www.ucan.av.org/services.htm (no longer available) (last accessed 26 April 2001).
- UPN 1999**
United Paramount Network (UPN). Scared Straight! 20 years later, 2009. www.youtube.com/watch?v=eKU-b9YhWn4 (accessed 9 February 2013).
- US Census 2012**
US Census Bureau. Law enforcement: courts, prisons, and arrests, 2012. www.census.gov/compendia/statab/cats/law-enforcement/courts-prisons/arrests.html (accessed 9 February 2013).
- US Department of Justice 2012**
US Department of Justice. FAQ: Why doesn't CrimeSolutions.gov make more use of findings from meta-analysis?. www.crimelutions.gov. 2012 (accessed 9 February 2013).
- US HCEL 1979**
US Congress, House Committee on Education and Labor. Oversight on Scared Straight. Hearings before the House Subcommittee on Human Resources, 96th Congress,

1st Session, 4 June 1979; Vol. General Printing Office:
Washington, DC.

References to other published versions of this review

Petrosino 2002

Petrosino A, Turpin-Petrosino C, Buehler J. "Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews* 2002, Issue 2. [DOI: 10.1002/14651858.CD002796]

Petrosino 2004

Petrosino A, Turpin-Petrosino C, Buehler J. "Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *Campbell Systematic Reviews* 2004; Vol. 2. [DOI: 10.4073/csr.2004.2]

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Cook 1992

Methods	Quasi-random assignment - researchers numbered court files and assigned all odd numbered ones to intervention group
Participants	176 juvenile delinquents ages 12-16 years under jurisdiction of 1 Mississippi county youth court, 36% white, 100% male
Interventions	Educational, prisoner-run 5-hour session, designed to be nonconfrontational
Outcomes	12 and 24 months' follow-up of official court record data, average offending rates and severity of offense School attendance and school dropout
Notes	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. No standard deviations reported with any mean data, no group percentages, attempts to retrieve these data from author and other primary documents failed. All available data seem to indicate a slightly negative impact for the program on crime measures

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	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
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done

Cook 1992 (Continued)

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Data retrieved from court system. No other information provided
Incomplete outcome data (attrition bias) All outcomes	High risk	24% lost in follow-up, no analysis to ensure groups still equivalent
Selective reporting (reporting bias)	Low risk	Masters' thesis with many findings in it, including the negative result for the intervention. The type of data reported could not be included, however, in the analyses
Other bias	Low risk	No problems reported with implementation

Finckenaer 1982

Methods	Random assignment	
Participants	81 delinquent or children ages 11-18 years at risk for delinquency, 50% had prior record of offending, 40% were white, 80% male	
Interventions	1 visit, a confrontational rap session lasting approximately 3 hours with inmates serving life sentence	
Outcomes	6-month follow-up of official complaints, arrests or adjudications. Severity of offense Attitudes: <ul style="list-style-type: none"> ● toward criminals ● toward crime ● toward law ● toward justice ● toward police ● toward prison ● toward punishment ● self-image 	
Notes	Randomization breakdown is cause for concern. Principal investigator does report additional analyses for agencies that followed protocol: 31% of the experimental group recidivated compared to 17% of the control group	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Randomization broke down, 6 of the 11 referral agencies violated assignment protocol, test for equivalence showed 59% of the experimental group had a prior record, only 40% of the control group

Finckenauer 1982 (Continued)

Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Researchers collected the data from court files, not program staff
Incomplete outcome data (attrition bias) All outcomes	Low risk	None reported
Selective reporting (reporting bias)	Low risk	It seems unlikely given the full length book treatment and the amount of findings reported
Other bias	Low risk	No problems with implementation reported

GERP&DC 1979

Methods	Random assignment
Participants	161 delinquent or children at risk for delinquency, 100% male, 84% white, ages 13-18 years
Interventions	Confrontational rap session with inmates
Outcomes	5-15 months' follow-up of contacts with police Piers Harris Children's Self-Concept Scale Jesness Inventory
Notes	Nothing in the report seems to indicate that the findings should be questioned

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Random assignment, no further information
Allocation concealment (selection bias)	Unclear risk	No other description of randomization

GERP&DC 1979 (Continued)

Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Study relied on subsequent police reports, but no information provided on blinding of outcome assessors
Incomplete outcome data (attrition bias) All outcomes	Low risk	No attrition reported
Selective reporting (reporting bias)	Low risk	Unknown
Other bias	Low risk	No implementation problems reported

Lewis 1983

Methods	Random assignment	
Participants	108 juvenile delinquents from 2 California counties, most with extensive prior record, ages 14-18 years, 100% male, mostly non-white	
Interventions	Total 3 visits (1 per week) including confrontational rap sessions, guided tours of prison and interaction with prisoners, review of pictures of prison violence	
Outcomes	<p>12-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</p> <p>Attitudes:</p> <ul style="list-style-type: none"> ● toward police ● toward school ● toward crime ● toward prison ● toward work camp <p>Semantic Differential Test</p>	
Notes	Over 100 moderating analyses performed on the data There is nothing in the study report to support any lack of confidence in the observed findings	
<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement

Lewis 1983 (Continued)

Random sequence generation (selection bias)	Low risk	Test for equivalence is satisfactory but age slightly favors the experimental group
Allocation concealment (selection bias)	Unclear risk	Not stated
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not used
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	2 researchers collected court data. Unknown if they were blind to youth conditions
Incomplete outcome data (attrition bias) All outcomes	Low risk	40% of an already small sample lost in follow-up, leaving 32 in the study
Selective reporting (reporting bias)	Low risk	This was based on the authors' Masters thesis. Many results reported, including null findings for intervention. But the outcome data were not possible to be used in the subsequent analysis
Other bias	Low risk	No implementation problems reported

Locke 1986

Methods	Random assignment
Participants	53 juvenile delinquents ages 14-19 years on probation from 3 Kansas counties, 65% white, 100% male
Interventions	Non-confrontational, educational interaction, tried to match juvenile with inmate
Outcomes	Minimum 6-month follow-up of self-reported crime and juvenile court and police records of official offending
Notes	No standard deviations reported with any mean data, no group percentages, attempts to retrieve these data from author and other primary documents failed The study appears to have severe attrition, limiting our confidence. The principal investigator reported no effect for treatment but do not provide enough data for computation of odds ratios or weighted mean differences
<i>Risk of bias</i>	

Locke 1986 (Continued)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomization used, test for equivalence satisfactory (though not stated if done after attrition)
Allocation concealment (selection bias)	Unclear risk	Not stated
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not used
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	2 researchers collected court data. Unknown if they were blind to youth conditions
Incomplete outcome data (attrition bias) All outcomes	High risk	40% of an already small sample lost in follow-up, leaving 32 in the study
Selective reporting (reporting bias)	Low risk	This was based on the authors' Masters thesis. Many results reported, including null findings for intervention. But the outcome data were not possible to be used in the subsequent analysis
Other bias	Low risk	No implementation problems reported

Michigan D.O.C. 1967

Methods	Assignment using random numbers table, data collectors were blind to assignment
Participants	60 juvenile delinquents from 1 Michigan county
Interventions	2 tours of a Michigan reformatory
Outcomes	6-month follow-up of official petition for delinquency or probation violation
Notes	Brief internal report that does not fully describe nature of intervention Juvenile home records used in follow-up; data investigators were blind to group allocation The 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
<i>Risk of bias</i>	

Michigan D.O.C. 1967 (Continued)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Random numbers tables used to allocate, no test for equivalence reported
Allocation concealment (selection bias)	Unclear risk	No description of concealment of allocation
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not possible to blind participants or personnel
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Juvenile home records used in follow-up; data investigators were blind to group allocation
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only two 2 participants lost
Selective reporting (reporting bias)	Low risk	Given the report was done by the Michigan Department of Corrections, and this was their program, it is highly unlikely they would choose to only report one 1 negative finding
Other bias	Low risk	No implementation problems reported

Orchowsky 1981

Methods	Random assignment
Participants	80 juvenile delinquents (with minimum 2 offenses), ages 13-20 years, 100% male
Interventions	Confrontational, inmate-run program, locked in cell, introduction by guard, 2-hour session with inmates
Outcomes	6-, 9- and 12-month follow-ups of official measures of offending including new court intakes, average number of court intakes, severity of offense
Notes	The massive attrition at 9 and 12 months also corresponds with positive results reported for the program after negative impact at 6 months. However, the tests for equivalence seem to indicate the groups were still comparable

Orchowsky 1981 (Continued)

<i>Risk of bias</i>		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random assignment used, test for equivalence satisfactory
Allocation concealment (selection bias)	Unclear risk	No description of concealment
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Juvenile court intake data is were the primary source but no description on how collected
Incomplete outcome data (attrition bias) All outcomes	Low risk	The study drops 41% at 9 months and 55% at 12 months, principal investigator PIs reports tests for equivalence at 9 and 12 months are satisfactory. We rate this as low risk because at first follow-up, there wa is little attrition
Selective reporting (reporting bias)	Low risk	Not likely given this is a government evaluation of its own program, and the results at the first follow-up are not positive
Other bias	Low risk	No implementation problems reported

Vreeland 1981

Methods	Randomly assigned to 1 of 4 groups
Participants	160 juvenile delinquents given probation by Dallas County Court, 100% male, 40% white, ages 15-17 years, averaged 2 or 3 prior offenses
Interventions	1-day orientation lasting 13 hours, including haircut and physical labor
Outcomes	6-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

Vreeland 1981 (Continued)

Notes	To remain consistent with other interventions in this review, we took the orientation group comparison with the no-treatment control group. However, the orientation plus counseling group was almost identical to the orientation only group in final results There is nothing in the report to lead us to question the findings
-------	--

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random assignment used, test for equivalence satisfactory
Allocation concealment (selection bias)	Unclear risk	No description provided
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Used court data and self-report, no other information provided
Incomplete outcome data (attrition bias) All outcomes	Low risk	No attrition for the two 2 groups (of the 4 in the experiment) reported
Selective reporting (reporting bias)	Low risk	This was a doctoral dissertation, and the study includes an array of data and analyses
Other bias	Low risk	No implementation problems reported

Yarborough 1979

Methods	Researchers randomly assigned participants according to random numbers table
Participants	227 juvenile delinquents under jurisdiction of courts in 4 Michigan counties
Interventions	Tour of facility, separated and take to cell for interaction with inmates, confrontational session with inmates, 1 visit 5 hours duration
Outcomes	3- and 6-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

Yarborough 1979 (Continued)

Notes	Extensive moderating analyses done The no-shows and its lack of attention in the report are concerning. Again, nothing in the report suggests anything other than a null or slightly negative effect for JOLT	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Research unit handled random assignment, good protocol in place, test for equivalence satisfactory
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding (performance bias and detection bias) All outcomes	Unclear risk	No description of monitoring of control group to determine if compensation was an issue. Probably unlikely that the control group received anything else but not specifically addressed
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not done
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Researchers collected data from court files but unknown if blind to conditions
Incomplete outcome data (attrition bias) All outcomes	High risk	The study has many no-shows who are dropped from analysis
Selective reporting (reporting bias)	Low risk	Government agency reported a negative result for its own program
Other bias	Low risk	No implementation problems reported

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Ashcraft 1970	Used a pre-post test without a control group
Bazemore 2004	Used a matched comparison group without randomization
Berry 1985	Used a matched comparison group without randomization

(Continued)

Blunkett 2008	No randomization, pre-post measures, or appropriate outcomes
Brodsky 1970	Used a pre-post design without a control group
Buckner 1983	Used a matched comparison group without randomization
Chesney-Lind 1981	Used a nonequivalent comparison group design without randomization
Dean 1982	Used randomization but did not include any measures of criminal behavior
Feinstein 2005	Did not include outcome measures relevant to this review
Gilman 1977	Used archival data from 3 sources for post-test only follow-ups without a control group
Langer 1980	Used a matched comparison group without randomization
Lloyd 1995	Case studies of 3-day visit programs in the UK. No control group is included
Mitchell 1986	Used pre-post data without a control group
Muhammed 1999	Used post-test data only with no control group
Nelson 1991	Used post-test only data without a control group
NSW BoS 1980	Used post-test only data without a control group
Nygaard 1980	Report on process and implementation data only. No follow-up or control group reported
O'Malley 1993	Process and implementation data on Australia's Victoria prison program. No control group
Portnoy 1986	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 1996	Used multivariate regression on county crime rates to estimate prevention impact of program, no control group or randomization employed
Shapiro 1978	Used post-test only data without a control group
Storvoll 1998	Process and implementation data are reported on Norway's Scared Straight program. No follow-up or control group included
Trotti 1980	Used post-test data of reactions of participants, without a control group
Wilson 2010	Inappropriate follow-up data, no randomization, no pre-post measures
Windell 2005	Descriptive report without adequate evaluative data or methods

DATA AND ANALYSES

Comparison 1. Intervention versus control, crime outcome

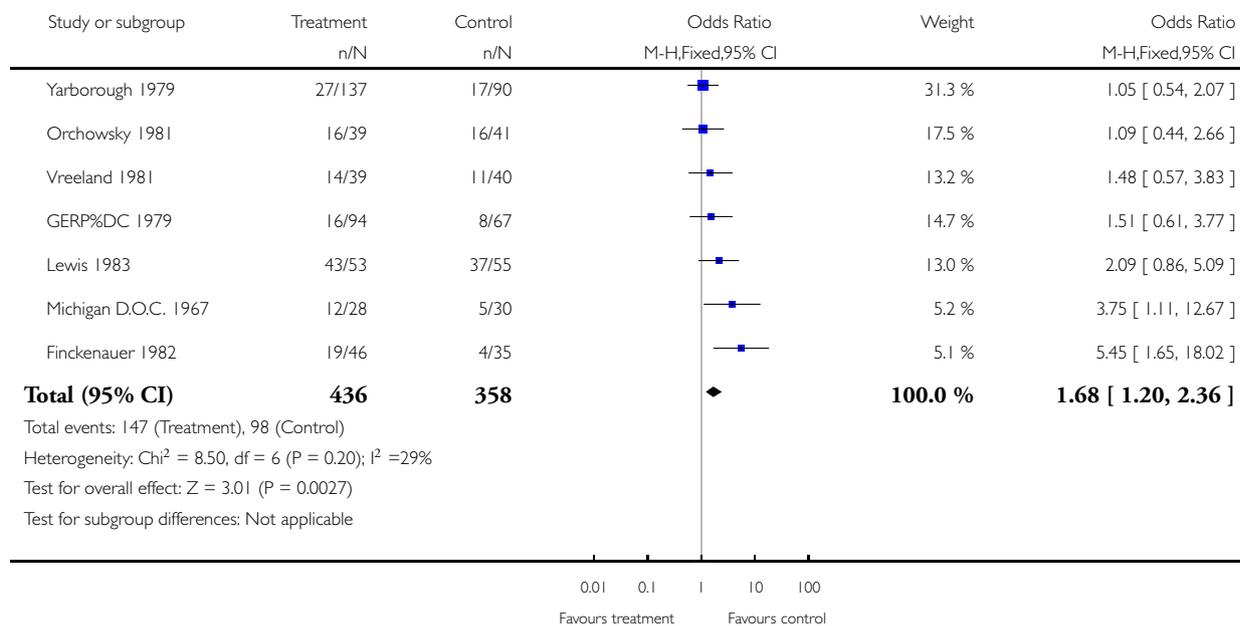
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Postintervention - group recidivism rates - official measures only (fixed-effect)	7	794	Odds Ratio (M-H, Fixed, 95% CI)	1.68 [1.20, 2.36]
2 Postintervention - group recidivism rates - official measures only (random-effects)	7	794	Odds Ratio (M-H, Random, 95% CI)	1.72 [1.13, 2.62]
3 Sensitivity analysis - excluding Finckenauer study	6	713	Odds Ratio (M-H, Random, 95% CI)	1.47 [1.03, 2.11]
4 Sensitivity analysis - excluding Yarborough study	6	567	Odds Ratio (M-H, Random, 95% CI)	1.96 [1.25, 3.08]
5 Sensitivity analysis - excluding both Finckenauer and Yarborough studies	5	486	Odds Ratio (M-H, Random, 95% CI)	1.68 [1.10, 2.58]

Analysis 1.1. Comparison 1 Intervention versus control, crime outcome, Outcome 1 Postintervention - group recidivism rates - official measures only (fixed-effect).

Review: 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency

Comparison: 1 Intervention versus control, crime outcome

Outcome: 1 Postintervention - group recidivism rates - official measures only (fixed-effect)

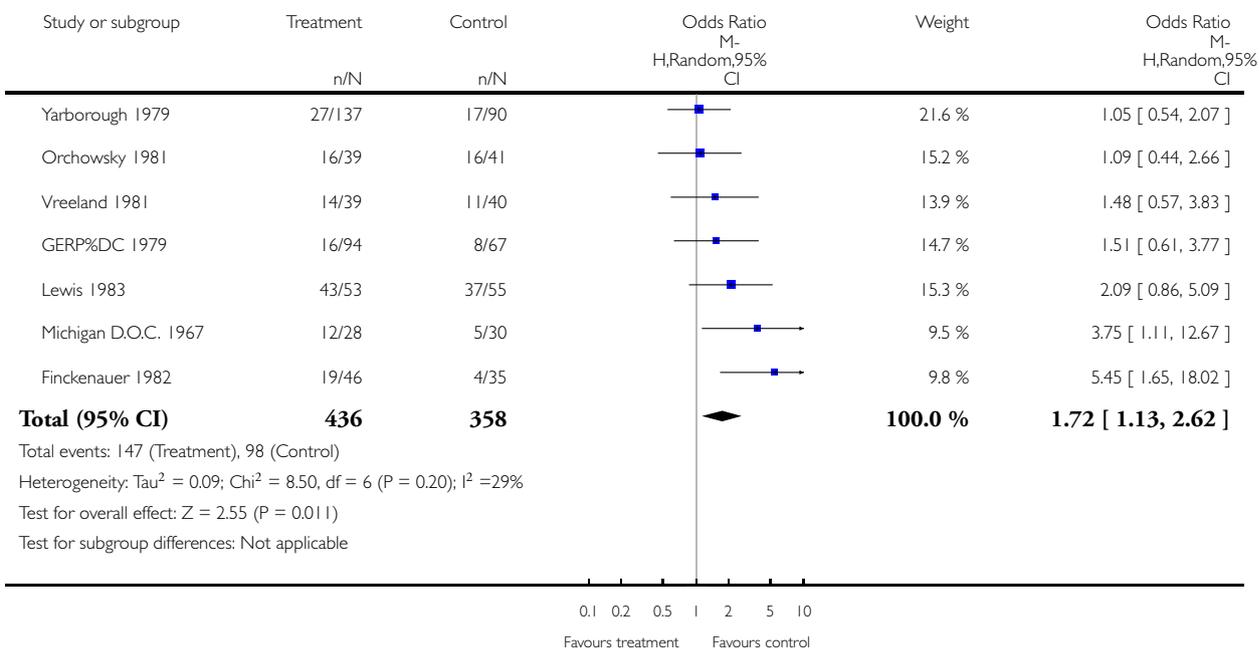


Analysis 1.2. Comparison 1 Intervention versus control, crime outcome, Outcome 2 Postintervention - group recidivism rates - official measures only (random-effects).

Review: 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency

Comparison: 1 Intervention versus control, crime outcome

Outcome: 2 Postintervention - group recidivism rates - official measures only (random-effects)

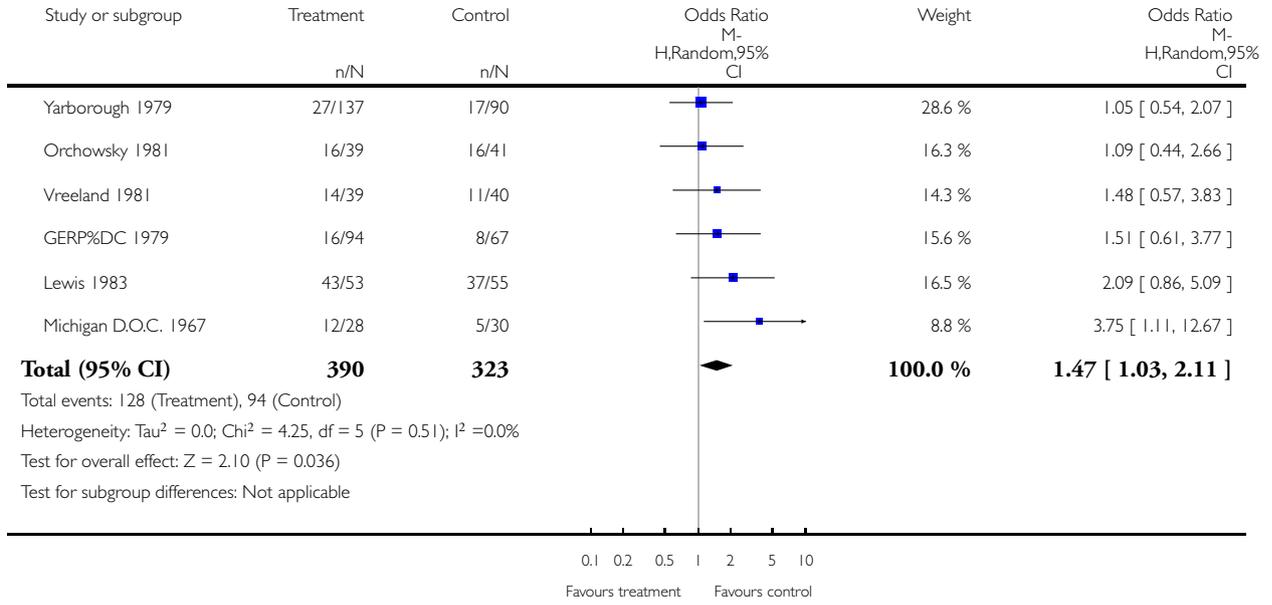


Analysis 1.3. Comparison 1 Intervention versus control, crime outcome, Outcome 3 Sensitivity analysis - excluding Finckenaer study.

Review: 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency

Comparison: 1 Intervention versus control, crime outcome

Outcome: 3 Sensitivity analysis - excluding Finckenaer study

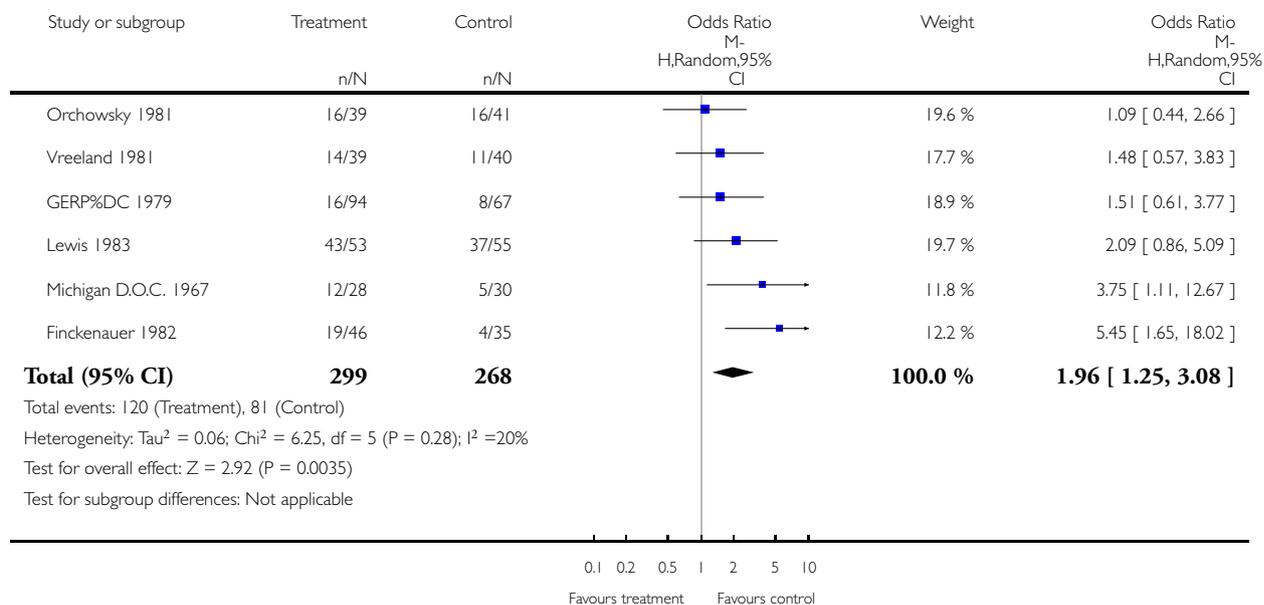


Analysis 1.4. Comparison 1 Intervention versus control, crime outcome, Outcome 4 Sensitivity analysis - excluding Yarborough study.

Review: 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency

Comparison: 1 Intervention versus control, crime outcome

Outcome: 4 Sensitivity analysis - excluding Yarborough study

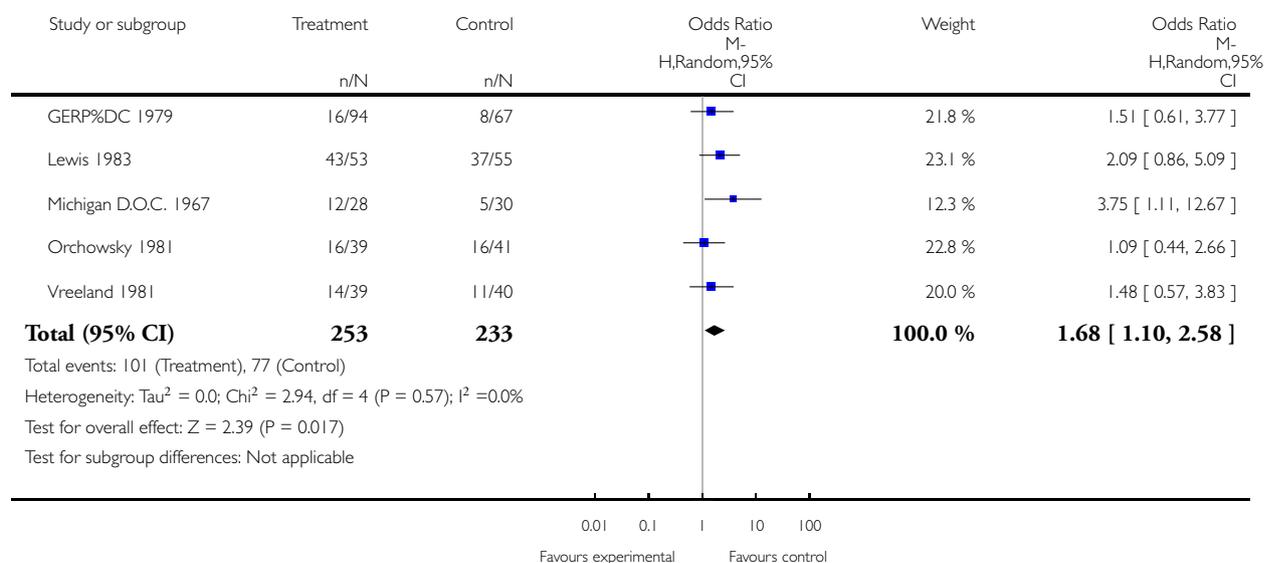


Analysis 1.5. Comparison 1 Intervention versus control, crime outcome, Outcome 5 Sensitivity analysis - excluding both Finckenaer and Yarborough studies.

Review: 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency

Comparison: 1 Intervention versus control, crime outcome

Outcome: 5 Sensitivity analysis - excluding both Finckenaer and Yarborough studies



ADDITIONAL TABLES

Table 1. Crime outcome data reported in original studies

Study Reference	At 3 months	At 6 months	At 9 months	At 12 months	Beyond 12 months
Michigan D.O.C. 1967		Percentage with new offense or new violation of probation			
GERP&DC 1979		Percentage subsequently contacted by police			
Yarborough 1979	Percentage with new offenses, type of offenses, percentage with new petitions, average offense rate and standard deviation	Percentage with new offenses, type of offenses, percentage with new petitions, average offense rate			

Table 1. Crime outcome data reported in original studies (Continued)

	tions, average weeks to new offense and standard deviations, number of days in detention and standard deviations	and standard deviations, average weeks to new offense and standard deviations, average days in detention and standard deviations			
Orchowsky 1981		Percentage with new intakes, average intakes (no standard deviations but test statistic), average severity score (no standard deviations but test statistic)	Percentage with new intakes, average intakes (with no standard deviations but test statistic) and average severity score (no standard deviations but test statistic)	Percentage with new intakes, average intakes (no standard deviations but test statistic), average severity score (no standard deviations but test statistic)	
Vreeland 1981		Percentage with new offenses (official measures), percentage with new offenses (self-reported data)			
Finckenaer 1982		Percentage new complaints, contacts or court appearances, average severity score (no standard deviation, but test statistic)			
Lewis 1983				Percentage arrested, percentage charged, average arrests (no standard deviation), average charges (no standard deviation), average time to first arrest (no standard deviation)	
Locke 1986		Only test statistic reported			
Cook 1992				Average offenses (no standard deviations)	Average offenses (no standard deviations)

Table 1. Crime outcome data reported in original studies (Continued)

				, average severity score (no standard deviations)	, average severity score (no standard deviations)
--	--	--	--	---	---

APPENDICES

Appendix I. Search terms used for all databases

'scared straight'
Prison orientation OR prison tour OR prison visit
Jail orientation OR jail tour OR jail visit
Reformatory orientation OR reformatory tour OR reformatory visit
Reformatory* orientation OR reformatory* tour OR reformatory* visit
'prisoner run' OR 'offender run' OR 'inmate run'
'prison awareness OR 'prison aversion' OR 'juvenile awareness'
'rap session' AND prisoner
'rap session' AND lifer
'rap session' AND inmate
'rap session' AND offender
speak out AND prisoner
speak out AND lifer
speak out AND inmate
speak out AND offender
confrontation AND prisoner

(Continued)

confrontation AND lifer
confrontation AND inmate
confrontation AND offender

Appendix 2. Search methods used for original review

For the original review we firstly identified randomized experiments from a larger review of field trials in crime reduction by the first author (Petrosino 1997). Petrosino used the following methods to find more than 300 randomized experiments: handsearch (that is, visually inspecting the entire contents) of 29 leading criminology or social science journals; checking the citations reported in the 'Registry of Randomised Experiments in Criminal Sanctions' (Weisburd 1990); 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); searches by information specialists of 18 bibliographic databases, including the National Criminal Justice Reference Service (NCJRS); an extensive mail campaign with over 200 researchers and 100 research centers; published solicitations in association newsletters; tracking of references in over 50 relevant systematic reviews and literature syntheses; and tracking of references in relevant bibliographies, books, articles and other documents. More detail about these search methods can be found in Petrosino 1995 and Petrosino 1997. The citations found in Petrosino 1997 covered literature with a publication date between January 1, 1945 and December 31, 1993. Seven randomized trials meeting the eligibility criteria were identified from this sample.

Second, we augmented this work with searches designed to uncover experiments missed by Petrosino 1997 and to cover more recent literature (1994 to 2001). These methods included: broad searches of the Campbell Collaboration Social, Psychological, Educational & Criminological Trials Register (C2-SPECTR) developed by the UK Cochrane Centre and then supervised by the University of Pennsylvania Graduate School of Education (Petrosino 2000a); check of citations from more recent systematic or traditional reviews to provide coverage of more recent studies (for example, Sherman 1997; Lipsey 1998); citation checking of documents relevant to Scared Straight and similar programs (for example, Finckenaer 1999); email correspondence with investigators; and broad searches of the Cochrane Controlled Trials Register (CENTRAL) in *The Cochrane Library* (Issue 1, 2002). By broad searches, we mean that we tried to first identify studies relevant to crime or delinquency and then visually scanned the citations or abstracts to see if any were relevant to this intervention.

Third, we decided to conduct a more specific search of the 14 additional electronic databases accessible to the authors and relevant to the topic area. Many of these include published and unpublished literature (for example, 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 Silver Platter databases not accessible at Harvard University or via the Internet. The bibliographic data bases and the years searched were:

- Criminal Justice Abstracts, 1968 to September 2001;
- Current Contents, 1993 to 2001;
- Dissertation Abstracts, 1981 to August 2001;
- Education Full Text, June 1983 to October 2001;
- ERIC (Education Resource Information Clearinghouse) 1966 to 2001;
- GPO Monthly (Government Printing Office Monthly), 1976 to 2001;
- MEDLINE 1966 to 2001;
- National Clearinghouse on Child Abuse and Neglect, to 2001;
- NCJRS (National Criminal Justice Reference Service), to 2001;
- Political Sciences Abstracts, 1975 to March 2001;
- PAIS International (Public Affairs Information Service), 1972 to October 2001;
- PsycINFO (Psychological Abstracts) 1987 to November 2001;
- Social Sciences Citation Index, February 1983 to October 2001;
- Sociofile (Sociological Abstracts and Social Planning And Development Abstracts) January 1963 to 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 (for example, 'random 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 citation, and did not vary:

- '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').

FEEDBACK

Feedback given on original review in 2003 - Meaning of equivalence at baseline

Summary

My question relates to information in the table describing the methodological quality of the included studies, where reference is made to 'tests for equivalence' at baseline. What does this mean? My concern is that it may refer to the use of tests of statistical significance to compare baseline characteristics following randomisation, a process which Altman (1985) has pointed out is absurd. Either chance (random allocation) was used to generate the comparison groups (in which case it makes no sense to use statistical tests to assess the probability that any differences reflect chance), or chance (random allocation) was not used.

Please clarify this, and provide more information, for each trial, about how the allocation schedule was generated, and what measures were taken to conceal the schedule from those recruiting participants into the trial. If this information has not been supplied by the authors of the reports, please make this explicit.

Ref. 1 - Altman DG. Comparability of randomised groups. *Statistician* 1985;34:125-136.

I certify that I have no affiliations with or involvement in any organisation or entity with a direct financial interest in the subject matter of my criticisms.

Reply

Reply to Sir Iain Chalmers' comment on our review, by Anthony Petrosino and Carolyn Turpin-Petrosino

We apologize for the unsatisfactory delay in responding to Sir Iain Chalmers' comment on our review. His question is most appreciated, and inspired us to query some of our more methodologically- and statistically-minded colleagues for advice. We have now had ample opportunity to mull over these responses.

He asked that we clarify what is meant by 'tests for equivalence at baseline'. Indeed, our reference is to statistical tests that are conducted by the experimental investigators to determine if randomization produced equivalent groups before the intervention or treatment is introduced. In his comment, Dr. Chalmers is correct when he states (referencing Altman 1985) that such 'pretests of group equivalence' are illogical because of randomization. But this only applies when we have confidence that randomization was carried out with full integrity.

Unfortunately, thorough description of how randomization was done and what efforts were taken to conceal such allocation are often missing in reports of experimental studies. This is particularly true of trials reported several decades ago; in our review, all of the studies were reported before 1993 and at least one was briskly reported in a short government document circa 1967. Sure enough, concealment and allocation was rated as 'unknown' in eight of the nine trials we included in our systematic review. Pretests of group equivalence increase our confidence (but does not guarantee) that randomization was successfully implemented.

Missing information is not the only problem. It is also the case that allocation in many criminological experiments is often left out of the hands of the investigators and is actually conducted by practitioners or treatment providers. Such individuals often have a good reason to corrupt the allocation schedule to ensure that particular cases end up in a certain group. Pretests of group equivalence are one way to determine if an intentional subversion of the allocation scheme has resulted in unhappy configurations of the groups.

Besides missing information and covert manipulation of allocation, there is another problem with criminological experiments that pretests of group equivalence can assist. Many justice experiments have very small samples. For example, the Locke et al study in our review (though it was not included in the meta-analysis) had 16 participants in each group. The laws of randomization naturally follow the laws of sampling probability. If you flip a valid coin 32 times, you may end up with 22 heads and 10 tails. Randomizing 32 participants to study groups may result in the experimental group receiving far more boys than girls when compared to the control group. To the extent that males are more likely to commit another crime than females, the experimental group is at a distinct disadvantage. Flipping a valid coin several hundred times is more likely to produce a near 50/50 split of heads and tails than 32 flips; random allocation of several hundred participants is more likely to produce balanced groups than assignment of 32 participants. Pretests of group equivalence, in this case, can identify situations where unintentional bias has produced unhappy configurations of groups. The methodological quality table contains our own subjective language of whether we thought the pretest results were 'satisfactory.' This should be changed. In our update of the Cochrane review, we will simply list if the pretests were done and whether the experimental investigators reported that pretest equivalence was confirmed.

Notes

- i. We especially thank Dr. Mark Lipsey and Dr. David Weisburd, among others, for their valuable input.
- ii. Of course, experimental investigators who have good a priori knowledge of a particular variable especially relevant to the outcome, can block on that variable to ensure equal distribution across study groups irrespective of randomization (in essence, they can randomize boys and girls separately into the study groups).

Contributors

Iain Chalmers, Director, UK Cochrane Centre, ichalmers@cochrane.co.uk

WHAT'S NEW

Last assessed as up-to-date: 30 June 2012.

Date	Event	Description
1 March 2013	New citation required but conclusions have not changed	Updated searches found no new studies suitable for inclusion. Listed six new studies in 'excluded studies' section. Added 'Risk of bias' tables. Conducted sensitivity analysis (excluding both Finckenauer and Yarborough studies)
1 January 2012	New search has been performed	Updated all searches

HISTORY

Protocol first published: Issue 4, 2000

Review first published: Issue 2, 2002

Date	Event	Description
22 September 2008	Amended	Converted to new review format.
25 May 2004	Amended	Response to feedback added: 25/05/04
26 February 2003	Feedback has been incorporated	Feedback added: 26/02/03
1 March 2002	New search has been performed	Minor update: 01/03/02
27 February 2002	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

Anthony Petrosino: searching for studies, screening studies, extracting data, conducting analyses, drafting review.

Carolyn Turpin-Petrosino: screening studies, extracting data, drafting review.

Meghan Hollis-Peel, updating searches, screening studies, drafting review.

Julia Lavenberg: updating searches, screening studies, drafting review.

DECLARATIONS OF INTEREST

Anthony Petrosino - the original review was supported in part by a consultancy to me from the University of Pennsylvania. This update is being supported in part by funding from the Campbell Collaboration, based in Oslo, Norway. I also received an honorarium in 2004 to 2005 for contributing an article summarizing Scared Straight to a special issue on randomized experiments by the *Annals of the American Academy of Political and Social Science*.

Carolyn Turpin-Petrosino - as spouse to lead author, I am also a beneficiary of funding Anthony received as a consultant or as an article contributor to the *Annals of the American Academy of Political and Social Science*.

Julia Lavenberg - I was supported as a consultant for some work on the update.

Meghan E. Hollis-Peel - I was supported as a consultant for some work on the update.

SOURCES OF SUPPORT

Internal sources

- American Academy of Arts and Sciences, USA.
- Harvard Graduate School of Education, USA.

External sources

- Smith-Richardson Foundation grant (to University of Pennsylvania), USA.
- Mellon Foundation grant (to AAA&S, Center for Evaluation), USA.
- Home Office, Research & Statistics Directorate (to Cambridge University), UK.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The main difference between the protocol and the review is that the protocol anticipated a range of outcomes (prevalence, incidence, severity and latency) at different time intervals, and the review only focused on prevalence outcomes (for example, percentage of youth in each group getting re-arrested) reported at first post-treatment follow-up.

NOTES

A publication based on the preliminary results of the original review was published in A. Petrosino, C. Petrosino and J. Finckenauer, 2000, *Crime & Delinquency*, 46, 1, 354-79.

The review is published in both the Cochrane and the Campbell Libraries.

INDEX TERMS

Medical Subject Headings (MeSH)

*Prisons; *Program Evaluation; Awareness; Juvenile Delinquency [*prevention & control; psychology]; Randomized Controlled Trials as Topic; Treatment Failure

MeSH check words

Adolescent; Humans; Young Adult