

— SYSTEMATIC REVIEW PROTOCOL —

EFFECTS OF CORRECTIONAL BOOT CAMPS  
ON OFFENDING: A CAMPBELL COLLABORATIVE  
SYSTEMATIC REVIEW, CRIMINAL  
JUSTICE REVIEW GROUP<sup>1</sup>

Doris L. MacKenzie,<sup>2</sup> David B. Wilson,<sup>3</sup> and Suzanne Kider

08 January 2003

---

<sup>1</sup>This work is supported in part by a grant from the Jerry Lee Foundation.

<sup>2</sup>2220 LeFrak Hall, University of Maryland, College Park, MD 20742-8235, 301.405.3008(w), 301.405.4733(fax), dmacken@crim.umd.edu

<sup>3</sup>Administration of Justice Program, 10900 University Blvd., MS 4F4, George Mason University, Manassas, VA 20110-2203, dwilsonb@gmu.edu

**Contents**

<b>1</b>	<b>Background for the Review</b>	<b>2</b>
<b>2</b>	<b>Objectives of the Review</b>	<b>3</b>
<b>3</b>	<b>Methods</b>	<b>4</b>
3.1	Criteria for inclusion and exclusion of studies in the review . . . . .	4
3.2	Search strategy for identification of relevant studies . . . . .	4
3.3	Description of methods used in the component studies . . . . .	5
3.4	Criteria for determination of independent findings . . . . .	5
3.5	Details of study coding categories . . . . .	6
3.6	Statistical procedures and conventions . . . . .	6
3.7	Treatment of qualitative research . . . . .	7
<b>4</b>	<b>Timeframe</b>	<b>7</b>
<b>5</b>	<b>Plans for Updating the Review</b>	<b>7</b>
<b>6</b>	<b>Acknowledgments</b>	<b>8</b>
<b>7</b>	<b>Statement Concerning Conflict of Interest</b>	<b>8</b>
<b>8</b>	<b>References</b>	<b>8</b>
<b>9</b>	<b>Tables</b>	<b>9</b>
<b>A</b>	<b>Eligibility Checklist (version 9/26/00)</b>	<b>10</b>
<b>B</b>	<b>Sample Status Screen from the Bibliographic Database Program</b>	<b>11</b>
<b>C</b>	<b>Coding Protocol</b>	<b>12</b>
<b>D</b>	<b>Sample Screen from the Coding Protocol Database</b>	<b>24</b>
<b>E</b>	<b>Sample Output from Meta-Analytic Macros</b>	<b>25</b>

## 1 Background for the Review<sup>4</sup>

Correctional boot camps, frequently called shock or intensive incarceration, were first opened in adult correctional systems in Georgia and Oklahoma in 1983. Since that time they have rapidly grown, first within adult systems and later in juvenile corrections. Today, correctional boot camps exist in federal, state and local juvenile and adult jurisdictions in the United States.

In the typical boot camp, participants are required to follow a rigorous daily schedule of activities including drill and ceremony and physical training. They arise early each morning and are kept busy most of the day. Correctional officers are given military titles and participants are required to use these titles when addressing staff. Staff and inmates are required to wear uniforms. Punishment for misbehavior is immediate and swift and usually involves some type of physical activity like push-ups. Frequently, groups of inmates enter the boot camps as squads or platoons. There is often an elaborate in-take ceremony where inmates are immediately required to follow the rules, respond to staff in an appropriate way, stand at attention and have their heads shaved. Many programs have graduation ceremonies for those who successfully complete the program. Frequently, family members and others from the outside public attend the graduation ceremonies.

The camps for adjudicated juveniles differ somewhat from the adult camps. Less emphasis is placed on hard labor and, as required by law, the camps provide juveniles with academic education. Juvenile camps are also apt to provide more therapeutic components. However, in many other aspects the juvenile camps are similar to adult camps with rigorous in-take procedures, shaved heads, drill and ceremony, physical training, immediate physical punishment for misbehavior (e.g., push-ups), and graduation ceremonies.

While there are some basic similarities among the correctional boot camps, the programs differ greatly in other aspects (MacKenzie and Hebert 1996). For example, the camps differ in the amount of focus given to the physical training and hard labor aspects of the program versus therapeutic programming such as academic education, drug treatment or cognitive skills. Some camps emphasize the therapeutic programming, others focus on discipline and rigorous physical training. Programs also differ in whether they are designed to be an alternative to probation or to prison. In some jurisdictions judges sentence participants to the camps, in others, participants are identified by department of corrections personnel from those serving terms of incarceration. Another difference among programs is whether the residential phase is followed by an aftercare or re-entry program designed to assist the participants with adjustment to the community.

---

<sup>4</sup>This section borrows heavily from MacKenzie, Wilson, and Kider (2001).

Despite their continuing popularity, correctional boot camps remain controversial. Primarily, the debate involves questions about the impact of the camps on the adjustment and behavior of participants while they are in residence and after they are released. According to the advocates, the atmosphere of the camps is conducive to positive growth and change (Clark and Aziz 1996, MacKenzie and Hebert 1996). In contrast, critics argue that many of the components of the camps are in direct opposition to the type of relationships and supportive conditions that are needed for quality therapeutic programming (Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen 1990, Gendreau, Little and Groggin 1996, Morash and Rucker 1990, Sechrest 1989).

Research examining the effectiveness of the correctional boot camps has focused on various potential impacts of the camps. Some have examined whether the camps change participant's attitudes, attachments to the community or impulsivity (MacKenzie, Styve, Gover, and Wilson, 2001, MacKenzie and Shaw 1990, MacKenzie and Souryal 1995). Others have examined the impact of the camps on the need for prison bed space (MacKenzie and Piquero 1994, MacKenzie and Parent 1991). However, the research receiving the most interest appears to be that examining the impact of the camps on recidivism. Past reviews of this research, such as MacKenzie (1997), have concluded that boot camps are ineffective in reducing recidivism with the caveat that those boot camps that emphasize therapeutic activities and aftercare may be successful.

## **2 Objectives of the Review**

The objective of this systematic review is to synthesize the extant empirical evidence (published or otherwise) on the effects of boot camps and boot camp like programs on the criminal behavior of convicted adult and juvenile offenders. The focus of this review is on the effectiveness of these programs with respect to recidivism. According to a survey of state correctional officials by Gowdy (1996), reduced correctional costs and recidivism were the primary goals of boot camps. The review does not examine research on the cost effects of these programs, nor does it review the potential secondary effects on outcomes such as antisocial attitudes.

### **3 Methods**

#### **3.1 Criteria for inclusion and exclusion of studies in the review**

The scope of this review is experimental and quasi-experimental evaluations of boot camp and boot camp like programs for juvenile and adult offenders that utilize a comparison group. The preliminary eligibility criteria are (a) that the study evaluated a correctional boot camp, shock incarceration, or intensive incarceration program (i.e., a residential program for convicted offenders that incorporates a militaristic environment and/or structured strenuous physical activity other than work); (b) that the study included a comparison group that received either probation or incarceration in an alternative facility, such as jail or prison (study design may be experimental or quasi-experimental; one-group research designs will not be eligible); (c) that the study participants were exclusively under the supervision of the criminal or juvenile justice system (i.e., convicted of or adjudicated for an offense); and (d) that the study reported a post-program measure of criminal behavior, such as arrest or conviction (the measure may be based on official records or self-report and may be reported on a dichotomous or continuous scale). These criteria will be modified if necessary after we interact with the literature. A form for evaluating the eligibility of studies is in Appendix A.

#### **3.2 Search strategy for identification of relevant studies**

Several strategies will be used to identify all studies, published or otherwise, that meet these criteria, including a keyword search of computerized databases, contact with authors working in this area, and examination of study registries. We have already completed a preliminary search of the following databases: Criminal Justice Periodical Index, Dissertation Abstracts Online, Government Publications Office Monthly Catalog, Government Publications Reference File, NCJRS, PsychINFO, Sociological Abstracts, Social SciSearch, and U.S. Political Science Documents. The keywords used were: boot camp(s), intensive incarceration, and shock incarceration. Several of the searched databases index nonpublished works. We will also search the Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register. These searches will be updated and we will identify and contact non U.S. researchers working in this area to help us locate international evaluations. We anticipate few such studies (with the exception of a couple of evaluations from Great Britain of which we are aware), for boot camps have been a predominately a U.S. phenomenon, although that may change in the coming decade.

All references identified as potentially eligible for this review will be entered into a bibliographic database program created in FileMaker Pro(TM) by D. B. Wil-

son and J. H. Derzon specifically for use in managing bibliographies for meta-analysis (a sample of the main status screen from this program is presented in Appendix B). Identified documents will be retrieved and two coders will assess eligibility through an examination of the full report.

### **3.3 Description of methods used in the component studies**

The methods used by the studies included in this review are variations on a treatment versus comparison group research design with a post-test and possible follow-up measurement points. In all cases the participant samples will be adjudicated juvenile delinquents or convicted adult offenders, a sample of whom will participate in a boot camp program and a sample of whom will participate in some other traditional correctional program, such as probation or short-term incarceration and parole.

The studies will vary with respect to the method of constructing the comparison group. A small number of studies will rely on random assignment of participants to the boot camp and comparison conditions. The remaining studies will be observational, relying on groups constructed by natural means. The common variations will be historical controls, adjacent jurisdictions, offenders eligible for boot camp who chose not to participate, and eligible offenders who did not participate due to limited space in the boot camp program. The studies will also vary with respect to the degree to which they employ statistical controls (matching, covariate analysis, etc.) to reduce the threat of selection bias.

All studies included in this review will include a post-program measure of criminal recidivism, that is, criminal behavior. These may include dichotomous indicators of arrest, conviction, or reincarceration or more differentiated indicators that specify specific types of offenses or the frequency of offenses. A few studies may also include self-report measures of criminal involvement.

A subset of studies will report on other outcomes, such as school performance, employment status, changes in antisocial attitudes, etc. We will not be coding these other outcomes. However, we will document all reported outcomes (see coding protocol).

### **3.4 Criteria for determination of independent findings**

A complication in conducting a systematic review of studies in this area is that most studies will report multiple indicators of recidivism, and some will report those indicators at multiple time points (e.g., 12 months post-program, 24 months post-program). The statistical methods outlined below require that the findings (effect sizes) be independent, that is, come from unique samples. Several strategies will be used to address this problem.

First, all findings will be coded and entered into the data file. Second, criteria will be developed to select the more general indicators of recidivism and those with the longest follow-up that are based on the full sample (i.e., not affected by attrition). Third, any multiple effect sizes remaining within a study will be averaged and the average will serve as the effect size for the primary analyses. Fourth, a secondary set of analyses will be conducted that will analyze all measures of arrest, conviction, and reinstitutionalization separately, averaging multiple indicators of each within a study. These analyses will provide a sensitivity check on the primary set of analyses.

### **3.5 Details of study coding categories**

A preliminary coding protocol has been developed for this project (see Appendix C) and is a modification of the coding protocol for two previously completed meta-analyses and two ongoing meta-analyses. This protocol recognizes the nested nature of effects within studies and as such is hierarchical in nature. Any number of effects can be coded for any number of dependent variables for each program-comparison contrast contained within a study. That is, a study may report multiple independent evaluations or boot-camp versus comparison group contrasts. Furthermore, a study may have multiple indicators of criminal involvement, such as arrest, conviction, and technical violation, measured at multiple time points following release from the program. The coding protocol and corresponding database are designed to accommodate this complexity.

The coding protocol captures aspects of the research design, including methodological quality, the boot camp program, the comparison group condition, the participant offenders, the outcome measures and the direction and magnitude of the observed effects. All studies will be coded by two independent coders and all coding differences will be resolved by either Dr. MacKenzie or Wilson. The uncorrected double coding will be used to assess interrater reliability. Only items with acceptable reliability (e.g., agreement over 70%) will be used in the analyses. The data will be maintained in a FileMaker Pro(TM) database. Coding can be performed either on the paper coding forms (Appendix C) or can be performed directly into the computer using database screens that match the paper coding form (see Appendix D for an example).

### **3.6 Statistical procedures and conventions**

The primary effect of interest is recidivism or a return to criminal activity on the part of the offender after leaving the program. Recidivism data are most typically reported dichotomously. As such, the natural index of effectiveness is the odds-ratio (see Fleiss, 1994) and is the index of effect that will be used in this review. If a study

reports recidivism on a continuous scale, then a standardized mean difference effect size will be computed and transformed into an equivalent odds-ratio (see Lipsey and Wilson, 2001, page 198).

The mean odds-ratio and homogeneity of effects across studies will be computed using the inverse variance weight method. A random effects model will be assumed and the random effects variance component will be estimated using the methods outlined by DerSimonian and Laird (1986). The computations will be performed using macros written by D. B. Wilson and available for use with SAS, SPSS, and Stata (Lipsey and Wilson, 2001). These macros use the same methods of computation as those used by the Cochrane Collaborative MetaView program (Deeks, 1999). Sample output from these macros is presented in Appendix E. In addition to an examination of overall effects, the relationship of study features, both substantive and methodological, to observed effects will be explored, both through categorical models and regression models.

### **3.7 Treatment of qualitative research**

At this time we have no plans to include qualitative research in this systematic review. We would welcome any collaborator with expertise in the area of qualitative research to contribute a review of qualitative evaluations to this synthesis.

## **4 Timeframe**

Most of the proposed work has already been completed and has been published (MacKenzie, Wilson, and Kider, 2001). Tasks to be completed are searching for more recent studies and a more comprehensive search for international studies. We will need to code any new studies identified and run new analyses. Finally, we need to produce a written product conforming to Campbell Collaborative guidelines. Given the extensive work already completed (29 coded studies representing 44 independent boot camp versus comparison contrasts), we anticipate completion of the Campbell review by the end of August, 2003.

## **5 Plans for Updating the Review**

We plan to update this review every two-to-three years in accordance with Campbell Collaborative guidelines.

## 6 Acknowledgments

We would like to thank the Jerry Lee Foundation for partial support of this project.

## 7 Statement Concerning Conflict of Interest

Drs. MacKenzie and Wilson, and Ms. Kider have no financial interest in any existing or planned boot camp program. Dr. MacKenzie has argued in prior publications that boot camps are ineffective, at least in the absence of therapeutic elements and aftercare components. Thus, the only potential conflict of interest is consistency with prior scholarly publications. The research team will strive to avoid any potential conflict.

## 8 References

- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., and Cullen, F. T. (1990). Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis. *Criminology*, 28, 369–404.
- Clark, C. L., & Aziz, D. W. (1996). Shock Incarceration in New York State: Philosophy, Results, and Limitations. In D. L. MacKenzie & E. E. Hebert (Eds.), *Correctional Boot Camps: A Tough Intermediate Sanction*. Washington, DC: National Institute of Justice.
- Deeks, J. (1999). *Statistical Methods Programmed into MetaView, Version 4*. The Cochrane Collaborative. Available online at [www.cochrane.org](http://www.cochrane.org).
- DerSimonian, R., & Laird, N. (1986). Meta-Analysis in Clinical Trials. *Controlled Clinical Trials*, 7, 177–188.
- Fleiss, J. L. (1994). Measures of Effect Size for Categorical Data. In H. Cooper & L. V. Hedges, *The Handbook of Research Synthesis*. New York: Russell Sage.
- Gendreau, P., Little, T., & Groggin, C. (1996). A Meta-analysis of the Predictors of Adult Offender Recidivism: What Works! *Criminology*, 34, 575–607.
- Gowdy, V. B. (1996). Historical Perspective. In D. L. MacKenzie & E. E. Hebert (Eds.), *Correctional Boot Camps: A Tough Intermediate Sanction*. Washington, DC: National Institute of Justice.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.

- MacKenzie, D. L. (2000). Reducing The Criminal Activities of Known Offenders and Delinquents: Crime Prevention in the Courts and Corrections. In L. W. Sherman, D. P. Farrington, B. C. Welsh, & D. L. MacKenzie (Eds.), *Evidence-based Crime Prevention*. Harwood Academic Publishers, United Kingdom.
- MacKenzie, D. L. (1997). Criminal Justice and Crime Prevention. In L. W. Sherman et al. (Eds.), *Preventing Crime: What Works, What Doesn't, What's Promising (A Report to the United States Congress)*. College Park, MD: Department of Criminology and Criminal Justice, University of Maryland.
- MacKenzie, D. L., & Herbert, E. E. (Eds.) (1996). *Correctional Boot Camps: A Tough Intermediate Sanction*. Washington, DC: National Institute of Justice.
- MacKenzie, D. L., & Parent, D. (1992). Boot Camp Prisons for Young Offenders. In J. M. Byrne, A. J. Lurigio, & J. Petersilia (Eds.), *Smart Sentencing: The Emergence of Intermediate Sanctions*. Newbury Park, CA: Sage Publications.
- MacKenzie, D. L., & Piquero, A. (1994). The Impact of Shock Incarceration Programs on Prison Crowding. *Crime and Delinquency*, 40, 222–249.
- MacKenzie, D. L., Shaw, J. W., & Gowdy, V. B. (1990). *Evaluation of Shock Incarceration in Louisiana, Executive Summary*. Washington, DC: National Institute of Justice.
- MacKenzie, D. L., Souryal, C. (1994). *Multi-Site Evaluation of Shock Incarceration: Executive Summary*. Washington, DC: National Institute of Justice.
- MacKenzie, D. L., Styve, G. J., Gover, A. R., & Wilson, D. B. (2001). The Impact of Boot Camps and Traditional Institutions on Juvenile Residents: Adjustment, Perception of the Environment and Changes in Social Bonds, Impulsivity, and Antisocial Attitudes. *Journal on Research in Crime & Delinquency*, 38, 279–313.
- MacKenzie, D. L., Wilson, D. B., & Kider, S. (2001). Effects of Correctional Boot Camps on Offending. *Annals of the American Academy of Political & Social Science*, 578, 126–143.
- Morash, M., & Rucker, L. (1990). A Critical Look at the Idea of Boot Camp as a Correctional Reform. *Crime & Delinquency*, 36, 204–222.
- Sechrest, D. D. (1989). Prison “Boot Camps” Do Not Measure Up. *Federal Probation*, 53, 15-20.

## 9 Tables

None.

**A Eligibility Checklist (version 9/26/00)**

First Author’s Last Name \_\_\_\_\_

Document Identification Number \_\_\_\_\_

Coder’s Initials \_\_\_\_\_

Date Eligibility Determined \_\_\_\_\_

Document Status

- Eligible
- Not Eligible
- Relevant Review

To be eligible, a study must meet the following criteria. Answer each question with a “yes” or “no.”

**Yes No**

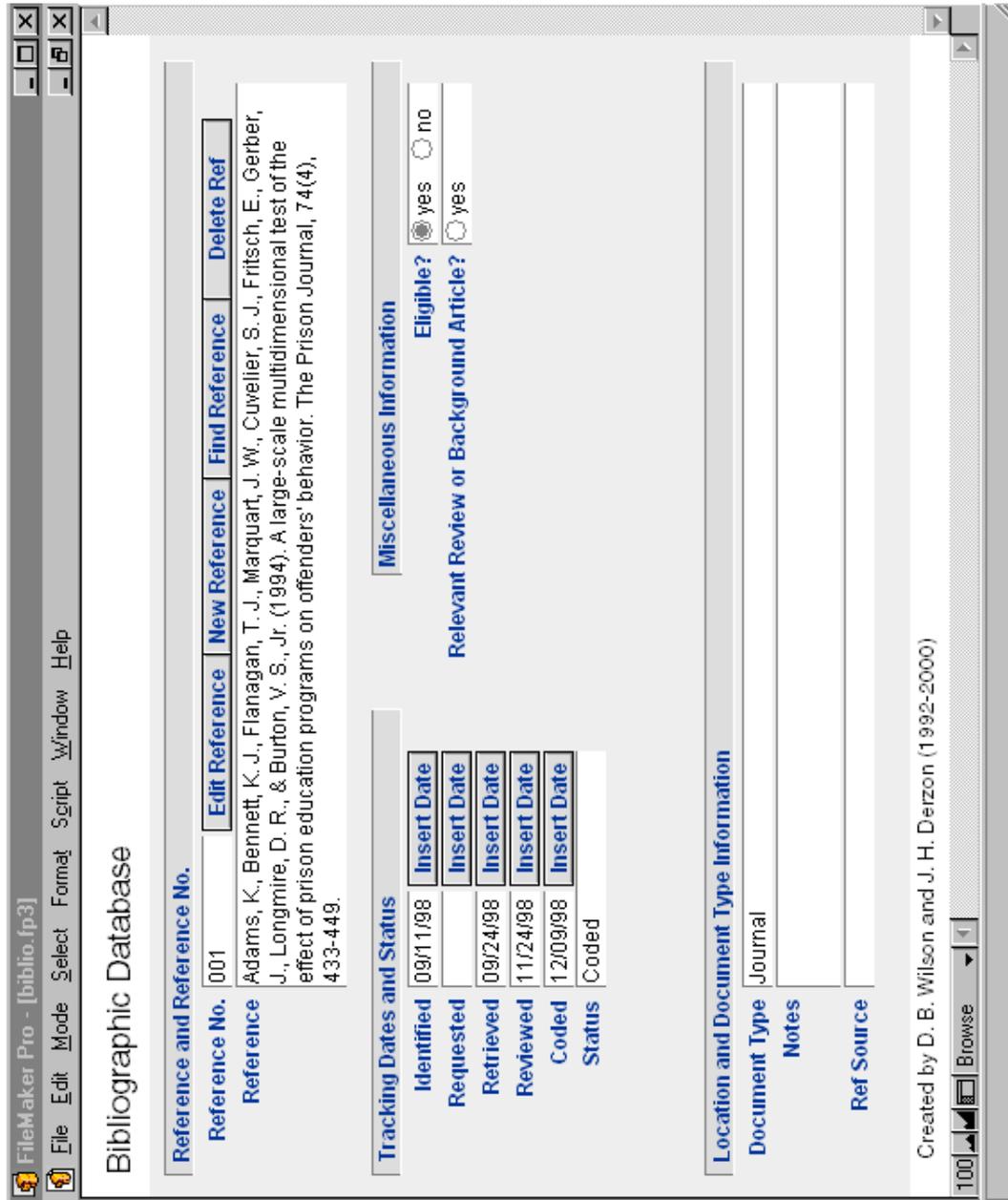
- The study evaluated a correctional boot camp, shock incarceration, or intensive incarceration program (i.e., a residential program for convicted offenders that incorporates a militaristic environment and/or structured strenuous physical activity other than work).
- The study included a comparison group that received either probation or incarceration in an alternative facility, such as jail or prison. Study design may be experimental or quasi-experimental. One-group research designs are not eligible.
- The study participants were exclusively under the supervision of the criminal or juvenile justice system (i.e., convicted of an offense). There is no age or gender restriction for this review.
- The study reported a post-program measure of criminal behavior, such as arrest or conviction. The measure may be based on official records or self-report and may be reported on a dichotomous or continuous scale.

For documents that do not meet the above criteria, answer the following questions:

- The document is not a quantitative evaluation study (i.e., this document does not report any data regarding the effects of a boot camp).
- The document is a review article relevant to this project (i.e., may have references to additional eligible studies or background information useful for preparing written manuscripts for this synthesis).

Notes:

## B Sample Status Screen from the Bibliographic Database Program



## C Coding Protocol

### Study Level Code Sheet

Use one study level code sheet for each study. Note that a single study may be reported in multiple documents. In such cases, the study identifier is the document number for the primary publication. List all other document numbers in items 2 through 4 below.

#### Identifying Information

1. Study (document) identifier studyid \_\_\_ \_\_\_ \_\_\_

If multiple documents were used to code this study, indicate the supplemental study ID numbers.

2. Cross references document identifier crosref1 \_\_\_ \_\_\_ \_\_\_

3. Cross references document identifier crosref2 \_\_\_ \_\_\_ \_\_\_

4. Cross references document identifier crosref3 \_\_\_ \_\_\_ \_\_\_

5. Coder's initials coder \_\_\_ \_\_\_ \_\_\_

6. Date coded date \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_

7. Author author  
 \_\_\_\_\_

8. Name of funding agent fund  
 \_\_\_\_\_

9. Geographic location of study location  
 \_\_\_\_\_

10. Date range for research (when the research was conducted, not published)  
 Start Year styear \_\_\_ \_\_\_ \_\_\_ \_\_\_  
 Finish Year styear \_\_\_ \_\_\_ \_\_\_ \_\_\_

11. Publication type pubtype \_\_\_  
 1. Book  
 2. Book Chapter  
 3. Govt' Report, Federal  
 4. Govt' Report, State/Local  
 5. Journal (peer reviewed)  
 6. Unpublished (tech report, convention paper, dissertation)

12. Number of different treatment-comparison contrasts included in this report mods \_\_\_

13. Is the same control/comparison group used in different contrasts? (1=yes;  
0=no; 8=na) same\_cg \_\_\_\_

### Treatment–Comparison Code Sheet

A study may report on multiple independent evaluations, such as independent treatment–comparison group contrasts, or may have a design that includes multiple interventions of interest contrasted with a single control group. Code each treatment–comparison contrast that is eligible for the review separately. Note that multiple treatment groups must have independent (non-overlapping) participants. A single comparison group may be contrasted with multiple treatment groups (in this case, code item 13 of the Study Level Code Sheet as “1”).

#### Identifying Information

- 14. Study (document) identifier studyid \_\_\_ \_\_\_ \_\_\_
- 15. Module identifier modid \_\_\_ \_\_\_ \_\_\_
- 16. Coder’s initials codermod \_\_\_ \_\_\_ \_\_\_

#### Program Description

- 17. Describe what happens to the treatment group  


---


---

progdes1
- 18. Describe what happens to the comparison group  


---


---

progdes2
- 19. Primary program type txtype \_\_\_ \_\_\_
  - 1. Boot-camp
  - 2. Intensive incarceration
  - 3. Shock incarceration
- 20. What happens to the comparison group? compgrp \_\_\_
  - 1. Probation
  - 2. Jail
  - 3. Prison
  - 4. Diverted Sentence
  - 5. Parole
  - 6. Probation and Prison
  - 7. Other \_\_\_\_\_
  - 8. Treatment Oriented Residential
  - 9. Cannot tell

21. Who delivers or provides the treatment? txstaff \_\_\_  
 1. Mental health professional  
 2. Criminal justice professional  
 3. Professional educator  
 5. Nonprofessional  
 6. Other \_\_\_\_\_  
 9. Cannot tell

22. Length of primary intervention in months (weeks/4.3)  
 1. Minimum txmon1 \_\_\_ \_\_\_ \_\_\_  
 2. Maximum txmon2 \_\_\_ \_\_\_ \_\_\_  
 3. Mean txmon3 \_\_\_ \_\_\_ \_\_\_  
 4. Fixed (same for all subjects) txmon4 \_\_\_ \_\_\_ \_\_\_

23. Length of aftercare or follow-up program component (weeks/4.3; record a 0 if no aftercare)  
txafterm \_\_\_ \_\_\_ \_\_\_

24. Was the program highly structured, that is, followed at set protocol? (1=yes; 0=no, 9=cannot tell) struct \_\_\_

25. Did the program remain consistent over time? (1=yes; 0=no, 9=cannot tell) consist \_\_\_

**Methodological Rigor**

26. Used control variables in statistical analyses to account for initial group differences (1=yes; 0=no) cntrlvar \_\_\_

27. Used subject level matching (1=yes; 0=no) matching \_\_\_

26-27. List of variables used to control/match on pretest differences.  
 \_\_\_\_\_  
 \_\_\_\_\_ cvarlist

28. Used random assignment to conditions (1=yes; 0=no) random \_\_\_

29. Measured prior criminal involvement; not necessarily arrest (1=yes; 0=no) pretest \_\_\_

30. Rate initial group similarity (7=highly similar; 1=highly dissimilar) simrate \_\_\_

anchors: 7—a randomized design with a large *N* or a small *N* design with matching; 5—a nonrandomized design with strong evidence of initial equivalence; 1—a nonrandomized design with a comparison group highly likely to be different from the treatment group or known differences that are related to future recidivism.

31. Was attrition discussed in the study report? (1=yes; 0=no) attrit1 \_\_\_
32. Is there a potential generalizability threat from overall attrition? (1=yes; 0=no; 9=cannot tell) attrit2 \_\_\_
33. Is there a potential threat from differential attrition? (1=yes; 0=no; 9=cannot tell) attrit3 \_\_\_
34. Did the statistical analysis of outcome effects attempt to control for differential attrition effects? (1=yes; 0=no; 9=cannot tell) attrit4 \_\_\_
35. Used statistical significance testing (1=yes; 0=no) sigtest \_\_\_
36. Global methodology rating methscor \_\_\_
1. A comparison group is present but lacks a demonstrated comparability to the treatment group.
  2. A comparison between two or more groups, one with and one without the program.
  3. Comparison between a program group and one or more control groups, controlling for other factors; or a nonequivalent comparison group that is only slightly different from the program group or a random assignment design with high attrition.
  4. Random assignment and analysis of comparable program and comparison groups, including controls for attrition.

Notes on Methodology

---

---

---

### Sample Level Code Sheet

Because a study may report results separately for distinct samples, a sample is a separate level in the coding scheme. For example, if a study reports the results separately for persons with prior arrests and persons without prior arrests, then each would have a separate sample level code sheet. Note that the treatment–comparison contrast is the same for the different samples. It is critical that each sample be independent of all other coded samples. Most studies will have only a single sample, that is, they will report their results for the full sample of individuals included in the study. A few may report the results broken-out by multiple variables, such as gender, ethnicity, and prior criminal history. Only one of these breakouts can be used in this coding scheme. Choose either the breakout with the most complete information or the breakout that is most relevant to our purposes (e.g., gender or age).

#### Identifying Information

- 37. Study (document) identifier studyid \_\_\_ \_\_\_ \_\_\_
- 38. Module identifier modid \_\_\_
- 39. Sample identifier (each sample within a study gets a unique number; most modules will have only a single sample) sampid \_\_\_ \_\_\_
- 40. Coder’s initials codersmp \_\_\_ \_\_\_ \_\_\_

#### Sample Description

- 41. Sample description treatment group (location, level of security, prior history)  
 \_\_\_\_\_  
 \_\_\_\_\_ sampdes1
- 42. Sample description comparison group (location, level of security, prior history)  
 \_\_\_\_\_  
 \_\_\_\_\_ sampdes2
- 43. Total *N* in treatment group at beginning of study txn \_\_\_ \_\_\_ \_\_\_ \_\_\_
- 44. Total *N* in comparison group at beginning of study cgn \_\_\_ \_\_\_ \_\_\_ \_\_\_

Note: The two items above must equal the total sample size prior to any attrition. If multiple samples per module are being coded, the sum across samples must equal the total sample size prior to any attrition.

45. Approximate age range of study participants age \_\_\_  
 1. Adolescent (12 to 18)  
 2. Young adult (19 to 25)  
 3. Adult (18+)  
 4. Adolescent and young adult  
 5. Adolescent and adult  
 9. Unspecified or cannot tell
46. Youngest age included in sample (99 if unknown) yngage \_\_\_ \_\_\_
47. Oldest age included in sample (99 if unknown) oldage \_\_\_ \_\_\_
48. Exact proportion of males in sample if reported males \_\_\_ . \_\_\_ \_\_\_ \_\_\_
49. Approximate gender description of sample sex \_\_\_  
 1. All males (>90%)  
 2. More males than females (60% to 90% males)  
 3. Roughly half males and half females  
 4. More females than males (60% to 90% females)  
 5. All females (>90%)  
 9. Cannot tell
50. Race/ethnicity of sample (999 if unknown)  
 Percentage African American raceaa \_\_\_ \_\_\_ \_\_\_  
 Percentage Hispanic racehisp \_\_\_ \_\_\_ \_\_\_  
 Percentage Asian raceasia \_\_\_ \_\_\_ \_\_\_  
 Percentage Native American racena \_\_\_ \_\_\_ \_\_\_  
 Percentage Caucasian racecauc \_\_\_ \_\_\_ \_\_\_
51. Offender type general categories samptype \_\_\_  
 1. Violent, person crimes  
 2. Nonviolent, nonperson crimes  
 3. Mixed

## Dependent Variable Code Sheet

Code each dependent variable reported in the study separately. The same dependent variable measured at multiple time points should be coded only once. That is, whether or not a participant was arrested by 6 months, 12 months, and 18 months is treated as a single dependent measure, arrest. For noncriminal behavior dependent variables, code only the outcome descriptor (item 56) and items 59 and 60. Do not code effect sizes for noncriminal behavior dependent variables.

### Identifying Information

53. Study (document) identifier studyid \_\_\_ \_\_\_ \_\_\_
54. Outcome identifier (each outcome within a study gets a unique number)  
outid \_\_\_ \_\_\_
55. Coder's initials coderdv \_\_\_ \_\_\_ \_\_\_

### Outcome Information

56. Outcome label (label used in report)
- 
- label
57. Recidivism construct represented by this measure. (1=yes; 0=no)
- |   |         |
|---|---------|
| arrest                                  | dv1 ___ |
| conviction                              | dv2 ___ |
| reinstitutionalization/reincarceration  | dv3 ___ |
| revocation                              | dv4 ___ |
| technical supervision violation         | dv5 ___ |
| other indicator of criminal involvement | dv6 ___ |
58. Specify types of offenses included in recidivism measure. (1=yes; 0=no)
- |   |             |
|---|-------------|
| all offenses ("no" for items b-h)       | dvtype1 ___ |
| drug offenses                           | dvtype2 ___ |
| person offenses, sexual                 | dvtype3 ___ |
| person offenses, nonsexual              | dvtype4 ___ |
| person offenses, unspecified            | dvtype5 ___ |
| property offenses                       | dvtype6 ___ |
| technical supervision or status offense | dvtype7 ___ |
| other (specify) _____                   | dvtype8 ___ |
59. Type of measurement scale scale \_\_\_
1. Dichotomy
  2. Trichotomy
  3. 4–9 discrete ordinal categories
  4. >9 discrete ordinal categories or continuous

60. Source of data source \_\_\_\_
1. Self-report
  2. Other report (e.g., teacher, parent)
  3. Official record (e.g., school, police, probation, court, institution)
  4. Other \_\_\_\_\_
  9. Cannot tell

### Effect Size Level Code Sheet

Code a separate effect size for each contrast between the treatment and comparison groups for each dependent variable.

#### Identifying Information

- 61. Study identifier studyid \_\_\_ \_\_\_ \_\_\_
- 62. Module identifier modid \_\_\_
- 63. Sample identifier sampid \_\_\_
- 64. Outcome identifier outid \_\_\_ \_\_\_
- 65. Effect size identifier (number each effect size within a study sequentially) esid \_\_\_ \_\_\_
- 66. Coder's initials coderes \_\_\_ \_\_\_ \_\_\_

#### Effect Size Information

- 67. Effect size type es\_type \_\_\_
  - 1. Baseline (pretest; prior to start of intervention)
  - 2. Post-test (first measurement point, post intervention)
  - 3. Follow-up (all subsequent measurement points, post intervention)
- 68. Which group does the raw effect favor (ignoring statistical significance)? es\_direc \_\_\_
  - 1. Treatment group
  - 2. Comparison group
  - 3. Neither (ES equals zero)
  - 9. Cannot tell (ES cannot be used if this option is selected)
- 69. Is this difference reported as statistically significant by the investigator? es\_sig \_\_\_
  - 0. No
  - 1. Yes
  - 8. Not tested
  - 9. Cannot tell

70. If tested, what type of statistical test was used? es\_test \_\_\_  
 1 *t*-test  
 2 *F*-test  
 3  $\chi^2$   
 4 Regression analysis, including logistic regression, etc.  
 5 Other \_\_\_\_\_  
 8 na  
 9 Cannot Tell

71. Time frame in months captured by measure (weeks/4.3)  
 Minimum es\_time1 \_\_\_ \_\_\_ \_\_\_  
 Maximum es\_time2 \_\_\_ \_\_\_ \_\_\_  
 Mean es\_time3 \_\_\_ \_\_\_ \_\_\_  
 Fixed (same for all subjects) es\_time4 \_\_\_ \_\_\_ \_\_\_

72. Time in months from end of program to measurement point (weeks/4.3)  
 Minimum es\_futi1 \_\_\_ \_\_\_ \_\_\_  
 Maximum es\_futi2 \_\_\_ \_\_\_ \_\_\_  
 Mean es\_futi3 \_\_\_ \_\_\_ \_\_\_  
 Fixed (same for all subjects) es\_futi4 \_\_\_ \_\_\_ \_\_\_

**Effect Size Data—All Effects**

73. Treatment group sample size for this effect size es\_txn \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_  
 74. Comparison group sample size for this effect size es\_cgn \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_

**Effect Size Data—Continuous Outcomes**

75. Treatment group mean es\_txm \_\_\_ \_\_\_ \_\_\_ . \_\_\_ \_\_\_ \_\_\_  
 76. Comparison group mean es\_cgm \_\_\_ \_\_\_ \_\_\_ . \_\_\_ \_\_\_ \_\_\_  
 77. Are the above means adjusted? (1=yes; 0=no) es\_madj \_\_\_  
 78. Treatment group standard deviation es\_txsd \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_  
 79. Comparison group standard deviation es\_cgsd \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_  
 80. Treatment group standard error es\_txse \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_  
 81. Comparison group standard error es\_cgse \_\_\_ \_\_\_ \_\_\_ \_\_\_ \_\_\_  
 82. *t*-value from an independent *t*-test or square root of *F*-value from a one-way analysis of variance with one *df* in the numerator (only two groups)  
 es.t \_\_\_ \_\_\_ . \_\_\_ \_\_\_ \_\_\_

83. Exact probability for a  $t$ -value from an independent  $t$ -test or  $F$ -value from a one-way analysis of variance with one  $df$  in the numerator

es\_t\_p \_\_\_ . \_\_\_ \_\_\_

84. Correlation coefficient (point biserial)

es\_rpb \_\_\_ . \_\_\_ \_\_\_

**Effect Size Data—Dichotomous Outcomes**

85. Treatment group; number successful

es\_txns \_\_\_ \_\_\_ \_\_\_

86. Comparison group; number successful

es\_cgns \_\_\_ \_\_\_ \_\_\_

87. Treatment group; proportion successful

es\_txps \_\_\_ . \_\_\_ \_\_\_

88. Comparison group; proportion successful

es\_cgps \_\_\_ . \_\_\_ \_\_\_

89. Are the above proportions adjusted for pretest variables? (1=yes; 0=no)

es\_padj \_\_\_

90. Logged odds-ratio

es\_lgodd \_\_\_ \_\_\_ . \_\_\_ \_\_\_

91. Standard error of logged odds-ratio

es\_selgo \_\_\_ \_\_\_ . \_\_\_ \_\_\_

92. Logged odds-ratio adjusted? (e.g., from a logistic regression analysis)

es\_oadj \_\_\_

93.  $\chi^2$  value with  $df = 1$  (2 by 2 contingency table)

es\_chisq \_\_\_ \_\_\_ . \_\_\_ \_\_\_

94. Correlation coefficient (phi)

es\_rphi \_\_\_ \_\_\_ . \_\_\_ \_\_\_

**Effect Size Data—Hand Calculated**

95. Hand calculated  $d$ -type effect size

es\_hand \_\_\_ . \_\_\_ \_\_\_

96. Hand calculated standard error of the  $d$ -type effect size

es\_sehd \_\_\_ . \_\_\_ \_\_\_

D Sample Screen from the Coding Protocol Database

FileMaker Pro - [es. (p3)]  
 File Edit Mode Select Format Script Window Help

**Effect Size Level Database**

biblio study	module outcome	sample effect size
--------------	----------------	--------------------

**identifying information**

Gordon, A. & Nicolaichuk, T. (1996). Applying the risk principle of sex offender treatment. Unpublished manuscript, Washington State Department of Corrections, Olympia, Washington.

27 Study identifier Outcome identifier Sexual reconivition  
 001 Module identifier Effect size identifier (number each ES within a study)  
 001 Sample identifier DBW Coder's initials

**effect size information**

Effect size type

1 baseline (pretest, prior to start of intervention)  
 2 post-test (first measurement point, post intervention)  
 3 follow-up (all subsequent measurement points, post intervention)

Which group does the raw effect (difference) favor (ignoring statistical significance)?

1 treatment group  
 2 comparison group  
 3 neither (ES equals zero)  
 9 cannot tell (ES cannot be used if this option is selected)

Is this difference reported as statistically significant by the investigator?

0 no  
 1 yes  
 8 not tested  
 9 cannot tell

Time frame in months captured by measure (weeks/4.3)

999 minimum  
 168 maximum  
 62 average  
 999 fixed (same for all subjects)

Notes

---

**descriptive data -- means and SD/SE**

	N	Mean	SD	SE
treatment group	80			
control group	116			

Are the means adjusted?  0  1

**descriptive data -- frequencies or percents**

	N	# successful	proportion successful
treatment group	80		.940
control group	116		.854

Are the proportions adjusted?  0  1

**inferential data**

t-test   
 p-value (t-test)   
 chi-square

**correlations**

correlation -- point-biserial   
 correlation -- phi

**effect size**

use this effect size in main analyses  
 computer calculated  0.501  
 hand calculated

100 | Browse

**E Sample Output from Meta-Analytic Macros<sup>5</sup>**

No. of obs =	26	Homogeneity Analysis				
Minimum obs =	-.64565	Q =	178.44			
Maximum obs =	1.763	df =	25			
Weighted SD =	0.414	p =	0.00000			
-----						
	Mean	-95%CI	+95%CI	SE	Z	P
-----						
Fixed effect	0.33552	0.27485	0.39620	0.03096	10.83902	0.00000
Random effect	0.47822	0.29999	0.65645	0.09093	5.25892	0.00000
-----						
Random effects variance component = 0.16033						
estimated via noniterative method of moments						

Figure 1: The Mean ES macro computes the overall inverse variance weighted effect size and associated statistics, such as the homogeneity statistic and confidence intervals.

<sup>5</sup>Available for SAS, SPSS, and Stata from: <http://mason.gmu.edu/~dwilsondb/ma.html>

```

Meta-Analytic Analog to the One-way ANOVA, Mixed Effects Model
-----
Source |           Q           df           P
-----
Between |          2.4712           3    0.48051
Within  |          31.2463          21    0.06965
-----
Total   |          33.7176          24    0.08981

Descriptive Fixed Effects Meta-Analytic Results by: txttype
-----
txttype |           Mean   St. Er.   [95% Conf. Int.]           z   P>|z|           n
-----
1       |          .23457   .21975   -.19612   .66527   1.0675   0.28576           4
2       |          .47803   .11987   .24309   .71296   3.9880   0.00007          13
3       |          .61208   .22794   .16533   1.0588   2.6853   0.00725           5
5       |          .20347   .23986   -.26665   .67360   .84828   0.39628           3
-----
Total   |          .42104   .08876   .24708   .59500   4.7437   0.00000          25

Fixed Effects Homogeneity Analysis by: txttype
-----
Source |           Q           df           P
-----
1       |          3.6537           4    0.64487
2       |          32.6494          13    0.00085
3       |          17.8359           5    0.00128
5       |           0.5530           3    0.96206
-----
Random effects variance component (via method of moments) = .1432373
    
```

Figure 2: The Meta F macro performs a categorical analysis of effect sizes, computing a separate inverse variance weighted mean and associated statistics for each level of a categorical variable. Both fixed and random effects models can be tested.

```

Meta-Analytic Random Intercept, Fixed Slopes Regression Analysis
-----
Source |           Q           df           P           No. of obs =    26
-----
Model  |          3.9464           2    0.13901           Mean ES = 0.4775
Residual |          36.0245          23    0.04102           R-squared = 0.0987
-----
Total  |          39.9709          25    0.02937

Variable |           Coef.   Std. Err.           z           P>|z|   [95% Conf. Interval]
-----
dvl     |          .111636   .194323   .574488   0.565638   -.269237   .492509
es_time3 |          .005718   .002976   1.92104   0.054727   -.000116   .011551
_cons   |          .084193   .233668   .360310   0.718615   -.373797   .542183
-----
Random effects variance component (via method of moments) = .1577332
    
```

Figure 3: The Meta Reg macro performs meta-analytic regression. Both fixed and random effects models can be tested.