

## Randomized experiments in criminology: What have we learned in the last two decades?

DAVID P. FARRINGTON\*

*Institute of Criminology, Cambridge University, Cambridge, UK*

*\*corresponding author: E-mail: dpf1@cam.ac.uk*

BRANDON C. WELSH

*Department of Criminal Justice, University of Massachusetts-Lowell, Massachusetts, USA*

**Abstract.** This paper aims to review randomized experiments in criminology with offending outcomes and reasonably large numbers that were published between 1982 and 2004. A total of 83 experiments are summarized, compared with only 35 published between 1957 and 1981: 12 on policing, 13 on prevention, 14 on corrections, 22 on courts, and 22 on community interventions. Randomized experiments are still relatively uncommon, but there have been more large-scale multi-site experiments and replication programs. There have also been several experiments in which 100 or more places were randomly assigned. Relatively few experiments (only 10 out of 83) were conducted outside the United States. Meta-analyses suggest that prevention methods, correctional therapy, batterer programs, drug courts, juvenile restitution and deterrent policing were effective in reducing offending, while Scared Straight and boot camp programs caused a significant increase in offending.

### Background

The defining feature of an experiment is that it investigates the impact of prospectively planned variations in an independent variable (e.g., arresting or not arresting spouse assaulters) on a dependent variable (e.g., reoffending). One of the major methodological problems facing experimental (and nonexperimental) research in criminology is to maximize internal validity: that is, to attribute observed variations in the dependent variable unambiguously to the effects of variations in the independent variable (e.g., an intervention). The main threats to internal validity are well known (see e.g., Farrington 2003c):

1. Selection: the effect reflects pre-existing differences between experimental and control conditions.
2. History: the effect is caused by some event occurring at the same time as the intervention.
3. Maturation: the effect reflects a continuation of pre-existing trends.
4. Instrumentation: the effect is caused by a change in the method of measuring the outcome.
5. Testing: the pretest measurement causes a change in the posttest measure.

6. Regression to the mean: where an intervention is implemented on units with unusually high scores (e.g., areas with high crime rates), natural fluctuation will cause a decrease in these scores on the posttest, which may be mistakenly interpreted as an effect of the intervention.
7. Differential attrition: the effect is caused by differential loss of units (e.g., people) from experimental compared to control conditions.
8. Causal order: it is unclear whether the intervention preceded the outcome.

In principle, a randomized experiment, in which units are assigned at random to experimental or control conditions, has the highest possible internal validity, because it can rule out all these threats (or alternative explanations of observed findings). Providing that enough units are assigned, randomization equates units in one condition with units in another condition on all possible (measured and unmeasured) extraneous variables that might influence the dependent variable. The control of extraneous variables by randomization in the behavioral and social sciences is similar to the control of extraneous variables in experiments in the physical sciences by holding boundary conditions (e.g., temperature or pressure) constant. In both cases, the aim is to disentangle the influence of the independent variable from all other possible influences on the dependent variable.

In light of the methodological advantages of randomized experiments, it is surprising that relatively few have been carried out in criminology. This is because of the many ethical, legal, and practical difficulties that face researchers who wish to mount a randomized experiment. This paper aims to provide a useful resource for researchers who wish to know about randomized experiments that have been carried out in criminology. We hope that evidence about past experiments that have been conducted successfully will help researchers to mount randomized experiments in the future.

Farrington (1983) reviewed randomized experiments on crime and justice that had been published in English between 1957 and 1981. Only experiments carried out in real-life settings were included, so laboratory experiments and simulation exercises were excluded. Methodological experiments, for example comparing different methods of administering a self-report questionnaire (Hindelang et al. 1981), were excluded. Within-subjects designs, where all treatments were given to the same people in a random order, were also excluded. Finally, experiments were only included if there were at least 50 persons assigned to each condition, since it was considered that with smaller numbers, the main benefit of randomization, in ensuring equivalence of persons on all measured and unmeasured extraneous variables, was not likely to be achieved.

Farrington (1983) presented summary tables showing key features of 22 community intervention experiments on prevention, juvenile diversion, probation, parole and community treatment and of 20 experiments on police, courts, and correctional treatment. The relatively small number of criminological experiments published up to that time is noteworthy. Most were carried out in the United States (31) or the United Kingdom (8), with two in Canada and one in Denmark. Only 35

experiments had offending outcomes and 25 of them found no significant effect of the intervention.

The main aim of the present paper is to review randomized experiments in criminology published since 1982 which have offending outcomes. The earlier review essentially focused on whether studies found significant effects of the intervention. Nowadays, however, there is much more emphasis on the effect size, and on systematic reviews and meta-analysis (Farrington and Welsh 2001; Lipsey and Wilson 2001), and so this paper includes meta-analyses of effect sizes.

### **Methodology**

The inclusion criteria for the present paper were essentially the same as in the 1983 review, with the additional requirement that the outcomes should include some measure of offending. Experiments were included in this paper if:

1. Units (persons or places) were randomly assigned to conditions. We did include experiments where persons were assigned alternately to conditions (Britt et al. 1992) but not studies where persons were assigned to one condition during one time period and to another condition during another time period (McDonald et al. 1992).
2. At least 50 units were initially assigned to each condition, or at least 100 units were initially assigned to two experimental conditions. Unfortunately, there was often a considerable drop-out between persons initially assigned and persons finally assessed (e.g., in follow-up interviews), and differential drop-out was one of the most important threats to internal validity. We report results for initially assigned units ("intent-to-treat") rather than for those who completed or actually received the treatment ("treatment-of-treated").
3. There was an outcome measure of offending (as opposed to self-reported drug use, childhood antisocial behavior or inmate misconduct, for example). Criminological experiments focusing on other outcomes, such as court dispositions (Greenwood and Turner 1993b), were excluded. There were often many outcome measures and in our tables we have reported only one summary measure (because of space limitations). We chose to report rearrests or reconvictions rather than self-reported offending because of the differential attrition problem with self-reports (e.g., Greenwood and Turner 1993a). We chose to report rearrests or reconvictions rather than revocations or return to prison because the latter measures are more dependent on agency decision-makers. We chose measures of crime in general rather than specific types of offending such as violent or property offences. Where information was given for several follow-up periods, we chose the longest one.
4. The experiment was published in English. There is a problem in defining what is meant by "published". We did not limit our review to books and journals but also included agency reports (e.g., California Youth Authority 1997).

Where an experiment was conducted on several sites, we report results for each site where possible (i.e., if the results were presented separately for each site and if the initial sample size at a site was at least 100).

### Searching for experiments

We did not have the resources to carry out an exhaustive systematic review of randomized experiments in criminology. Searching all the major abstracting systems would produce many thousands of references that would need to be obtained and read for possible inclusion. However, we searched key books (e.g., Sherman et al. 2002; Welsh and Farrington 2005), key systematic reviews (e.g., Braga 2001; MacKenzie et al. 2001; Tong and Farrington 2005), and key journals (e.g., *Criminal Justice and Behavior*; *Justice Quarterly*), selected online databases such as NCJRS Abstracts and SPECTR (Petrosino et al. 2000) and we consulted leading researchers.

Petrosino et al. (2003a) reported that there were 267 distinct criminological experiments published or available in English between 1945 and 1993. The annual average number of experiments increased sharply from 1.8 in 1961–1965 to 9.4 in 1971–1975 but then levelled off to 9.4 in 1986–1990 and 11.6 in 1991–1993. This research suggests that, since 1983, there have been about 10 experiments per year reported in English but some of these would not meet our inclusion criteria (e.g., because of small numbers or articles not published).

Weisburd et al. (2001) analyzed the Maryland report on the effectiveness of criminological interventions (Sherman et al. 1997). They excluded school-based interventions, few of which had offending outcomes. They found 308 studies with offending outcomes, of which 46 scored 5 on the Maryland Scientific Methods Scale (Farrington et al. 2002) because they were randomized controlled trials. However, there were also some randomized experiments among the 28 evaluations scored 4 (“quasi-experiments”), which had been downgraded from 5 to 4 because of differential drop-out, for example.

Of the 46 studies coded as randomized experiments, 22 reported a nonsignificant effect, 17 found a desirable effect, and seven found an undesirable effect of the intervention. Some of these 46 experiments would not meet our inclusion criteria because they were carried out before 1983 or were based on small numbers. Nevertheless, the two reviews by Petrosino et al. (2003b) and Weisburd et al. (2001) suggest that the number of criminological experiments conducted since 1983 that would meet our inclusion criteria is probably between 40 and 150.

Despite this surprisingly large number of randomized experiments, it is still true that only a small fraction of criminological evaluations use the experimental method. For example, only five out of 44 evaluations of boot camps (MacKenzie et al. 2001), only three out of 32 evaluations of prison-based drug treatment (Mitchell et al. 2005) and only four out of 42 evaluations of drug courts (Wilson et al. 2005) used randomized experiments.

### Summarizing results

We classified experiments into five categories:

1. Policing
2. Prevention
3. Corrections
4. Court
5. Community

This is to some extent an arbitrary classification system and some experiments could perhaps be placed in more than one category (e.g., the experiment by Killias et al. (2000) comparing community and institutional sentences, which we placed in the “community” category). Nevertheless, we thought that it was useful to organize the experiments according to their main focus. Because of space limitations, it is only possible to provide brief details of each experiment in this paper. For more details, the references should be consulted.

The tables summarize:

1. Author, publication date, location of research.
2. Description and initial size of sample.
3. Experimental and control conditions: In order to simplify the presentation, we have sometimes combined or eliminated conditions in order to reveal the key contrast.
4. Results: We report the percentage who became offenders in the different conditions, where this was given. In other cases, we report offending rates. We do not report the results of multivariate analyses.
5. Effect size: For ease of understanding, we report the percentage offending in the experimental condition compared with the control condition. This is the relative difference rather than the absolute difference in percentages. For example, if 50% of the control group and 20% of the experimental group reoffended, the percentage decrease would be 60%  $[(50-20)/50]$ . Positive values of the percentage difference indicate that there was a higher offending rate in the experimental condition compared with the control condition. We also report the standardized mean difference  $d$ , which is used as the measure of effect size in later meta-analyses. This is related to the odds ratio (OR) as follows (see Lipsey and Wilson 2001, p. 202):

$$d = 0.55 \ln (\text{OR})$$

where  $\ln (\text{OR})$  is the natural logarithm of OR.

### Policing experiments

Table 1 summarizes 12 policing experiments. We begin with these because of the great historical importance of the Minneapolis Domestic Violence Experiment

Table 1. Policing experiments (12).

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Sherman and Berk (1984), Minneapolis	330 domestic violence suspects	E = arrest C = advise or separate	6 months rearrests: E 6.5% (93) C 14.8% (237)	0.52 (−56)*
Sherman et al. (1992), Milwaukee	1,200 domestic violence suspects	E = arrest C = warning	7–9 months rearrests: E 20.5% (624) C 23.2% (297)	0.09 (−12)
Hirschel et al. (1992), Charlotte	686 domestic violence suspects	E = arrest C = citation, advise or separate	6 months rearrests: E 18.2% (214) C 15.6% (436)	−0.10 (+17)
Berk et al. (1992), Colorado Springs	1,658 domestic violence suspects	E = arrest C = protection order, counselling, restore order	6 months rearrests: E 19.2% (421) C 19.3% (1,158)	0.00 (−1)
Pate and Hamilton (1992), Miami	907 domestic violence suspects	E = arrest C = no arrest	6 months rearrests: E 19.1% (465) C 20.6% (442)	0.05 (−7)
Dunford (1990), Omaha	247 domestic violence suspects (offender absent)	E = arrest C = no arrest	12 months rearrests: E 10.8% (111) C 20.6% (136)	0.42 (−48)*
Dunford et al. (1990), Omaha	330 domestic violence suspects (offender present)	E = arrest C = mediate or separate	6 months rearrests: E 11.9% (109) C 10.0% (221)	−0.11 (+20)
Sherman and Weisburd (1995), Minneapolis	110 hot spots	E = increased patrol C = normal patrol	12 months crime calls: EB 6531 (55) EA 6931 (55) CM 6491 (55) CA 7702 (55)	0.06 (−11)*
Sherman and Rogan (1995), Kansas City	207 hot spots	E = crack house raid C = no raid	30 days calls for service: EB 1059 (104) EA 865 (104) CB 1037 (103) CA 929 (103)	0.05 (−9)
Mazerolle et al. (1998), Oakland (CA)	100 high-risk places	E = civil remedies C = police response	Observed drug-selling: EB 6% (50) EA 4% (50) CB 10% (50) CA 44% (50)	1.31 (−85)*

Table 1. Continued.

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size <math>d</math> (%)</i>
Abrahamse et al. (1991), Phoenix	480 repeat offenders	E = police targeting C = no targeting	6 months convictions: E 82.6% (270) C 75.4% (253)	0.24 (–10) <sup>+</sup>
Eck and Wartell (1998), San Diego	121 rental properties	E = police letter or meeting C = no police action	30 months crimes: EM 3.2 (79) ES 4.1 (79) CM 5.1 (42) CS 6.3 (42)	0.38 (–37)*

Effect size shows standardized mean difference ( $d$ ), with relative percentage difference between experimental and control conditions in parentheses.

\* $p < 0.05$ ,

<sup>+</sup> $p < 0.10$ .

E: Experimental, C: Control, EB: Experimental Before, CB: Control Before, EA: Experimental After, CA: Control After, EM: Experimental Mean, CM: Control Mean, ES: Experimental SD, CS: Control SD.

(Sherman and Berk 1984). This showed that arresting the perpetrators caused fewer repeat incidents of spouse assault than ordering the offender to leave the premises for 8 hours or offering advice and mediation. These results were welcomed by the US Department of Justice and used to encourage police forces to arrest male perpetrators of domestic violence rather than dealing with them in other ways (Meeker and Binder 1990; Sherman and Cohn 1989). According to Sherman (1992, p. 103), “the publicity helped to gain acceptance for the idea of randomized experiments.”

The results of the original Minneapolis experiment encouraged the National Institute of Justice to fund an ambitious program of research to see if the findings could be replicated in different settings. The results from published evaluations in Milwaukee (Sherman et al. 1992), Charlotte (Hirschel et al. 1992), Colorado Springs (Berk et al. 1992), Miami (Pate and Hamilton 1992) and Omaha (offender absent: Dunford 1990; offender present: Dunford et al. 1990) are shown in Table 1. In some cases, the percentage of offenders in experimental and control conditions who were rearrested were taken from Garner et al. (1995), because they were not published in the original reports. (For a more recent research synthesis of these experiments, see Maxwell et al. 2002.)

Table 1 shows that the results of these seven experiments on the effects of arrest on spouse assault were quite variable. Only the original Minneapolis experiment and the offender-absent experiment in Omaha found that arrest had significantly desirable effects. (The significance levels are based on our calculated  $d$  values.) Rearrests decreased by 56% in Minneapolis and by 48% in Omaha. In attempting to reconcile all these results, Sherman (1992) concluded that offenders with a lower stake in conformity (e.g., unemployed, unmarried, African-American) tended to get worse after arrest, whereas those with a greater stake in conformity tended to improve (i.e., offend less).

The next two experiments in Table 1 investigated the effects of deterrent police activity in “hot spots”, that is, high crime locations (see Braga 2001). These experiments are noteworthy because the unit that was randomized was an area, not an individual. Sherman and Weisburd (1995) found that increased police patrolling in hot spots caused a significant (11%) decrease in crime calls for service, while Sherman and Rogan (1995) reported that police raids on crack houses caused some decrease (9%) in calls for service. Several other randomized experiments based on places were excluded from this review because the number of units randomized was less than 100 (e.g., Braga et al. 1999; Weisburd and Green 1995).

In a third randomized experiment targeted on places, Mazerolle et al. (1998) compared the impact in controlling social disorder of civil remedies (police working with city agency representatives to inspect drug nuisance properties, coerce landlords to clean up blighted properties, post “no trespassing” signs, enforce civil law codes and municipal regulatory rules, and initiate court proceedings against property owners who failed to comply with civil law citations) *versus* traditional police tactics (surveillance, arrests and field interrogations). Observations of street blocks showed that conditions improved in the experimental places compared with the control places. In the most direct measure of offending, the mean number of males selling drugs on experimental blocks decreased from 3 before to 2 afterwards, while the mean number selling drugs on control blocks increased from 5 before to 22 afterwards. These figures are treated as prevalence measures in order to calculate a (very large)  $d$  value in Table 1.

Abrahamse et al. (1991) investigated the impact of a Repeat Offender Program that gave special attention to certain offenders to try to increase their probability of being convicted. The results were desirable, since the experimental offenders were 10% more likely to be convicted. This is the only example in Table 1 where a higher conviction rate is a desirable outcome and this is indicated by a positive  $d$  value and a negative sign before the 10%. Finally, Eck and Wartell (1998) randomly assigned 121 rental properties to receive either a letter or a visit from the police or no treatment. The police actions were intended to reduce drug dealing and there were indeed 37% fewer crimes reported in the experimental places during the 30-month follow-up period.

### Prevention experiments

Farrington and Welsh (2003) carried out a meta-analysis of family-based prevention experiments. Table 2 summarizes 13 prevention experiments that fulfilled our inclusion criteria (e.g., for initial sample size and offending outcomes). They are roughly ordered according to the age of the children, from the youngest upwards. It can be seen that five of the 13 experiments found that the intervention had significantly desirable effects in reducing later offending.

Olds et al. (1998) investigated the effects of a home visiting program on pregnant women. The home visitors (nurses) gave the women advice about child-rearing, infant development, nutrition and the need to avoid alcohol and drugs.



Table 2. Prevention experiments (13).

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Olds et al. (1998), Elmira (NY)	400 pregnant women	E = home visits C = no home visits	15 years arrests: E 16.6% (176) C 36.0% (148)	0.47 (−54)*
Schweinhart et al. (2005), Ypsilanti (MI)	123 children age 3–4	E = preschool C = no preschool	Felony arrests up to age 40: E 31% (58) C 48% (65)	0.39 (−35)
Campbell et al. (2002), USA	111 children age 3	E = intensive preschool C = usual preschool	Felony conviction up to age 21: E 8% (53) C 12% (51)	0.27 (−33)
Mills et al. (2002), Washington	206 children age 5	E = cognitive preschool C = usual preschool	SR offending up to age 15: EM 12.62 (90) ES 14.27 (90) CM 11.05 (81) CS 13.66 (81)	−0.11 (+14)
Tremblay et al. (1996), Montreal	319 boys age 7	E = child skills training plus parent training C = no treatment or just attention	SR arrest up to age 15: E 14.0% (43) C 30.1% (123)	0.54 (−53)*
Harrell et al. (1999), USA	671 adolescents age 11–13	E = risk focused prevention C = no treatment	12 months recorded crime: E 28.0% (264) C 34.3% (236)	0.16 (−18)
Borduin et al. (1995), Columbia (MO)	176 delinquents age 12–17	E = MST C = individual therapy	4 years arrests: E 26.1% (92) C 71.4% (84)	1.08 (−63)*
Henggeler et al. (1997), South Carolina	155 delinquents age 10–17	E = MST C = probation	1.7 years arrest rate: EM 0.89 (70) ES 1.39 (70) CM 1.20 (70) CS 3.11 (70)	0.12 (−26)
Henggeler et al. (1999), Charleston (SC)	116 psychiatric adolescents, mean age 13	E = MST C = hospitalization	4 months arrest rate: EB 0.46 (57) EA 0.33 (57) CB 0.30 (56) CA 0.27 (56)	0.21 (−20)
Henggeler et al. (2002), Charleston (SC)	118 delinquents age 12–17	E = MST C = usual community services	4 years conviction rate: EM 0.34 (43) ES 0.43 (43) CM 0.77 (37) CS 1.34 (37)	0.39 (−56)*

Table 2. Continued.

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Leschied and Cunningham (2002), London (Ontario)	409 delinquents age 12–17	E = MST C = probation	12 months criminal convictions: E 41.2% (211) C 37.6% (198)	–0.09 (+10)
Grossman and Tierney (1998), USA	1,138 youths age 10–16	E = Big Brothers Big Sisters C = no treatment	12 months SR violence: E 41% (487) C 41% (472)	0 (0)
Schochet et al. (2001), USA	15,386 youths age 16–24	E = Job Corps C = no Job Corps	40 months arrests: E 28.8% (6,828) C 32.6% (4,485)	0.10 (–12)*

Effect size shows standardized mean difference (*d*), with relative percentage difference between experimental and control conditions in parentheses.

\* $p < 0.05$ .

E: Experimental, C: Control, SR: Self-report, MST: Multisystemic therapy, EB: Experimental Before, CB: Control Before, EA: Experimental After, CA: Control After, EM: Experimental Mean, CM: Control Mean, ES: Experimental SD, CS: Control SD.

Hence, this was a general parent education program. A 15-year follow-up showed that the children of visited mothers were arrested at a significantly (54%) lower rate than the children of nonvisited mothers.

Schweinhart et al. (2005) carried out the longest follow-up of the effects of an intervention. In the famous Perry project, experimental children attended a cognitively-oriented preschool program that was designed to increase their thinking and reasoning abilities and school achievement. The experimental and control children were followed up to age 40, and the results showed that the experimental children had 35% fewer arrests. Similarly, in the Abercedarian project, Campbell et al. (2002) found that an intensive cognitively-oriented preschool curriculum led to 33% fewer convictions up to age 21 (compared with a regular preschool program). However, encouraging results were not obtained in the preschool experiment of Mills et al. (2002).

Tremblay et al. (1996) evaluated the success of a multi-modal program including child skills training and parent management training. The program proved to be effective in reducing later offending (by 53%), and in fact the desirable effects increased over time. Another multi-modal program, termed “Children at Risk”, was evaluated by Harrell et al. (1999) in five sites across the United States. The intervention was designed to reduce the number of risk factors to which adolescents were exposed, through family services, skills training, mentoring, education and after school activities. The program was effective in reducing offending (by 18%) and the researchers concluded that the main effects were through reducing peer risk factors.

The next five experiments in Table 2 evaluated multisystemic therapy (MST), which is a multimodal intervention designed for serious juvenile offenders. The particular type of treatment is chosen according to the needs of the young person and it may include individual, family, peer, school and community interventions (including parent training and skills training). Four of the five trials of MST, all carried out by Henggeler (the originator of this treatment) and his colleagues, found that the intervention was effective in reducing later offending (Borduin et al. 1995; Henggeler et al. 1997, 1999, 2002). The results obtained by Borduin et al. (1995) showed a remarkably large effect size ( $d$ ) of 1.08, and a 63% reduction in the prevalence of arrests. However, the one large-scale independent evaluation of MST, by Leschied and Cunningham (2002) in Canada, did not find that it was effective in reducing later convictions (compared with the usual community services, which typically involved probation supervision); the MST group were 10% more likely to be convicted.

The final two experiments evaluated large-scale community-based interventions in multi-site studies. Grossman and Tierney (1998) found that the mentoring program “Big Brothers Big Sisters” had some beneficial effects, for example in decreasing the onset of illegal drugs and alcohol use. However, it had no effect according to the only prevalence measure (of self-reported violence). Schochet et al. (2001) studied the impact of “Job Corps” (a program including basic education, vocational training, support services and job placement assistance) and found that it caused statistically significant reductions in the arrest rate.

### **Correctional experiments**

Historically, many of the pioneering experiments in criminology were concerned with the effectiveness of correctional interventions. Table 3 summarizes 14 experiments carried out in correctional facilities. The first two, by Lewis (1983) and Cook and Spirrisson (1992), evaluated the “Scared Straight” program, in which adult prisoners harangue young delinquents about the terrors of imprisonment in an attempt to deter them from offending. Both experiments had negative effect sizes; the percentage convicted in Lewis (1983) was 21% greater in the Scared Straight condition, while the decrease in offending rates (after *versus* before) in Cook and Spirrisson (1992) was 7% less in the Scared Straight condition. These results are quite typical. A systematic review by Petrosino et al. (2003b) found that reoffending was greater after Scared Straight in every one of seven evaluations. They concluded (p. 58) that “doing nothing would have been better than exposing juveniles to the program.”

The next four experiments, three by Peters et al. (1997) and one by the California Youth Authority (1997), evaluated the effectiveness of juvenile boot camps, which were also intended to deter juveniles from offending. None of these four experiments found a significant desirable effect of a boot camp and one (Peters et al. 1997, Cleveland) found that it led to a significantly (44%) increased prevalence of offending compared with confinement in youth services facilities.

Table 3. Correctional experiments (14).

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Lewis (1983), California	108 male delinquents age 14–18	E = Scared Straight C = no treatment	12 months arrests: E 81.1% (53) C 67.3% (55)	−0.41 (+21)
Cook and Spirrison (1992), Mississippi	176 male delinquents, mean age 15	E = Scared Straight C = no treatment	12 months offenses: EB 1.32 (97) EA 0.43 (97) CB 1.25 (79) CA 0.38 (79)	−0.03 (+7)
Peters et al. (1997), Cleveland	About 354 male delinquents	E = Boot Camp C = confinement	9 months convictions: E 72% (182) C 50% (172)	−0.52 (+44)*
Peters et al. (1997), Denver	About 230 male delinquents	E = Boot Camp C = confinement or probation	9 months convictions: E 39% (124) C 36% (106)	−0.07 (+8)
Peters et al. (1997), Mobile	About 526 male delinquents	E = Boot Camp C = probation (mostly)	9 months convictions: E 28% (187) C 31% (339)	0.08 (−10)
CYA (1997), California	632 male delinquents, mean age 17	E = Boot Camp C = confinement	12 months arrests: E 60.7% (313) C 58.0% (243)	−0.06 (+5)
Greenwood and Turner (1993a), Ohio	150 male delinquents, age 15 or over	E = Paint Creek C = usual training school	12 months arrests: E 50.7% (73) C 61.3% (75)	0.24 (−17)
Robinson (1995), Canada	4072 male offenders, mean age 29	E = Reasoning and Rehabilitation C = Other	12 months convictions: E 21.3% (1,673) C 24.8% (369)	0.11 (−14)
Ortmann (2000), Germany	228 male prisoners	E = social therapy prison C = usual prison	5 years convictions: E 60.4% (111) C 67.9% (112)	0.18 (−11)
Armstrong (2003), Maryland	256 male inmates, mean age 21	E = moral reconation therapy C = no treatment	2.3 years convictions: E 64.5% (110) C 64.7% (102)	0 (0)
Inciardi et al. (1997), Delaware	367 drug-involved inmates	E = therapeutic community C = no treatment	18 months arrests: E 43% (179) C 53.9% (180)	0.24 (−20)*

Table 3. Continued.

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Dugan and Everett (1998) Washington	145 drug-involved inmates	E = therapeutic community C = no treatment	2 year arrests: EM 4.5 (61) ES 6.3 (61) CM 3.4 (56) CS 5.2 (56)	-0.20 (+32)
Wexler et al. (1999), San Diego	715 drug-involved inmates, mean age 30	E = therapeutic community C = no treatment	3 years reincarceration: E 68.9% (289) C 75.1% (189)	0.17 (-8)
Marques et al. (1994), California	229 volunteer male sex offenders	E = Cognitive behavioral treatment C = No treatment	34 months sex arrests: E 10.4% (106) C 13.4% (97)	0.16 (-22)

Effect size shows standardized mean difference (*d*), with relative percentage difference between experimental and control conditions in parentheses.

\**p* < 0.05.

E: Experimental, C: Control, EB: Experimental Before, EA: Experimental After, CB: Control Before, CA: Control After, EM: Experimental Mean, ES: Experimental SD, CM: Control Mean, CS: Control SD.

Again, these results are not untypical. A systematic review of 29 evaluations of boot camps by MacKenzie et al. (2001) reported that boot camp participants had lower recidivism in nine studies, higher recidivism in eight studies and were no different in 12 studies. They concluded that “a meta-analysis found no overall significant differences in recidivism between boot camp participants and comparison samples” (p. 126).

The next four experiments, by Greenwood and Turner (1993a), Robinson (1995), Ortmann (2000) and Armstrong (2003), evaluated therapeutic programs for inmates. Paint Creek and “Reasoning and Rehabilitation” were both cognitive-behavioral programs, while social therapy and moral recondition therapy were more akin to psychotherapy. None of the four evaluations found significant desirable effects on offending, although the results were encouraging in the first three cases. A systematic review of 26 evaluations of the “Reasoning and Rehabilitation” program by Tong and Farrington (2005) concluded that it was effective in reducing reconvictions, with a weighted mean odds ratio of 1.16, corresponding to a *d* value of 0.08. Also, Wilson et al. (2004) concluded that Reasoning and Rehabilitation, moral recondition therapy and other cognitive-behavioral programs were all effective according to “higher quality” evaluations, with *d* values of 0.16, 0.33 and 0.49, respectively.

The next three experiments in Table 3 evaluated prison-based therapeutic community programs for drug-involved inmates. We could not include Richards et al. (2003) because they only reported the average number of days to arrest in each condition for arrestees, not the fraction who were arrested or the mean arrest rate of

persons in each condition. Inciardi et al. (1997) found that therapeutic community treatment had significantly desirable effects (a 20% decrease in arrests), while Wexler et al. (1999) also obtained encouraging results (an 8% decrease). However, arrest rates were 32% higher for experimental inmates in the Dugan and Everett (1998) evaluation. A systematic review of 31 evaluations of prison-based drug treatment by Mitchell et al. (2005) concluded that it was effective in reducing arrests, with a weighted mean odds ratio of 1.29, corresponding to a  $d$  value of 0.14.

The final experiment in Table 3, by Marques et al. (1994), evaluated a cognitive-behavioral treatment program for sex offenders. Those who volunteered for the treatment were placed in pairs matched on age, offense type and previous criminal history, and the members of each pair were randomly assigned to experimental or control conditions. However, it seems that only offenders who were actually transferred to the treatment program had a matched control offender. The treated group had a lower reoffending rate (for sex crimes) than the controls.

### Court experiments

Table 4 summarizes results obtained in 22 court experiments. We begin with three experiments on court-mandated treatment of male domestic violence offenders. The first experiment, by Feder and Dugan (2002), is historically important because of the battle involved in implementing it. The local prosecutor's office sought an injunction against the experiment on legal and ethical grounds, namely (a) judicial misuse of discretion (a legal issue), and (b) that it was unethical to deny treatment based on chance. In response, Feder (1998) argued that it was unethical to mandate an intervention (counseling) that had not been rigorously tested to establish its possible positive and negative consequences.

Feder (1998, p. 7) noted that "though many associations were supportive of our research, only one – the American Society of Criminology – stepped forward and agreed to serve as a friend of the court". The three ASC Presidents who were involved in this issue argued that "scholarly societies have an obligation to uphold and promulgate the principle that random assignment to treatment options is the best scientific method for determining the effectiveness of options such as those proposed in this case" (Short et al. 2000, p. 295). Eventually, the prosecutor withdrew the injunction and the experiment was completed.

In the three experiments, Davis et al. (2000) found that court-mandated psycho-educational treatment of male batterers caused a significant 44% decrease in reoffending against the same victim, while Dunford (2000) reported an encouraging 22% decrease in repeat assaults. (Note: Only results obtained with Dunford's cognitive-behavioral interventions are shown in Table 4.) However, Feder and Dugan (2002) found that counseling had no effect on rearrests. In their systematic review of court-mandated batterer treatment programs, Feder and Wilson (2005) concluded that they caused a moderate reduction in reoffending (weighted mean  $d = 0.26$ ). However, they also pointed out that these beneficial effects may apply only to a select and motivated group of convicted batterers.

Table 4. Court experiments (22).

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Feder and Dugan (2002), Florida	404 domestic violence defendants, mean age 35	E = counselling + probation C = probation	12 months arrests: E 24% (230) C 24% (174)	0.01 (0)
Davis et al. (2000), Brooklyn (NY)	380 convicted male spouse abusers, median age 31	E = psychotherapy C = community service	12 months arrests: E 14.4% (194) C 25.8% (186)	0.40 (−44)*
Dunford (2000), San Diego	861 Navy spouse assaulters, mean age 27	E = cognitive behavioral C = no treatment	12 months arrests: E 3.1% (321) C 4% (150)	0.14 (−22)
Britt et al. (1992), Pima County (AZ)	231 pretrial defendants	E = drug testing C = usual supervision	Pretrial arrests: EM 0.04 (153) ES 0.19 (153) CM 0.12 (78) CS 0.32 (78)	0.33 (−67)*
Britt et al. (1992), Maricopa County (AZ)	234 pretrial defendants	E = drug testing C = release	Pretrial arrests: EM 0.25 (118) ES 0.43 (118) CM 0.24 (116) CS 0.43 (116)	−0.02 (+4)
Britt et al. (1992), Maricopa County (AZ)	890 pretrial defendants	E = drug testing C = usual supervision/release	Pretrial arrests: EM 0.45 (425) ES 0.50 (425) CM 0.37 (465) CS 0.48 (465)	−0.16 (+22)*
Goldkamp and Jones (1992), Milwaukee	737 pretrial defendants	E = drug testing C = usual supervision	3 months arrests: E 13.6% (382) C 18.4% (332)	0.20 (−26)
Goldkamp and Jones (1992), Maryland	596 pretrial defendants	E = drug testing C = usual supervision	4 months arrests: E 10.4% (298) C 11.8% (289)	0.08 (−12)
Deschenes et al. (1995a), Maricopa County (AZ)	639 drug offenders	E = drug court C = probation	12 months arrests: E 31.3% (176) C 32.6% (454)	0.03 (−4)
Gottfredson et al. (2003), Baltimore	235 drug arrestees, mean age 35	E = drug court C = usual court	2 years arrests: E 66.2% (139) C 81.3% (96)	0.44 (−19)*
Marlowe et al. (2003), Wilmington (DE)	197 drug offenders, mean age 24	E = frequent hearings C = as needed	12 months SR arrests: E 17.6% (68) C 36.5% (52)	0.55 (−52)*
Klein (1986), USA	306 juvenile arrestees	E = court petition C = release	27 months arrests: E 72.8% (81) C 48.8% (82)	−0.57 (+49)*

Table 4. Continued.

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Schneider (1986), Boise (ID)	181 adjudicated juveniles	E = restitution C = weekend detention	22 months court referrals: E 53% (86) C 59% (95)	0.12 (−10)
Schneider (1986), Washington (DC)	411 adjudicated juveniles	E = mediation C = probation	32 months court referrals: E 54% (274) C 63% (137)	0.20 (−14)
Schneider (1986), Georgia	257 adjudicated juveniles	E = restitution C = probation or incarceration	36 months court referrals: E 47.6% (147) C 56.4% (110)	0.19 (−16)
Schneider (1986), Oklahoma	298 adjudicated juveniles	E = restitution C = probation or incarceration	24 months court referrals: E 49.5% (220) C 52% (78)	0.07 (−5)
Strang and Sherman (2005), Canberra	269 juvenile property offenders	E = restorative conference C = usual court	12 months arrests: E 32.3% (127) C 31.1% (122)	−0.03 (+4)
Strang and Sherman (2005), Canberra	143 juvenile shoplifters	E = restorative conference C = usual court	12 months arrests: E 28.6% (77) C 42.4% (66)	0.34 (−33)
Strang and Sherman (2005), Canberra	121 violent offenders	E = restorative conference C = usual court	12 months arrests: E 32.3% (62) C 37.3% (59)	0.12 (−13)
Strang and Sherman (2005), Canberra	900 drunk drivers	E = restorative conference C = usual court	12 months arrests: E 16.2% (450) C 14.0% (450)	−0.10 (+16)
McGarrell et al. (2000), Indianapolis	261 juvenile offenders	E = restorative conference C = diversion	12 months arrests: E 30.8% (130) C 41.2% (131)	0.25 (−25)
McCold and Wachtel (1998), Bethlehem (PA)	143 juvenile property offenders	E = restorative conference C = usual court	12 months arrests: E 33.3% (90) C 20.8% (53)	−0.36 (+60)

Effect size shows standardized mean difference (*d*) with relative percentage difference between experimental and control conditions in parentheses.

\**p* < 0.05.

E: Experimental, C: Control, EM: Experimental Mean, ES: Experimental SD, CM: Control Mean, CS: Control SD, SR: Self-reported.



The next five experiments evaluated the impact of frequent pretrial drug testing on arrests while on bail. Britt et al. (1992) found significantly desirable effects (a 67% reduction) in one experiment, significantly undesirable effects (a 22% increase) in another and no effect in a third. Goldkamp and Jones (1992) reported encouraging results in Milwaukee (a 28% reduction in arrests) and Maryland (a 12% reduction).

The next three experiments evaluated the effectiveness of drug courts. Compared with the usual court treatment, Gottfredson et al. (2003) found that drug courts caused a significant 19% decrease in rearrests, but Deschenes et al. (1995a) reported no effect of drug courts compared with probation. Marlowe et al. (2003) evaluated the impact of frequent judicial status hearings for drug court clients compared with hearings as needed. Rearrests were significantly (52%) fewer in the frequent hearing group, but these results were based on self-reports in interviews with about two-thirds of the offenders. In a systematic review of 42 drug court evaluations, Wilson et al. (2005) concluded that drug court programs were effective in reducing offending, with a weighted mean odds ratio of 1.79, corresponding to a  $d$  value of 0.32.

The next experiment, by Klein (1986), evaluated the effectiveness of court petitions compared with diversion programs for juveniles. Table 4 only shows the comparison of the most extreme conditions, that is court petition compared with release. The reoffending rate was significantly (49%) greater for the court cases, in agreement with labeling theory rather than deterrence.

The remaining 10 experiments evaluated restitution and restorative justice programs. Schneider (1986) carried out four experiments on restitution. In all four cases, restitution (compared with probation or incarceration) caused a nonsignificant decrease in reoffending. Strang and Sherman (2005) carried out four experiments on restorative conferences compared with usual court appearances in Canberra, Australia. The experiments showed mixed results, with a substantial decrease in reoffending for juvenile shoplifters, a small decrease for violent offenders under age 30, no effect on juvenile property offenders, and a small increase in reoffending for drunk drivers.

Conflicting results were also obtained in two other evaluations of restorative conferences that were reviewed by Strang and Sherman (2005). McGarrell et al. (2000) reported a desirable 25% reduction in reoffending by juvenile offenders, but McCold and Wachtel (1998) found an undesirable 60% increase in reoffending by juvenile property offenders. However, many cases were not treated as assigned in this last experiment.

### **Community experiments**

Table 5 summarizes the results obtained in 22 community experiments. The first eight entries are drawn from a 14-site evaluation of intensive supervision conducted by Petersilia and Turner (1993a, b). The initial sample size was greater than 100 in seven sites, and so these are included in Table 5 as separate experiments. Intensive supervision had desirable effects on recidivism (a 40%

Table 5. Community experiments (22).

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Petersilia and Turner (1993a), Seattle	173 drug-involved offenders, mean age 31	E = intensive supervision C = usual supervision	12 months arrests: E 46.1% (89) C 35.7% (84)	-0.24 (+29)
Petersilia and Turner (1993a), Des Moines (IA)	115 drug-involved offenders, mean age 30	E = intensive supervision C = usual supervision	12 months arrests: E 23.7% (59) C 28.6% (56)	0.14 (-17)
Petersilia and Turner (1993a), Dallas	221 parolees, mean age 28	E = intensive supervision C = usual supervision	12 months arrests: E 39.2% (130) C 29.7% (91)	-0.24 (+32)
Petersilia and Turner (1993a), Houston	458 parolees, mean age 29	E = intensive supervision C = usual supervision	12 months arrests: E 43.9% (239) C 40.2% (219)	-0.09 (+9)
Petersilia and Turner (1993a), Contra Costra (CA)	170 drug offenders, mean age 28	E = intensive supervision C = usual supervision	12 months arrests: E 29% (85) C 27% (85)	-0.06 (+7)
Petersilia and Turner (1993a), Ventura (CA)	168 probationers, mean age 30	E = intensive supervision C = usual supervision	12 months arrests: E 32% (82) C 53% (86)	0.50 (-40)*
Petersilia and Turner (1993a) Los Angeles	100 probationers, mean age 29	E = intensive supervision C = usual supervision	12 months arrests: E 32% (51) C 30% (49)	-0.02 (+7)
Petersilia and Turner (1993a), seven other sites	357 adult offenders, mean age 28	E = intensive supervision C = usual supervision	12 months arrests: E 35% (182) C 18% (175)	-0.47 (+94)*
Deschenes et al. (1995b), Minnesota	124 adult offenders, mean age 27	E = intensive supervision C = prison	2 years arrests: E 50% (76) C 50% (48)	0 (0)
Deschenes et al. (1995b), Minnesota	176 adult offenders, mean age 29	E = intensive supervision C = usual supervision	12 months arrests: E 15% (95) C 21% (81)	0.24 (-29)
Rhodes and Gross (1997), Portland (OR)	696 substance-using arrestees	E = intensive supervision C = no supervision	6 months SR arrests: E 22% (185) C 26.5% (376)	0.13 (-17)
Barton and Butts (1990), Detroit	511 committed juveniles	E = intensive supervision C = placement	2 years arrests: EM 3.69 (326) CM 3.58 (160)	-0.02 (+3)

Table 5. Continued.

<i>Publication, location</i>	<i>Initial sample</i>	<i>Conditions</i>	<i>Results (N)</i>	<i>Effect size d (%)</i>
Fagan (1990), four sites in USA	229 violent delinquents, mean age 17	E = reintegration C = usual program	3 years felony rearrests: E 40.4% (104) C 38.5% (91)	-0.04 (+5)
Greenwood et al. (1993), Detroit + Pittsburgh	187 delinquents	E = intensive aftercare C = usual supervision	12 months arrests: E 34.4% (96) C 32.2% (90)	-0.05 (+7)
Land et al. (1990), North Carolina	174 status offenders	E = intensive supervision C = usual supervision	12 months court referrals: E 20.4% (49) C 28.1% (57)	0.23 (-27)
Sontheimer and Goodstein (1993), Philadelphia	106 male juvenile offenders, mean age 17	E = intensive aftercare C = usual aftercare	3 months arrests: E 20.5% (44) C 34.8% (46)	0.40 (-41)
van Voorhis et al. (2004), Georgia	468 male parolees	E = Reasoning and Rehabilitation C = No Reasoning and Rehabilitation	9 months arrests: E 37.8% (229) C 42.4% (232)	0.10 (-11)
Killias et al. (2000), Switzerland	123 convicted offenders	E = community service C = prison	2 years convictions: EB 92.9% (84) EA 21.4% (84) CB 84.6% (39) CA 25.6% (39)	0.60 (-24)+
Swanson et al. (2001), North Carolina	216 mentally ill persons, mean age 39	E = outpatient commitment C = community release	12 months arrests: E 18.6% (102) C 19.3% (114)	0.02 (-4)
McAuliffe (1990), USA, Hong Kong	168 opioid addicts, mean age 31	E = relapse prevention C = no program	12 months SR arrests: E 23.1% (91) C 28.4% (74)	0.15 (-19)
Latessa and Moon (1992), USA	271 drug-involved offenders, mean age 28	E = outpatient acupuncture C = no acupuncture	arrests: E 20.1% (184) C 20.7% (87)	0.02 (-3)
Kling et al. (2005), USA	3,079 persons aged 15-20	E = move to better area C = no vouchers	5 years arrests: E 32.9% (1,233) C 31.8% (943)	-0.03 (+3)

Effect size shows standardized mean difference (*d*) with relative percentage difference between experimental and control conditions in parentheses.

\* $p < 0.05$ , + $p < 0.10$ .

E: Experimental, C: Control, EM: Experimental Mean, CM: Control Mean, EB: Experimental Before, EA: Experimental After, CB: Control Before, CA: Control After, SR: Self-reported.

decrease) in only one of these sites (Ventura). The initial sample size was about 50 in the other seven sites, and so they are amalgamated into one experiment. Intensive supervision had undesirable effects on recidivism (a 94% increase) in these seven sites. Over all 14 sites, 37% of experimental and 33% of control persons were rearrested, a nonsignificant difference. Petersilia and Turner (1993a, p. 5) speculated that the increased surveillance in the experimental condition may have increased the probability that crimes would be detected and arrests made.

In Minnesota, Deschenes et al. (1995b) evaluated two experiments on intensive supervision and found that it did not significantly reduce arrest rates compared with either prison or the usual supervision. In Portland (Oregon), Rhodes and Gross (1997) evaluated an intensive program for substance-abusing arrestees, in which case managers coordinated drug treatment programs, employment and job training agencies, housing programs, and community health services in trying to help clients in a non-judgmental way. The program had hopeful (but not significant) results, since it reduced self-reported arrests by 17%.

The next five experiments evaluated intensive supervision, community reintegration, or aftercare for juvenile offenders. The effects on recidivism were small in the experiments by Barton and Butts (1990), Fagan (1990) and Greenwood et al. (1993). They were substantial – a 27% reduction – in the study by Land et al. (1990) and also – a 41% reduction – in the study by Sontheimer and Goodstein (1993).

The final six experiments are quite heterogeneous. Van Voorhis et al. (2004) evaluated the impact of the “Reasoning and Rehabilitation” program on parolees, and found little effect. Killias et al. (2000) reported that community service, compared with imprisonment, caused a near-significant 24% decrease in convictions. There was little effect of involuntary outpatient commitment (requiring patients to comply with recommended treatment) compared with community release in the experiment by Swanson et al. (2001), of a relapse prevention program for drug addicts in the evaluation by McAuliffe (1990), or of acupuncture in the study by Latessa and Moon (1992).

In the final community experiment (Kling et al. 2005), poor families were given vouchers to enable them to move to better areas of five cities in the “Moving to Opportunity” program. The effects of this move on the offending of their children were investigated. There was little effect on the prevalence of arrests overall, but there were desirable effects on the number of arrests of females for violent and property crimes and undesirable effects on the number of arrests of males for property crimes. The authors speculated that brothers and sisters might have responded differently to new neighborhood environments; possibly females reacted to their more affluent schoolmates by trying harder in school, while males reacted with resentment, stealing from their classmates and not trying in school.

### **Meta-analyses**

Counting the number of these experiments that produced significantly desirable results is rather depressing. This was true for five out of 12 policing experiments

(with another nearly significant), five out of 13 prevention experiments, only one out of 14 correctional experiments (with another producing significantly undesirable results), four out of 22 court experiments (with another two producing significantly undesirable results), and only one out of 22 community experiments (with one nearly significant and one producing significantly undesirable results). Adding up, only 16 out of 83 experiments produced significantly desirable results, with two nearly significant and four producing significantly undesirable results. While 16 out of 83 is much greater than the chance expectation of 4, these results do not seem impressive.

However, we should not conclude from these findings that “nothing works,” because this counting method is unsatisfactory. As mentioned, a statistically significant result can reflect either a large effect in a small sample or a small effect in a large sample. It is important to measure effect size. Meta-analysis makes it possible to estimate the weighted mean effect size in a number of studies.

Table 6 shows the results of meta-analyses carried out where there were a number of reasonably similar experiments. We will assume that weighted mean effect sizes that are statistically significant (or nearly significant) and greater in magnitude than .10 indicate noteworthy effects. On this basis, we conclude that recent experiments show that prevention methods in general, and MST in particular, are effective in reducing offending. However, Scared Straight and Boot Camp programs cause an increase in offending. Correctional therapy, batterer

Table 6. Meta-analyses of effects of interventions.

	<i>No. of studies</i>	<i>No. sig.</i>	<i>Weighted mean d</i>	<i>Sig. of d</i>
Arresting domestic violence suspects	7	2	0.05	ns
Police targeting	4	2	0.07	0.011
All police	11	4	0.06	0.006
MST	5	2	0.20	0.063
All prevention	13	5	0.15	0.0001
Scared Straight/Boot Camp	6	1*	−0.13	0.063
Therapy for offenders	5	0	0.16	0.003
Therapy for drug inmates	3	1	0.14	0.059
Correctional therapy	8	1	0.16	0.0004
Court batterer programs	3	1	0.18	0.054
Pretrial drug testing	5	2*	0.00	ns
Drug court	3	2	0.19	0.023
Juvenile restitution	4	0	0.15	0.026
Restorative conferences	6	0	0.03	ns
All restorative	10	0	0.08	0.072
Intensive supervision adults	11	2*	−0.04	ns
Intensive supervision juveniles	5	0	0.02	ns
All intensive supervision	16	2*	−0.02	ns

\*One result in unfavorable direction.

Sig.: Significant, MST: Multisystemic Therapy.

programs, drug courts and juvenile restitution are effective in reducing reoffending. There are indications that police targeting of “hot-spot” places is effective in reducing crime, but the effect size is small.

In calculating weighted mean effect sizes and their confidence intervals, fixed effects models were used if the effect sizes were not significantly heterogeneous according to Q. Where effect sizes were significantly heterogeneous (for MST, all prevention, Scared Straight/boot camps, pretrial drug testing, and intensive supervision with adults), random effects models were used.

### Conclusions

The good news is that there has been a big increase in the number of randomized experiments in criminology, from 35 with offending outcomes in the Farrington (1983) review – covering 1957–1981 – to 83 with offending outcomes in the last two decades. Also, there have been many more large-scale multi-site experiments, and important replication programs. In addition, there are now several large experiments where the unit of randomization is the area; 100 or more places were randomly assigned in four experiments on policing. There are still relatively few experiments conducted outside the United States. Of the 83 experiments, four were carried out in Australia, three in Canada, one in Germany, one in Switzerland and one partly in Hong Kong. Not one experiment was carried out in the United Kingdom in the last two decades, compared with eight that were conducted before 1982 (see Farrington 2003a, b). However, eight experiments on restorative conferencing are currently being conducted in the United Kingdom by Strang and Sherman (2005). As in the earlier review, most evaluations in the present review found nonsignificant effects, but meta-analyses revealed important conclusions.

The present paper does not present an exhaustive systematic review. Only experiments based on reasonably large numbers (at least 100) and published in English were included. Because of limitations of space, the summaries of each experiment are very abbreviated. Whereas many outcomes were often reported in each evaluation, we have quoted only one, usually based on the fraction rearrested in each condition. It might be argued that self-reports of offending are less biased, but unfortunately the self-reports (where available) suffered from differential attrition and also from their highly skewed distributions; the fact that standard deviations often exceeded means meant that *d* values were problematic. We also did not attempt to report the complex multivariate analyses that researchers often carried out, for example to disaggregate results for subgroups (e.g., males *versus* females) or to control for pre-existing differences between conditions. Such statistical control should be unnecessary in a randomized experiment, because the randomization should equate conditions on all measured and unmeasured variables. In short, our main aim in this paper was to review randomized experiments that have been conducted in criminology and provide

very brief summaries of their key features. Because we focus only on one outcome, our conclusions are sometimes less positive than those drawn by the original experimenters.

Randomized experiments still present many challenges to researchers. There are often problems of getting permission and cooperation from practitioners, leading to case flow problems and difficulties in carrying through the randomization successfully. There are often serious differences between treatments assigned and treatments delivered. Differential attrition is a great problem, although it might be overcome by randomizing within matched pairs and dropping both members of a pair if one member cannot be followed up. We hope that the review of past experiments in Farrington (1983) and in the present paper will help researchers to mount randomized experiments in the future.

It is still true that the reporting of experiments is poor, with even basic information about number and gender of participants often missing. There is a great need for better specification of the nature of the intervention and of how the controls were treated so that causal conclusions can be drawn. There is a need for better measures of offending, including the number, seriousness, and cost of crimes per time at risk, so that comparable measures of effect size can be calculated in different experiments. Farrington (1983) ended his review with a checklist of items that reviewers should report, and since then the CONSORT statement has been developed for medical research (Moher et al. 2001) and adopted for American Psychological Association journals.

Our conclusions are optimistic. While randomized experiments are still relatively uncommon in criminology, the number of reasonably large-scale experiments with offending outcomes has more than doubled between 1957–1981 and 1982–2004. As Weisburd (2003) pointed out, there is a moral imperative for randomized experiments in crime and justice because of our professional obligation to provide valid answers to questions about the effectiveness of interventions. We hope that in future other countries will join the United States in carrying out more experiments in criminology.

### **Acknowledgements**

Earlier versions of this paper were presented at the Fourth Annual Jerry Lee Crime Prevention Symposium, in Washington, DC (February 2004) and at the Societies of Criminology First Key Issues Conference, in Paris, France (May 2004). We are very grateful to the following persons for providing helpful information: Marcelo Aebi, Todd Armstrong, Alison Cunningham, Elizabeth Deschenes, Terry Fain, Lynette Feder, David Festinger, Denise Gottfredson, Peter Greenwood, Adele Harrell, David Hirschel, Martin Killias, Mark Lipsey, Friedrich Losel, Jens Ludwig, William McAuliffe, Douglas Marlowe, Lorraine Mazerolle, Ojmarh Mitchell, Lawrence Schweinhart, Lawrence Sherman, Heather Strang, Richard Tremblay, Susan Turner, Pat van Voorhis and David Wilson.

## References

- Abrahamse, A. F., Ebener, P. A., Greenwood, P. W., Fitzgerald, N. & Kosin, T. E. (1991). An experimental evaluation of the Phoenix Repeat Offender Program. *Justice Quarterly* 8, 141–168.
- Armstrong, T. A. (2003). The effect of moral reconnection therapy on the recidivism of youthful offenders: A randomized experiment. *Criminal Justice and Behavior* 30, 668–687.
- Barton, W. H. & Butts, J. A. (1990). Viable options: Intensive supervision programs for juvenile delinquents. *Crime and Delinquency* 36, 238–256.
- Berk, R. A., Campbell, A., Klap, R. & Western, B. (1992). A Bayesian analysis of the Colorado Springs spouse abuse experiment. *Journal of Criminal Law and Criminology* 83, 170–200.
- Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M. & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology* 63, 569–578.
- Braga, A. A. (2001). The effects of hot spots policing on crime. *Annals of the American Academy of Political and Social Science* 578, 104–125.
- Braga, A. A., Weisburd, D. L., Waring, E. J., Mazerolle, L. G., Spelman, W. & Gajewski, F. (1999). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology* 37, 541–580.
- Britt, C. L., Gottfredson, M. R. & Goldkamp, J. S. (1992). Drug testing and pretrial misconduct: An experiment on the specific deterrent effects of drug monitoring defendants on pretrial release. *Journal of Research in Crime and Delinquency* 29, 62–78.
- California Youth Authority. (1997). *LEAD: A boot camp and intensive parole program: The final impact evaluation*. Sacramento, CA: Department of the Youth Authority.
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J. & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abercrombie project. *Applied Developmental Science* 6, 42–57.
- Cook, D. D. & Spurrison, C. L. (1992). Effects of a prisoner-operated delinquency deterrence program: Mississippi's Project Aware. *Journal of Offender Rehabilitation* 17, 89–99.
- Davis, R. C., Taylor, B. G. & Maxwell, C. D. (2000). *Does batterer treatment reduce violence? A randomized experiment in Brooklyn*. Washington, DC: National Institute of Justice.
- Deschenes, E. P., Turner, S. & Greenwood, P. W. (1995a). Drug court or probation? An experimental evaluation of Maricopa County's drug court. *Justice System Journal* 18, 55–73.
- Deschenes, E. P., Turner, S. & Petersilia, J. (1995b). A dual experiment in intensive community supervision: Minnesota's prison diversion and enhanced supervised release programs. *Prison Journal* 75, 330–356.
- Dugan, J. R. & Everett, R. S. (1998). An experimental test of chemical dependency therapy for jail inmates. *International Journal of Offender Therapy and Comparative Criminology* 42, 360–368.
- Dunford, F. W. (1990). System-initiated warrants for suspects of misdemeanor domestic assault: A pilot study. *Justice Quarterly* 7, 631–653.
- Dunford, F. W. (2000). The San Diego Navy Experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology* 68, 468–476.



- Dunford, F. W., Huizinga, D. & Elliott, D. S. (1990). The role of arrest in domestic assault: The Omaha police experiment. *Criminology* 28, 183–206.
- Eck, J. E. & Wartell, J. (1998). Improving the management of rental properties with drug problems: A randomized experiment. In L. G. Mazerolle & J. Roehl (Eds.), *Civil remedies and crime prevention* (pp. 161–185). Monsey, NY: Criminal Justice Press.
- Fagan, J. A. (1990). Treatment and reintegration of violent juvenile offenders: Experimental results. *Justice Quarterly* 7, 233–263.
- Farrington, D. P. (1983). Randomized experiments on crime and justice. In M. Tonry & N. Morris (Eds.), *Crime and justice*, vol. 4 (pp. 257–308). Chicago: University of Chicago Press.
- Farrington, D. P. (2003a). A short history of randomized experiments in criminology: A meager feast. *Evaluation Review* 27, 218–227.
- Farrington, D. P. (2003b). British randomized experiments on crime and justice. *Annals of the American Academy of Political and Social Science* 589, 150–167.
- Farrington, D. P. (2003c). Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Science* 587, 49–68.
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W. & Welsh, B. C. (2002). The Maryland scientific methods scale. In L. W. Sherman, D. P. Farrington, B. C. Welsh & D. L. Mackenzie (Eds.), *Evidence-based crime prevention* (pp. 13–21). London: Routledge.
- Farrington, D. P. & Welsh, B. C. (Eds.). (2001). What works in preventing crime? Systematic reviews of experimental and quasi-experimental research. *Annals of the American Academy of Political and Social Science*, vol. 578.
- Farrington, D. P. & Welsh, B. C. (2003). Family-based prevention of offending: A meta-analysis. *Australian and New Zealand Journal of Criminology* 36, 127–151.
- Feder, L. (1998). Using random assignment in social science settings. *Professional Ethics Report* 11(1), 1–7.
- Feder, L. & Dugan, L. (2002). A test of the efficacy of court-mandated counselling for domestic violence offenders: The Broward experiment. *Justice Quarterly* 19, 343–375.
- Feder, L., Wilson, D. B. (2005). Mandated batterer programs to reduce domestic violence. In B. C. Welsh & D. P. Farrington (Eds.), *Preventing crime: What works for children, offenders, victims and places*. Belmont, CA: Wadsworth, in press.
- Garner, J., Fagan, J. & Maxwell, C. (1995). Published findings from the spouse assault replication program: A critical review. *Journal of Quantitative Criminology* 11, 3–28.
- Goldkamp, J. S. & Jones, P. R. (1992). Pretrial drug-testing experiments in Milwaukee and Prince George's County: The context of implementation. *Journal of Research in Crime and Delinquency* 29, 430–465.
- Gottfredson, D. C., Najaka, S. S. & Kearly, B. (2003). Effectiveness of drug treatment courts: Evidence from a randomized trial. *Criminology and Public Policy* 2, 171–196.
- Greenwood, P. W., Deschenes, E. P. & Adams, J. (1993). *Chronic juvenile offenders: Final results from the Skillman aftercare experiment*. Santa Monica, CA: Rand.
- Greenwood, P. W. & Turner, S. (1993a). Evaluation of the Paint Creek Youth Center: A residential program for serious delinquents. *Criminology* 31, 263–279.
- Greenwood, P. W. & Turner, S. (1993b). Private presentence reports for serious juvenile offenders: Implementation issues and impacts. *Justice Quarterly* 10, 229–243.
- Grossman, J. B. & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation Review* 22, 403–426.
- Harrell, A. V., Cavanagh, S. E. & Sridharan, S. (1999). *Evaluation of the Children at Risk program: Results 1 year after the end of the program*. (Research in Brief). Washington, DC: National Institute of Justice.

- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G. & Hanley, J. H. (1997). Multisystemic theory with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology* 65, 821–833.
- Henggeler, S. W., Rowland, M. D., Randall, J., Ward, D. M., Pickrel, S. G., Cunningham, P. B., Miller, S. L., Edwards, J., Zealberg, J. J., Hand, L. D., & Santos, A. B. (1999). Home-based multisystemic therapy as an alternative to the hospitalization of the youths in psychiatric crisis: Clinical outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry* 38, 1331–1339.
- Henggeler, S. W., Clingempeel, W. G., Brondino, M. J. & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry* 41, 868–874.
- Hindelang, M. J., Hirschi, T. & Weis, J. G. (1981). *Measuring delinquency*. Beverly Hills, CA: Sage.
- Hirschel, J. D., Hutchison, I. W. & Dean, C. W. (1992). The failure of arrest to deter spouse abuse. *Journal of Research in Crime and Delinquency* 29, 7–33.
- Inciardi, J. A., Martin, S. S., Butzin, C. A., Hopper, R. M. & Harrizon, L. D. (1997). An effective model of prison-based treatment for drug-involved offenders. *Journal of Drug Issues* 27, 261–278.
- Killias, M., Aebi, M. & Ribeaud, D. (2000). Does community service rehabilitate better than short-term imprisonment? Results of a controlled experiment. *Howard Journal* 39, 40–57.
- Klein, M. W. (1986). Labeling theory and delinquency policy: An experimental test. *Criminal Justice and Behavior* 13, 47–79.
- Kling, J. R., Ludwig, J. & Katz, L. F. (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *Quarterly Journal of Economics* 120, in press.
- Land, K. C., McCall, P. L. & Williams, J. R. (1990). Something that works in juvenile justice: An evaluation of the North Carolina court counselors' intensive protective supervision randomized experimental project, 1987–1989. *Evaluation Review* 14, 574–606.
- Latessa, E. J. & Moon, M. M. (1992). The effectiveness of acupuncture in an outpatient drug treatment program. *Journal of Contemporary Criminal Justice* 8, 317–331.
- Leschied, A. & Cunningham, A. (2002). *Seeking effective interventions for serious young offenders: Interim results of a four-year randomized study of multisystemic therapy in Ontario, Canada*. London, Canada: London Family Court Clinic.
- Lewis, R. V. (1983). Scared straight – California style: Evaluation of the San Quentin Squires program. *Criminal Justice and Behavior* 10, 209–226.
- Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- MacKenzie, D. L., Wilson, D. B. & Kider, S. B. (2001). Effects of correctional boot camps on offending. *Annals of the American Academy of Political and Social Science* 578, 126–143.
- Marlowe, D. B., Festinger, D. S., Lee, P. A., Schepise, M. M., Hazzard, J. E. R., Merrill, J. C., Mulvaney, F. D. & McLellan, A. T. (2003). Are judicial status hearings a key component of drug court? During-treatment data from a randomized trial. *Criminal Justice and Behavior* 30, 141–162.
- Marques, J. K., Day, D. M., Nelson, C. & West, M. A. (1994). Effects of cognitive-behavioral treatment on sex offender recidivism: Preliminary results from a longitudinal study. *Criminal Justice and Behavior* 21, 28–54.

- Maxwell, C. D., Garner, J. H. & Fagan, J. A. (2002). The preventive effects of arrest on intimate partner violence: Research, policy, and theory. *Criminology and Public Policy* 2, 51–79.
- Mazerolle, L. G., Roehl, J. & Kadleck, C. (1998). Controlling social disorder using civil remedies: Results from a randomized field experiment in Oakland, California. In L. G. Mazerolle & Jan Roehl (Eds.), *Civil remedies and crime prevention* (pp. 141–159). Monsey, NY: Criminal Justice Press.
- McAuliffe, W. E. (1990). A randomized controlled trial of recovery training and self-help for opioid addicts in New England and Hong Kong. *Journal of Psychoactive Drugs* 22, 197–209.
- McCold, P. & Wachtel, T. (1998). *Restorative policing experiment: The Bethlehem Pennsylvania police family group conferencing project*. Pipersville, PA: Community Service Foundation.
- McDonald, D. C., Greene, J. & Worzella, C. (1992). *Day fines in American courts: The Staten Island and Milwaukee experiments*. Washington, DC: National Institute of Justice.
- McGarrell, E., Olivares, K., Crawford, K. & Kroovand, N. (2000). *Returning justice to the community: The Indianapolis restorative justice experiment*. Indianapolis: Hudson Institute Crime Control Policy Center.
- Meeker, J. W. & Binder, A. (1990). Experiments as reforms: The impact of the “Minneapolis Experiment” on police policy. *Journal of Police Science and Administration* 17, 147–153.
- Mills, P. E., Cole, K. N., Jenkins, J. R. & Dale, P. S. (2002). Early exposure to direct instruction and subsequent juvenile delinquency: A prospective examination. *Exceptional Children* 69, 85–96.
- Mitchell, O., MacKenzie, D. L., Wilson, D. B. (2005). Incarceration-based drug treatment. In B. C. Welsh & D. P. Farrington (Eds.), *Preventing crime: What works for children, offenders, victims and places*. Belmont, CA: Wadsworth, in press.
- Moher, D., Schulz, K. F. & Altman, D. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association* 285, 1987–1991.
- Olds, D. L., Henderson, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P. & Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized controlled trial. *Journal of the American Medical Association* 280, 1238–1244.
- Ortmann, R. (2000). The effectiveness of social therapy in prison: A randomized experiment. *Crime and Delinquency* 46, 214–232.
- Pate, A. & Hamilton, E. E. (1992). Formal and informal deterrents to domestic violence: The Dade County spouse assault experiment. *American Sociological Review* 57, 691–697.
- Peters, M., Thomas, D. & Zamberlan, C. (1997). *Boot camps for juvenile offenders*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention (Program Summary).
- Petersilia, J. & Turner, S. (1993a). *Evaluating intensive supervision probation/parole: Results of a nationwide experiment*. Washington DC: National Institute of Justice (Research in Brief).
- Petersilia, J. & Turner, S. (1993b). Intensive probation and parole. In M. Tonry (Ed.), *Crime and justice*, vol. 17 (pp. 281–335). Chicago: University of Chicago Press.
- Petrosino, A. J., Boruch, R. F., Rounding, C., McDonald, S. & Chalmers, I. (2000). The Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR) to facilitate the preparation and maintenance of systematic

- reviews of social and educational interventions. *Evaluation Research in Education* 14, 293–307.
- Petrosino, A. J., Boruch, R. F., Farrington, D. P., Sherman, L. W. & Weisburd, D. (2003a). Toward evidence-based criminology and criminal justice: Systematic reviews, the Campbell Collaboration, and the Crime and Justice Group. *International Journal of Comparative Criminology* 3, 42–61.
- Petrosino, A. J., Turpin-Petrosino, C. & Buehler, J. (2003b). Scared straight and other juvenile awareness programs for preventing juvenile delinquency: A systematic review of the randomized experimental evidence. *Annals of the American Academy of Political and Social Science* 589, 41–62.
- Rhodes, W. & Gross, M. (1997). *Case management reduces drug use and criminality among drug-involved arrestees: An experimental study of an HIV prevention intervention*. Washington, DC: National Institute of Justice.
- Richards, H. J., Casey, J. O. & Lucente, S. W. (2003). Psychopathy and treatment response in incarcerated female substance abusers. *Criminal Justice and Behavior* 30, 251–276.
- Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders*. Ottawa: Correctional Service of Canada (Research Report R-41).
- Schneider, A. L. (1986). Restitution and recidivism rates of juvenile offenders: Results from four experimental studies. *Criminology* 24, 533–552.
- Schochet, P. Z., Burghardt, J. & Glazerman, S. (2001). *National Job Corps study: The impacts of job corps on participants' employment and related outcomes*. Princeton, NJ: Mathematica Policy Research.
- Schweinhart, L. J., Montie, J., Zongping, X., Barnett, W. S., Belfield, C. R. & Nores, M. (2005). *Lifetime effects: The High/Scope Perry preschool study through age 40*. Ypsilanti, MI: High/Scope Press.
- Sherman, L. W. (1992). *Policing domestic violence: Experiments and dilemmas*. New York: Free Press.
- Sherman, L. W. & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American Sociological Review* 49, 261–272.
- Sherman, L. W. & Cohn, E. G. (1989). The impact of research on legal policy: The Minneapolis domestic violence experiment. *Law and Society Review* 23, 117–144.
- Sherman, L. W. & Rogan, D. P. (1995). Deterrent effects of police raids on crack houses: A randomized controlled experiment. *Justice Quarterly* 12, 755–781.
- Sherman, L. W. & Weisburd, D. (1995). General deterrent effects of police patrol in crime “hot spots”: A randomized controlled trial. *Justice Quarterly* 12, 625–648.
- Sherman, L. W., Schmidt, J. D., Rogan, D. P., Smith, D. A., Gartin, P. R., Cohen, E. G., Collins, D. J. & Bacich, A. R. (1992). The variable effects of arrest on criminal careers: The Milwaukee domestic violence experiment. *Journal of Criminal Law and Criminology* 83, 137–169.
- Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J. E., Reuter, P. & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: Office of Justice Programs.
- Sherman, L. W., Farrington, D. P., Welsh, B. C. & MacKenzie, D. L. (Eds.). (2002). *Evidence-based crime prevention*. London: Routledge.
- Short, J. F., Zahn, M. A. & Farrington, D. P. (2000). Experimental research in criminal justice settings: Is there a role for scholarly societies? *Crime and Delinquency* 46, 295–298.
- Sontheimer, H. & Goodstein, L. (1993). An evaluation of juvenile intensive aftercare probation: Aftercare versus system response effects. *Justice Quarterly* 10, 197–227.

- Strang, H. & Sherman, L. W. (2005). Restorative justice to reduce victimization. In B. C. Welsh & D. P. Farrington (Eds.), *Preventing crime: What works for children, offenders, victims and places*. Belmont, CA: Wadsworth, in press.
- Swanson, J. W., Borum, R., Swartz, M. S., Hiday, V. A., Wagner, H. R. & Burns, B. J. (2001). Can involuntary outpatient commitment reduce arrests among persons with severe mental illness? *Criminal Justice and Behavior* 28, 156–189.
- Tong, L. S. J. & Farrington, D. P. (2005). How effective is the “Reasoning and Rehabilitation” program in reducing reoffending? A meta-analysis of evaluations in four countries. *Psychology, Crime and Law*, in press.
- Tremblay, R. E., Masse, L. C., Pagani, L. & Vitaro, F. (1996). From childhood physical aggression to adolescent maladjustment: The Montreal prevention experiment. In R. D. Peters & R. J. McMahon (Eds.), *Preventing childhood disorders, substance use and delinquency* (pp. 268–298). Thousand Oaks, CA: Sage.
- van Voorhis, P., Spruance, L. M., Ritchey, P. N., Listwan, S. J. & Seabrook, R. (2004). The Georgia cognitive skills experiment: A replication of Reasoning and Rehabilitation. *Criminal Justice and Behavior* 31, 282–305.
- Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review* 27, 336–354.
- Weisburd, D. & Green, L. (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly* 12, 711–735.
- Weisburd, D., Lum, C. M. & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *Annals of the American Academy of Political and Social Science* 578, 50–70.
- Welsh, B. C., Farrington, D. P. (Eds.). (2005). *Preventing crime: What works for children, offenders, victims and places*. Belmont, CA: Wadsworth, in press.
- Wexler, H. K., Melnick, G., Lowe, L. & Peters, J. (1999). Three-year reincarceration outcomes for Amity in-prison therapeutic community and aftercare in California. *Prison Journal* 79, 321–336.
- Wilson, D. B., Allen, L. C., MacKenzie, D. L. (2004). A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Criminal Justice and Behavior*, in press.
- Wilson, D. B., Mitchell, O. MacKenzie, D. L. (2005). A systematic review of drug court effects on recidivism. *Criminology and Public Policy*, in press.

### About the authors

**David P. Farrington**, O.B.E., is a Professor of Psychological Criminology at the Institute of Criminology, Cambridge University. He is a Fellow of the British Academy, of the Academy of Medical Sciences, of the British Psychological Society and of the American Society of Criminology, and an Honorary Life Member of the British Society of Criminology and of the Division of Forensic Psychology of the British Psychological Society. He is Co-Chair of the Campbell Collaboration Crime and Justice Group, a member of the Board of Directors of the International Society of Criminology, joint editor of *Cambridge Studies in Criminology* and of the journal *Criminal Behaviour and Mental Health*, and a member of the editorial boards of 13 other journals. He received B.A., M.A. and Ph.D. degrees in psychology from Cambridge University, the Sellin-Glueck Award of the American Society of Criminology for international contributions to criminology, and the Sutherland Award of the American Society of Criminology for outstanding contributions to criminology.

**Brandon C. Welsh** is an Associate Professor in the Department of Criminal Justice at the University of Massachusetts Lowell. He received his Ph.D. in Criminology from Cambridge University. He is an author or editor of five books, including *Evidence-Based Crime Prevention* (Routledge 2002, with Lawrence Sherman, David Farrington, and Doris MacKenzie), and was as an editor of a special issue of the *Annals of the American Academy of Political and Social Science*, entitled “What Works in Preventing Crime? Systematic Reviews of Experimental and Quasi-Experimental Research” (2001). Recent articles have appeared in *Criminology & Public Policy*, *Justice Quarterly*, *Youth Violence and Juvenile Justice*, and *British Journal of Criminology*.