

DOES CORRECTIONAL TREATMENT WORK? A CLINICALLY RELEVANT AND PSYCHOLOGICALLY INFORMED META-ANALYSIS*

D.A. ANDREWS
IVAN ZINGER
ROBERT D. HOGE
Carleton University

JAMES BONTA
Ottawa-Carleton Detention Centre

PAUL GENDREAU
Centracare Saint John, New Brunswick

FRANCIS T. CULLEN
University of Cincinnati

Careful reading of the literature on the psychology of criminal conduct and of prior reviews of studies of treatment effects suggests that neither criminal sanctioning without provision of rehabilitative service nor servicing without reference to clinical principles of rehabilitation will succeed in reducing recidivism. What works, in our view, is the delivery of appropriate correctional service, and appropriate service reflects three psychological principles: (1) delivery of service to higher risk cases, (2) targeting of criminogenic needs, and (3) use of styles and modes of treatment (e.g., cognitive and behavioral) that are matched with client need and learning styles. These principles were applied to studies of juvenile and adult correctional treatment, which yielded 154 phi coefficients that summarized the magnitude and direction of the impact of treatment on recidivism. The effect of appropriate correctional service (mean phi = .30) was significantly ($p < .05$) greater than that of unspecified correctional service (.13), and both were more effective than inappropriate service (-.06) and non-service criminal sanctioning (-.07). Service was effective within juvenile and adult corrections, in studies published before and after 1980, in randomized and nonrandomized designs, and in diversionary, community, and residential programs (albeit, attenuated in residential settings). Clinical sensitivity and a psychologically informed perspective on crime

* This paper is dedicated to Daniel Glaser, Ted Palmer, and Marguerite Q. Warren.

may assist in the renewed service, research, and conceptual efforts that are strongly indicated by our review.

During the 1970s, the ideological hegemony of the individualized treatment ideal suffered a swift and devastating collapse (Rothman, 1980). Previously a code word for "doing good," rehabilitation came to be seen by liberals as a euphemism for coercing offenders and by conservatives as one for letting hardened criminals off easily. Although the public's belief in rehabilitation was never eroded completely (Cullen et al., 1988), defenders of treatment were branded scientifically and politically naive apologists for the socially powerful, self-serving human service professionals, or curious relics of a positivistic past. Thus, a number of jurisdictions in the United States (Cullen and Gilbert, 1982) and Canada (Andrews, 1990; Leschied et al., 1988) embarked on sentencing reforms that undercut the role of rehabilitation in justice and corrections.

The decline of the rehabilitative ideal cannot be attributed to a careful reading of evidence regarding the effectiveness of rehabilitative treatment. As will be shown, reviews of the effectiveness literature routinely found that a substantial proportion of the better-controlled studies of rehabilitative service reported positive effects, and did so for programs that operated within a variety of conditions established by criminal sanctions, such as probation or incarceration. We will also show that criminal sanctions themselves were typically found to be only minimally related to recidivism. Thus, rather than a rational appreciation of evidence, the attack on rehabilitation was a reflection of broader social and intellectual trends. This is evident upon consideration of the particular historical timing and intensity of the attack on rehabilitation.

First, the rapidly changing sociopolitical context of the decade preceding the mid-1970s propelled conservatives to seek "law and order," while liberals attached to class-based perspectives on crime became discouraged about the benevolence of the state and the promise of direct intervention (Allen, 1981; Cullen and Gendreau, 1989). Second, an emerging social science, informed by labelling and critical/Marxist approaches, embraced antipsychological and often anti-empirical themes (Andrews, 1990; Andrews and Wormith, 1989). These emergent perspectives played an important role in legitimating the decision of many academic criminologists and juridical policymakers to declare rehabilitation fully bankrupt. Most noteworthy was Robert Martinson's (1974:25) conclusion that "the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism." In short order, with the blessing of a major academy of science (Séchrest et al., 1979), the notion that "nothing works" became accepted doctrine (Walker, 1989). "Nothing works" satisfied conservative political reactions to the apparent disorder of the 1960s, liberal sorrow over perceived failures of the Great Society,

and the ideological persuasions of those academicians whose truly social visions of deviance asserted that only radical social change could have an impact on crime.

In the 1980s, however, rehabilitation and respect for evidence made at least a modest comeback. As will be noted, a number of revisionist scholars have observed that the marriage of conservative politics and leftist social science—in both its “discouraged liberal” and “critical/Marxist” versions—has neither improved justice nor increased crime control. In any case, it is our thesis that evidence of effective treatment was there from the earliest reviews, now is mounting, and constitutes a persuasive case against the “nothing works” doctrine.

Even so, criticisms of rehabilitation are not in short supply. As Walker (1989:231) comments: “It is wishful thinking to believe that additional research is going to uncover a magic key that has somehow been overlooked for 150 years.” Other scholars—as exemplified most notably and recently by Whitehead and Lab (1989; Lab and Whitehead, 1988)—continue to participate in the scientific exchange on intervention and to present evidence ostensibly bolstering the “nothing works” message.

Whitehead and Lab’s (1989) report is very much in the tradition of the reviews and conclusions that are challenged in this paper. Before detailing our position, however, we note that the Whitehead and Lab review is important for several reasons. First, having searched the psychological, sociological, and criminological journals, they produced an impressively complete set of controlled evaluations of juvenile treatment for the years 1975 to 1984. They coded the setting of treatment and distinguished among diversion programs (within and outside the juvenile justice system), probation and other community-based programs, and residential programming. Moreover, they coded type of treatment within these settings as either behavioral or nonbehavioral and considered recency (year of publication) and quality of research design. Focused exclusively upon evaluations employing recidivism as an outcome variable, their conclusions actually had to do with crime control. Clearly then, the negative conclusion of Whitehead and Lab is worthy of serious consideration by those in criminal justice.

Most serious, and unlike most earlier reviews—including the Martinson (1974) review—portions of the Whitehead and Lab (1989) paper support a very firm version of “nothing works.” That is, the methodological, clinical, and sampling caveats typically listed by earlier reviewers were discounted systematically in Whitehead and Lab (1989). Regarding quality of the research, the more rigorous studies were reported to find correctional treatment to have effects even more negative than did the less rigorous studies. As to standards of effectiveness, Whitehead and Lab advised that their standard (a phi coefficient of .20 or greater) was so generous that evidence favorable to treatment would certainly have emerged had positive evidence, in fact, existed. In

regard to type of treatment, they admitted that behavioral forms of intervention may be effective with outcomes other than recidivism, but they found behavioral treatment to be no more effective than nonbehavioral approaches in the control of recidivism.

Our meta-analysis includes, but is not confined to, the Whitehead and Lab (1989) sample of studies. Challenging sweeping conclusions regarding program ineffectiveness, we reaffirm a line of analysis for developing meaningful conclusions on the conditions under which programs will work. Our challenge is informed by considerations of research and theory on the causes of crime and by research and theory on behavioral influence processes. In particular, a growing number of scholars and practitioners now agree with what was always the starting point of the Gluecks (1950), the Grants (1959), Glaser (1974), and Palmer (1975): The effectiveness of correctional treatment is dependent upon what is delivered to whom in particular settings. Certainly that has been our view¹ and the view of many other reviewers and commentators.²

CLINICALLY RELEVANT AND PSYCHOLOGICALLY INFORMED PROGRAMMING, EVALUATION, AND META-ANALYSIS

The psychology of criminal conduct recognizes multiple sources of variation in criminal recidivism (Andrews, 1980, 1983; Andrews and Kiessling, 1980; Andrews et al., 1990; Cullen and Gendreau, 1989; Hoge and Andrews, 1986; Palmer, 1983; Warren, 1969). These major sources of variation are found through analyses of the main and interactive effects of (a) preservice characteristics of offenders, (b) characteristics of correctional workers, (c) specifics of the content and process of services planned and delivered, and (d) intermediate changes in the person and circumstances of individual offenders. Logically, these major sources of variation in outcome reside within the conditions established by the specifics of a judicial disposition or criminal sanction. Thus, there is little reason to expect that variation among settings or sanctions will have an impact on recidivism except in interaction with offender characteristics and through the mediators of intervention process and intermediate change. We develop this "criminal sanction" hypothesis first and then compare it with hypotheses regarding the effectiveness of a correctional service approach that attends to preservice case characteristics, to

1. Andrews (1980, 1983, 1990), Andrews and Kiessling (1980), Andrews et al. (1990), Cullen and Gendreau (1989), Gendreau and Ross (1979, 1981, 1987).

2. Basta and Davidson (1988), Currie (1989), Garrett (1985), Geismar and Wood (1985), Greenwood and Zimring (1985), Izzo and Ross (1990), Lipsey (1989), Martinson (1979), Mayer et al. (1986), Palmer (1983), Ross and Fabiano (1985).

the process and content of intervention, and to intermediate change within particular sanctions.

IN THEORY, WHY SHOULD CRIMINAL SANCTIONING WORK?

A focus upon variation in official disposition is a reflection of one or more of the three sets of theoretical perspectives known as *just deserts*, *labelling*, and *deterrence*. The just deserts or justice set is not overly concerned with recidivism, but on occasion the assumption surfaces that unjust processing may motivate additional criminal activity (Schur, 1973:129). It appears, however, that the devaluation of rehabilitation—in the interest of increasing “just” processing—has been associated with increased punishment and decreased treatment but not with reduced recidivism (Cullen and Gilbert, 1982; Leschied et al., 1988).

The labelling and deterrence perspectives actually yield conflicting predictions regarding the outcomes of different dispositions (Rausch, 1983). Labelling theory suggests that less involvement in the criminal justice system is better than more (because the stigma is less), while deterrence theory suggests the opposite (because fear of punishment is greater). The assumptions of both labelling (Andrews and Wormith, 1989; Wellford, 1975) and deterrence (Gendreau and Ross, 1981) have been subjected to logical and empirical review, and neither perspective is yet able to offer a well-developed psychology of criminal conduct. Basic differentiations among and within levels and types of sanctions have yet to be worked out (Smith and Gartin, 1989), type of offender is likely a crucial moderating variable (Klein, 1986), and the social psychology of “processing” is only now being explored (Link et al., 1989).

IN FACT, DOES CRIMINAL SANCTIONING WORK?

To our knowledge, not a single review of the effects of judicial sanctioning on criminal recidivism has reached positive conclusions except when the extremes of incapacitation are tested or when additional reference is made to moderators (e.g., type of offender) or mediators (e.g., the specifics of intervention). Reading Kirby (1954), Bailey (1966), Logan (1972), and Martinson (1974) reveals the obvious but unstated fact that their negative conclusions regarding “treatment” reflected primarily the negligible impact of variation in sanctions such as probation and incarceration. Thus, we agree with Palmer (1975): The main effects of criminal sanctions on recidivism have been slight and inconsistent.

This hypothesis is extended to judicial “alternatives,” because there are no solid reasons for expecting alternative punishments, such as community service or restitution, to have an impact on recidivism. Any anticipated rehabilitative benefit of “alternatives” is based on the hope that offenders will learn that crime has negative consequences, and yet the enhancement of cognitive

and interpersonal skills (e.g., future-orientation and perspective-taking) are dependent upon systematic modeling, reinforcement, and graduated practice (Ross and Fabiano, 1985). Given little reason to expect much from the incidental learning opportunities provided by such sanctions as restitution, correctional treatment service is a crucial supplement to a criminal justice approach that is preoccupied with avoiding stigma while delivering "just" and "innovative alternative" punishment.

CORRECTIONAL TREATMENT SERVICES

Reviewers of the literature have routinely found that at least 40% of the better-controlled evaluations of correctional treatment services reported positive effects (Andrews et al., 1990). For example, considering only the better-controlled studies, the proportion of studies reporting positive evidence was 75% ($3/4$) in Kirby (1954), 59% ($13/22$) in Bailey (1966), 50% ($9/18$) in Logan (1972), 78% ($14/18$) in Logan when Type of Treatment \times Type of Client interactions are considered, 48% ($39/82$) in Palmer's (1975) retabulation of studies reviewed by Martinson (1974), 86% ($82/95$) in Gendreau and Ross (1979), and 47% ($40/85$) in Lab and Whitehead (1988). This pattern of results strongly supports exploration of the idea that some service programs are working with at least some offenders under some circumstances, and we think that helpful linkages among case, service, and outcome are suggested by three principles known as risk, need, and responsivity (Andrews et al., 1990).

THE RISK PRINCIPLE AND SELECTION OF LEVEL OF SERVICE

The risk principle suggests that higher levels of service are best reserved for higher risk cases and that low-risk cases are best assigned to minimal service. In the literature at least since the Gluecks (1950), the risk principle has been restated on many occasions (e.g., Glaser, 1974). Although the parameters remain to be established, evidence favoring the risk principle continues to grow (Andrews et al., 1990). In brief, when actually explored, the effects of treatment typically are found to be greater among higher risk cases than among lower risk cases. This is expected unless the need and/or responsivity principles are violated.

THE NEED PRINCIPLE AND SELECTION OF APPROPRIATE INTERMEDIATE TARGETS

Risk factors may be static or dynamic in nature, and psychology is particularly interested in those dynamic risk factors that, when changed, are associated with *subsequent* variation in the chances of criminal conduct. Clinically, dynamic risk factors are called *criminogenic needs*, and guidelines for their assessment are described elsewhere (Andrews, 1983; Andrews et al., 1990).

The most promising intermediate targets include changing antisocial attitudes, feelings, and peer associations; promoting familial affection in combination with enhanced parental monitoring and supervision; promoting identification with anticriminal role models; increasing self-control and self-management skills; replacing the skills of lying, stealing, and aggression with other, more prosocial skills; reducing chemical dependencies; and generally shifting the density of rewards and costs for criminal and noncriminal activities in familial, academic, vocational, and other behavioral settings.³ Theoretically, modifying contingencies within the home, school, and work by way of an increased density of reward for noncriminal activity may reduce motivation for crime and increase the costs of criminal activity through having more to lose (Hunt and Azrin, 1973).

Less-promising targets include increasing self-esteem without touching antisocial propensity (e.g., Wormith, 1984), increasing the cohesiveness of antisocial peer groups (e.g., Klein, 1971), improving neighborhood-wide living conditions without reaching high-risk families (the East Side, Midcity, and other community projects in Klein, 1971, and Schur, 1973), and attempts to focus on vague personal/emotional problems that have not been linked with recidivism (Andrews and Kiessling, 1980).

THE RESPONSIVITY PRINCIPLE AND SELECTION OF TYPE OF SERVICE

The responsivity principle has to do with the selection of styles and modes of service that are (a) capable of influencing the specific types of intermediate targets that are set with offenders and (b) appropriately matched to the learning styles of offenders. We begin with the general literature on the treatment of offenders and then turn to specific Responsivity \times Service interactions.

Responsivity: General principles of effective service. Drawing upon our earlier review (Andrews et al., 1990), appropriate types of service typically, but not exclusively, involve the use of behavioral and social learning principles of interpersonal influence, skill enhancement, and cognitive change. Specifically, they include modeling, graduated practice, rehearsal, role playing, reinforcement, resource provision, and detailed verbal guidance and explanations (making suggestions, giving reasons, cognitive restructuring). Elsewhere (Andrews and Kiessling, 1980), we describe the applications of these practices as (a) use of authority (a "firm but fair" approach and definitely not interpersonal domination or abuse), (b) anticriminal modeling and reinforcement (explicit reinforcement and modeling of alternatives to procriminal styles of thinking, feeling, and acting), and (c) concrete problem solving and

3. For example, Andrews et al. (1990), Andrews and Wormith (1989), Glueck and Glueck (1950), Johnson (1979), Loeber and Stouthamer-Loeber (1987), Wilson and Herrnstein (1985).

systematic skill training for purposes of increasing reward levels in anticriminal settings. High levels of advocacy and brokerage are also indicated as long as the receiving agency actually offers appropriate service. Finally, Andrews and Kiessling (1980) recommended that service deliverers relate to offenders in interpersonally warm, flexible, and enthusiastic ways while also being clearly supportive of anticriminal attitudinal and behavioral patterns. Interestingly, social learning approaches receive strong, albeit indirect, support from the prediction literature on the causal modeling of delinquency (Akers and Cochran, 1985; Jessor and Jessor, 1977).

Responsivity: Ineffective service. Some types and styles of services should be avoided under most circumstances (Andrews et al., 1990). Generally, programming for groups is to be approached very cautiously because the opening up of communication within offender groups may well be criminogenic (Andrews, 1980). In group and residential programming, clinicians must gain control over the contingencies of interaction so that anticriminal, rather than procriminal, patterns are exposed and reinforced (Buehler et al., 1966). For example, Agee's (1986) programmatic structures supporting positive change may be contrasted with the failure of unstructured, peer-oriented group counseling and permissive, relationship-oriented milieu approaches. The failure of these unstructured approaches is well documented in open community settings (e.g., Faust, 1965; Klein, 1971), in group homes operating according to the essentially nondirective guidelines of "guided group interaction" (Stephenson and Scarpitti, 1974:Ch. 8), in hospitals (Craft et al., 1966), and in prisons (Kassebaum et al., 1971; Murphy, 1972). There are also no convincing theoretical grounds for believing that young people will be "scared straight" (Finckenauer, 1982). Fear of official punishment is not one of the more important correlates of delinquency (Johnson, 1979), and yelling at people is counter to the relationship principle of effective service (Andrews, 1980).

Finally, traditional psychodynamic and nondirective client-centered therapies are to be avoided within general samples of offenders (Andrews et al., 1990). These therapies are designed to free people from the personally inhibiting controls of "superego" and "society," but neurotic misery and overcontrol are not criminogenic problems for a majority of offenders. Authorities such as Freud (in his introductory lectures on psychoanalysis, 1953) and the Gluecks (in their classic *Unraveling*, 1950) warned us about evocative and relationship-dependent psychodynamic approaches with antisocial cases.

Specific responsivity considerations. The success of highly verbal, evocative, and relationship-dependent services seems to be limited to clients with high levels of interpersonal, self-reflective, and verbal skill. The "I-Level" (Harris, 1988) and "Conceptual Level" (Reitsma-Street and Leschied, 1988) systems

provide guidance regarding the types of offenders who may respond in positive ways to services that are less structured than those we have been describing as appropriate for antisocial samples in general.

SUMMARY

Our clinically relevant and psychologically informed principles of treatment predict that criminal sanctioning without attention to the delivery of correctional service will relate to recidivism minimally. Additionally, we suggest that the delivery of services, regardless of criminal sanction or setting, is unproductive if those services are inconsistent with the principles of risk, need, and responsivity. Positively, we predict that appropriate treatment—treatment that is delivered to higher risk cases, that targets criminogenic need, and that is matched with the learning styles of offenders—will reduce recidivism.

METHOD

SAMPLES OF STUDIES

We subjected 45 of the 50 studies included in the Whitehead and Lab (1989) review to content and meta-analysis.⁴ The Whitehead and Lab sample included only studies of juvenile treatment that appeared in professional journals between 1975 and 1984 and that presented effects of treatment on binary (less-more) measures of recidivism. Studies that focused on imprisonment or the treatment of substance abuse were not included.

We also explored a second sample of studies in order to check on the generalizability of any findings based on the Whitehead and Lab sample. Sample 2 included 35 studies in our research files as of February 1989 that were not included in the Whitehead and Lab set but had employed binary measures of recidivism. Studies in sample 2 date from the 1950s through 1989, but they are not purported to be a representative sample of any particular time period. Sample 2 provides a convenient means of exploring, albeit tentatively, how well conclusions based on the Whitehead and Lab sample may generalize to adult samples.

ESTIMATES OF TREATMENT EFFECT

The Whitehead and Lab sample yielded a total of $87\ 2 \times 2$ contingency

4. Douds and Collingwood (1978) and Collingwood and Genthner (1980) were excluded because their samples appeared to overlap those of either Collingwood et al. (1976) or Williams (1984). Similarly, Fo and O'Donnell (1975) was dropped because of overlap with O'Donnell et al. (1979). The Baer et al. (1975) report on Outward Bound was excluded because the independent variable did not involve variation in service. Beal and Duckro (1977) was dropped because the outcome seemed to be court proceedings on the offense that led to a program referral.

tables reflecting the strength and direction of the association between two levels of treatment and recidivism-nonrecidivism. Whitehead and Lab, on the other hand, tabled a single phi coefficient for each study. With our approach, distinct phi coefficients were computed when distinct samples and distinct treatments were reported in a paper (e.g., Klein et al., 1977), and rather than compare two "appropriate" styles of service, we compared each service with its respective control (e.g., Jesness, 1975; Mitchell, 1983; in the latter study we estimated that the experimental recidivists were averaging twice the number of new offenses found among control recidivists). Tests of Type of Offender \times Type of Treatment interactions were represented only incidentally in Whitehead and Lab. In our report, services to higher and lower risk cases yield separate estimates of treatment effects.

Sample 2 yielded 67 treatment-recidivism tables, 44 based on studies of juveniles and 23 based on adults. (Romig's 1976 analysis of parole supervision is entered as part of the Whitehead and Lab sample, and the analysis of months incarcerated is entered as part of sample 2). The studies and treatment comparisons are outlined in detail in the appendix (Table A1) for readers who may wish to reconstruct our analyses. Phi was employed as the measure of treatment effect because it provides a convenient summary of the direction and magnitude of the association between two binary variables, is equivalent to the Pearson product-moment coefficient, is more conservative than gamma, and was used by Whitehead and Lab.

CONTENT ANALYSIS

The potential covariates of phi estimates were coded as follows:

1. Setting: The Whitehead and Lab codes for setting were accepted uncritically: nonsystem diversion, system diversion, probation/parole/community corrections, and institutional/residential. Preliminary analyses confirmed that the effects on phi coefficients of the three different community settings were statistically indistinguishable. Hence, setting was employed as a two-level, community-residential factor in further analyses. Table A1, however, includes the elaborate code.
2. Year of publication: before the 1980s/in the 1980s.
3. Quality of research design: Studies employing random assignment were coded "stronger design." Nonrandom assignment was coded "weaker design," except when information on risk factors (e.g., prior offense or "bad attitude") allowed the computation of separate treatment comparisons for lower and higher risk cases. When risk was so controlled, the design was coded "stronger."
4. Sample of studies: Whitehead and Lab/sample 2.
5. Justice system: Juvenile system/adult system.

6. Behavioral intervention: Programs described as behavioral by the authors of an evaluation study were coded "behavioral," as were those that systematically employed behavioral techniques.⁵
7. Type of treatment: Following the principles discussed above, the four levels of type of treatment were as follows:
 - a. Criminal sanctions: This code involved variation in judicial disposition, imposed at the front end of the correctional process and not involving deliberate variation in rehabilitative service (e.g., restitution, police cautioning versus regular processing, less versus more probation, and probation versus custody).
 - b. Inappropriate correctional service: Inappropriate service included (1) service delivery to lower risk cases and/or mismatching according to a need/responsivity system, (2) nondirective relationship-dependent and/or unstructured psychodynamic counseling, (3) all milieu and group approaches with an emphasis on within-group communication and without a clear plan for gaining control over procriminal modeling and reinforcement, (4) nondirective or poorly targeted academic and vocational approaches, and (5) "scared straight."
 - c. Appropriate correctional service: Appropriate service included (1) service delivery to higher risk cases, (2) all behavioral programs (except those involving delivery of service to lower risk cases), (3) comparisons reflecting specific responsivity-treatment comparisons, and (4) nonbehavioral programs that clearly stated that criminogenic need was targeted and that structured intervention was employed.⁶

5. The interventions of Hackler and Hagan (1975) were coded as nonbehavioral. William's (1984) Dallas program was coded behavioral in our study, in line with Whitehead and Lab's coding of the Collingwood et al. (1976) report on the same program as behavioral. Both studies of restitution were coded nonbehavioral in our study (only one of which was coded nonbehavioral by Whitehead and Lab). The Ross and Fabiano behavioral skills program was coded as unspecified because it was a comparison condition for a more appropriate program.

6. Treatments admitted to the "appropriate" category by criterion "4" were appropriate according to the principles of need and responsivity (although some readers might disagree): Kelly et al. (1979) encouraged delinquents to explore alternative values and behavior patterns; the transactional program (Jesness, 1975) established individualized targets based on criminogenic need; the family counseling program of McPherson et al. (1983) targeted discipline and self-management; Bachara and Zaba (1978) focused on specific learning problems; Shore and Massimo (1979) studied very intensive, highly individualized, vocationally oriented counseling. Some difficult calls, which we ultimately coded as unspecified, included the following: Druckman's (1979) family counseling, which hinted at a nondirective client-centered approach but lacked a clear statement of same; the paraprofessional advocacy program of Seidman et al. (1980), Wade et al.'s (1977) family program, and Sowles and Gill's (1970) counseling programs all included references to both

- d. Unspecified correctional service: Unspecified service was a residual set for those comparisons involving treatments that we could not confidently label appropriate or inappropriate.

HYPOTHESES

Our first hypothesis is that Type of Treatment is the major source of variation in estimates of effect size (phi coefficients).⁷ Specifically, the contributions of Type of Treatment to the prediction of effect size will exceed the predictive contributions of year of publication, quality of design, setting, behavioral-nonbehavioral intervention, justice system (juvenile or adult), and sample of studies examined.

Our second hypothesis is that appropriate correctional service will yield an average estimate of impact on recidivism that is positive and exceeds those of criminal sanctions, unspecified service, and inappropriate service.

RESULTS AND DISCUSSION

A preliminary comparison of the two samples of studies was conducted on various control variables. The comparisons reflected an obvious concern that any systematic differences between the Whitehead and Lab sample and sample 2 be documented. Overall, apart from the inclusion of studies of adult treatment in sample 2, the two samples of studies were found to be reasonably comparable across the various potential predictors of treatment effect size explored in this paper (see row 2 of the intercorrelation matrix in Table 1).⁸

appropriate and inappropriate elements. Some "treatments" in Rausch (1983) may have involved unspecified service components, but they were assigned to the criminal sanction set in the spirit of the Rausch analysis of labelling and deterrence theory.

7. Reliability and validity in coding the type of treatment are obvious concerns. One of our ongoing research efforts involves building a psychometrically sound instrument that can be used to assess the correctional appropriateness not simply of printed program descriptions but also of ongoing programs. The psychometrics of this instrument will be the focus of future reports. For now, we have indicated in Table A1 what comparisons were assigned to what categories, and they are thereby appropriately and easily the focus of critical review.

8. The Whitehead and Lab sample ($n = 87$) and sample 2 ($n = 67$) were virtually identical in the proportion of tests falling in the three categories of treatment services: inappropriate ($^{20/87}$ vs. $^{16/67}$), unspecified ($^{16/87}$ vs. $^{16/67}$), appropriate ($^{30/87}$ vs. $^{24/67}$). The nonsignificant trend was an underrepresentation of comparisons involving criminal sanctions in sample 2 ($^{21/87}$ vs. $^9/67$, $r = .08$). Because the Whitehead and Lab sample was limited to studies of juveniles, there was an expected and substantial correlation between Justice System and Sample of Studies ($\phi = .48$, $p < 0.01$). Not as obviously deducible from the descriptions of the samples provided in the methods section, sample 2 included a statistically significant overrepresentation of institution-based treatments ($\phi = .21$, $p < .05$).

Table 1. Intercorrelation Matrix, Correlations with Phi Coefficients (N = 154), and Mean Phi Coefficients at Each Level of Each Variable

	A Type of Treatment	B Sample of Studies	C Justice System	D Year of Publication	E Quality of Design	F Setting
A.		.08	.01	-.14	.10	.11
B.			.48**	.11	.14	.21*
C.				.23*	.15	-.01
D.					-.10	-.33**
E.						-.17
Simple Unadjusted Correlation with Phi (Mean Phi = .104, SD = .234)						
	.69**	.18*	.02	.09	-.03	-.07
Unadjusted Mean Phi Coefficient (n) at Each Level of Each Variable						
1.	-.07 (30)	.07 (87)	.10 (131)	.08 (76)	.11 (81)	.11 (119)
2.	-.06 (38)	.15 (67)	.11 (23)	.13 (78)	.10 (73)	.07 (35)
3.	.13 (32)					
4.	.30 (54)					
<i>F</i> Values for Unadjusted Effects						
	45.62**	5.27*	0.49	1.33	0.11	0.74
Partial Correlation with Phi, Controlling for Other Variables						
	.72**	.15*	.02	.18*	-.07	-.16*
Adjusted Mean Phi Coefficient (n) at Each Level of Each Variable						
1.	-.08 (30)	.07 (87)	.10 (131)	.06 (76)	.11 (81)	.12 (119)
2.	-.07 (38)	.14 (67)	.11 (23)	.14 (78)	.08 (73)	.03 (35)
3.	.10 (32)					
4.	.32 (54)					
<i>F</i> Values for Adjusted Effects						
	57.15**	6.99*	0.33	9.80**	1.18	7.43**

* $p < .05$ ** $p < .01$

Note: The levels of the variables are as follows: Type of Treatment (criminal sanctions, inappropriate service, unspecified service, appropriate service), Sample of Studies (Whitehead and Lab, sample 2), Justice System (juvenile, adult), Year of Publication (before 1980, 1980s), Quality of Research Design (weaker, stronger), and Setting (community, institutional/residential).

A qualitative and nonparametric summary of findings is appended, but here the hypotheses are tested directly.

HYPOTHESIS 1: RELATIVE PREDICTIVE POTENTIAL OF TYPE OF TREATMENT

Inspection of the first column of Table 1 reveals that the correlation between Type of Treatment and phi coefficients was strong ($\eta = .69$) and, with simultaneous control introduced for each of the other variables through

analysis of covariance techniques in a multiple classification analysis, the correlation increased to .72 (Beta). The only other significant unadjusted predictor of phi coefficients was Sample of Studies (.18, unadjusted; .15, adjusted). With controls for Type of Treatment introduced, the magnitude of correlation with phi coefficients increased to significant levels for Year of Publication (from .09 to .18) and for Setting (from $-.07$ to $-.16$).

Comparisons from sample 2, recency of publication and community-based treatment, were each associated with relatively positive effects of treatment. These trends, however, were overwhelmed by Type of Treatment. In a stepwise multiple regression, the only variables contributing significantly ($p < .05$) to variation in phi estimates were Type of Treatment (beta = .69) and Year of Publication (beta = .19), $F(2/151) = 68.01$, $p < .000$, adjusted R square = .47. In summary, our first hypothesis was strongly supported: Type of Treatment was clearly the strongest of the correlates of effect size sampled in this study.

HYPOTHESIS 2: THE IMPORTANCE OF APPROPRIATE CORRECTIONAL SERVICE

As described above, the main effect of Type of Treatment on phi estimates was strong and positive, with or without adjustment for control variables. Scheffe tests confirmed that the mean phi coefficient for appropriate correctional service (.30, $n = 54$) was significantly ($p < 0.05$) greater than that for criminal sanctions ($-.07$, $n = 30$), inappropriate service ($-.06$, $n = 38$), and unspecified service (.13, $n = 32$). In addition, Scheffe tests revealed that the average effect of unspecified correctional service significantly exceeded the mean phi coefficients for criminal sanctions and inappropriate service.

Mean phi coefficients for each of the four types of treatment are presented in Table 2 at each of the two levels of the various control variables. Inspection reveals a robust correlation between Type of Treatment and effects on recidivism at each level of Sample of Studies, Justice System, Year of Publication, Design, and Setting.

The only variable to interact significantly ($p < 0.05$) with Type of Treatment was Year of Publication. It appears that criminal sanctions yielded more negative phi estimates in the earlier literature than in the more recent literature ($-.16$ versus $-.02$, $F[1/28] = 8.98$, $p < .006$). This reflects a greater representation of residential studies in the earlier years (the negative implications of residential programs will be discussed below). More interestingly, studies of appropriate correctional treatment in the 1980s yielded a much higher mean phi estimate than did earlier studies of appropriate treatment (.40 versus .24, $F[1/52] = 8.40$, $p < .005$). Most likely, this reflects three trends. First, the earlier studies included what are now recognized to be unsophisticated applications of token economy systems (see Ross and

McKay, 1976). Second, studies of the 1980s paid greater attention to cognitive variables (Ross and Fabiano, 1985). Third, the positive effects of short-term behavioral family counseling have been replicated in the 1980s (Gordon et al., 1988). In summary, Hypothesis 2 was supported to a stronger degree than was initially anticipated: Both appropriate and unspecified correctional services were significantly more effective in reducing recidivism than were criminal sanctions and inappropriate service.

NOTE ON BEHAVIORAL INTERVENTION

The use of behavioral methods was a major element in the coding of appropriateness according to the principle of responsivity. Not surprisingly, in view of our coding rules, 95% ($38/41$) of the behavioral treatments were coded as appropriate treatment and 70% ($38/54$) of the appropriate treatments were behavioral. Thus, the correlation between Behavioral Intervention and Type

Table 2. The Effect of Type of Treatment on Recidivism at Each Level of the Control Variables: Mean Phi Coefficients (N)

	Criminal Sanctions	Correctional Service		
		Inapp.	Unspec.	Appropriate
Sample of Studies				
Whitehead and Lab	-.04 (21)	-.11 (20)	.09 (16)	.24 (30)
Sample 2	-.13 (9)	-.02 (18)	.17 (16)	.37 (24)
Justice System				
Juvenile	-.06 (26)	-.07 (31)	.13 (29)	.29 (45)
Adult	-.12 (4)	-.03 (7)	.13 (3)	.34 (9)
Year of Publication				
Before the 1980s	-.16 (10)	-.09 (22)	.17 (11)	.24 (33)
1980s	-.02 (20)	-.03 (16)	.11 (21)	.40 (21)
Quality of Research Design				
Weaker	-.07 (21)	-.04 (10)	.15 (18)	.32 (26)
Stronger	-.07 (9)	-.08 (22)	.11 (14)	.29 (28)
Setting				
Community	-.05 (24)	-.14 (31)	.12 (27)	.35 (37)
Institution/Res.	-.14 (6)	-.15 (7)	.21 (5)	.20 (17)
Behavioral Intervention				
No	-.07 (30)	-.06 (36)	.13 (31)	.27 (16)
Yes	—	-.09 (2)	.23 (1)	.31 (38)
Overall Mean Phi	-.07 (30)	-.06 (38)	.13 (32)	.30 (54)
S.D.	.14	.15	.16	.19
Mean Phi Adjusted for Other Variables	-.08 (30)	-.07 (38)	.10 (32)	.32 (54)

of Treatment was substantial ($r = .62$). As expected, Behavioral Intervention, on its own, yielded a significantly greater mean phi coefficient than did non-behavioral treatment. The mean phi coefficients were .29 (SD = .23, $n = 41$) and .04 (SD = .20, $n = 113$) for behavioral and nonbehavioral interventions, respectively ($F [1/152] = 46.09, p < .000, \text{Eta} = .48$). Once controls were introduced for Type of Treatment, however, the contribution of Behavioral Intervention was reduced to nonsignificant levels, $F (1/151) < 1.00, \text{Beta} = .07$. It appears, then, that use of behavioral methods contributes to the reduction of recidivism, but those contributions are subsumed by the broader implications of risk, need, and responsivity as represented in our Type of Treatment variable.

NOTE ON RESIDENTIAL PROGRAMMING

The minor but statistically significant adjusted main effect of setting is displayed in column six of Table 1. This trend should not be overemphasized, but the relatively weak performance of appropriate correctional service in residential facilities is notable from Table 2 (mean phi estimate of .20 compared with .35 for treatment within community settings, $F [1/52] = 5.89, p < .02$). In addition, inappropriate service performed particularly poorly in residential settings compared with community settings ($-.15$ versus $-.04, F [1/36] = 3.74, p < .06$). Thus, it seems that institutions and residential settings may dampen the positive effects of appropriate service while augmenting the negative impact of inappropriate service. This admittedly tentative finding does not suggest that appropriate correctional services should not be applied in institutional and residential settings. Recall that appropriate service was more effective than inappropriate service in all settings.

CONCLUSIONS

The meta-analysis has revealed considerable order in estimates of the magnitude of the impact of treatment upon recidivism. As predicted, the major source of variation in effects on recidivism was the extent to which service was appropriate according to the principles of risk, need, and responsivity. Appropriate correctional service appears to work better than criminal sanctions not involving rehabilitative service and better than services less consistent with our a priori principles of effective rehabilitation. This review has convinced us that the positive trends that we and others detected in the literature of the 1960s and early 1970s were indeed worthy of serious application and evaluation. There is a reasonably solid clinical and research basis for the political reaffirmation of rehabilitation (Cullen and Gilbert, 1982).

The importance of clinical and theoretical relevance in programming and in meta-analysis has been demonstrated—the sanction and treatment services should be differentiated, and the action in regard to recidivism appears to

reside in appropriate treatment. Much, however, remains to be done. We look forward to critiques and revisions of the principles of risk, need, and responsivity as stated and applied herein. What comparisons were assigned to what analytic categories is described in our report and is thereby easily and appropriately the focus of critical review (see note 7). Reserved for future reports are the many issues surrounding therapeutic integrity (Gendreau and Ross, 1979), the measurement of recidivism (Andrews, 1983), and methodological issues such as sample size (Lipsey, 1989). Similarly, we anticipate exploring in detail the value of alternatives to ordinary least squares analyses (for now, nonparametric tests of Type of Treatment are appended). Gender effects and the treatment of sex offenders, substance abusers, and inmates of long-term institutions require detailed analyses. Toward these ends, our meta-analytic data base is being extended. Our focus here, however, remains on type of service and effect size.

Of immediate concern is the meaning of an average phi coefficient of .30 for comparisons involving appropriate correctional service. First, until convinced otherwise, we will assume that an average phi of .30 is more positive, clinically and socially, than the mean effects of the alternatives of sanctioning without regard for service or servicing without regard to the principles of effective correctional service. Casual review of recidivism rates will reveal that, on average, appropriate treatment cut recidivism rates by about 50% (in fact, the mean reduction was 53.06%, $SD = 26.49$). Thus, we do not think that the positive effects are "minimal". Second, the correlation between effect size estimates and type of treatment approached .70. Correlations of this magnitude are unlikely to reflect "lucky outliers" (Greenwood, 1988), although more systematic sources of error may indeed inflate correlation coefficients. Third, issues surrounding the assessment of the clinical and social significance of diverse measures of effect size are indeed worthy of ongoing research. Future reports on our expanding data bank will compare various estimates of effect size, including some direct estimates of clinical/social significance. For now, we are interested in discovering ethical routes to strengthened treatment effects, but we are not talking about magical cures.

Critics of rehabilitation are correct when they note that the average correlation between treatment and recidivism is not 1.00. At the same time, critics might be asked to report on the variation that their "preferred" variable shares with recidivism. For example, if their preferred variable is social class, they may be reminded that some reviewers have estimated that the average correlation between class and crime is about $-.09$ (Tittle et al., 1978). If their preferred approach is incapacitation or community crime prevention, they may be reminded of the minimal effects so far reported for these strategies (Rosenbaum, 1988; Visher, 1987). Critics, be they supporters of social class or incapacitation, likely will respond with examples of particular studies that yielded high correlations with indicators of crime. We remind them that

the largest correlations are no better estimates of the average effect than are the least favorable estimates. We also remind them that the positive evidence regarding appropriate rehabilitative service comes not from cross-sectional research—the typical research strategy of critics of rehabilitation—but from deliberate and socially sanctioned approximations of truly experimental ideals. Finally, we remind the critics that one can be interested in the effects of class, punishment, and prevention programs on individual and aggregated crime rates while maintaining multiple interests and without letting one interest justify dismissal of the value of another.

This meta-analysis has done more than uncover evidence that supported our a priori biases regarding the importance of appropriate correctional service. The finding that the effects of inappropriate service appeared to be particularly negative in residential settings while the positive effects of appropriate service were attenuated was something of a surprise. While sensitive to the difficulties of working with antisocial groups, we did not predict this incidental affirmation of a widely shared preference for community over residential programming. Institutions and group homes, however, remain important components of correctional systems and hence active but thoughtful service is indicated. The literature should be carefully scrutinized in order to avoid inappropriate service, and follow-up services in the community may be necessary in order to maximize effectiveness. Finally, the suppressive impact of residential programming suggests that the negative effects of custody are better established than we anticipated.

The effect of the quality of the research design on estimates of effect size was relatively minor. Even if some design problems do inflate effect size estimates (Davidson et al., 1984; Lipsey, 1989), the interesting finding was that comparisons involving more and less rigorous research designs agreed as to what types of treatment were most effective. Program managers and frontline clinicians who find truly randomized groups to be practically or ethically impossible may consider conducting an evaluation that approximates the ideals of a true experiment. In particular, we strongly endorse the use of designs that introduce controls for the preservice risk levels of clients and that actually report on risk \times service interactions. In addition, even evaluations that rely upon comparisons of clients who complete or do not complete treatment may be valuable.

Finally, the number of evaluative studies of correctional service should increase dramatically over the next decade. Although millions of young people were processed by juvenile justice systems during the past decade, the total number of papers in the Whitehead and Lab (1989) set that involved systematic study of appropriate service was 21. Were it not for behavioral psychologists, the number of papers involving appropriate service would have been nine. From a positive perspective, there is renewed interest, vigor, and sensitivity in the study of the psychology of criminal conduct (Andrews and

Wormith, 1989; Loeber and Stouthamer-Loeber, 1987; Wilson and Herrnstein, 1985) and of correctional service and prevention (e.g., Andrews et al., 1990; Cullen and Gendreau, 1988; Currie, 1989; Gendreau and Ross, 1987). There are solid reasons to focus in ethical and humane ways on the client and the quality of service delivered within just dispositions.

REFERENCES

- Adams, R. and H.J. Vetter
1982 Social structure and psychodrama outcome: A ten-year follow-up. *Journal of Offender Counseling, Services, and Rehabilitation* 6:111-119.
- Agee, V.L.
1986 Institutional treatment programs for the violent juvenile. In S. Apter and A. Goldstein (eds.), *Youth Violence: Program and Prospects*. Elmsford, N.Y.: Pergamon.
- Akers, R.L. and J.K. Cochran
1985 Adolescent marijuana use: A test of three theories of deviant behavior. *Deviant Behavior* 3:323-346.
- Alexander, J.F., B. Cole, R.S. Schiavo, and B.V. Parsons
1976 Systems-behavioral intervention with families of delinquents: Therapist characteristics, family behavior, and outcome. *Journal of Consulting and Clinical Psychology* 44:556-664.
- Allen, F.A.
1981 *The Decline of the Rehabilitative Ideal: Penal Policy and Social Purpose*. New Haven: Yale University Press.
- Andrews, D.A.
1980 Some experimental investigations of the principles of differential association through deliberate manipulations of the structure of service systems. *American Sociological Review* 45:448-462.
1983 The assessment of outcome in correctional samples. In M.L. Lambert, E.R. Christensen, and S.S. DeJulio (eds.), *The Measurement of Psychotherapy Outcome in Research and Evaluation*. New York: John Wiley & Sons.
1990 Some criminological sources of antirehabilitation bias in the Report of the Canadian Sentencing Commission. *Canadian Journal of Criminology*. Forthcoming.
- Andrews, D.A. and J.J. Kiessling
1980 Program structure and effective correctional practices: A summary of the CaVIC research. In R.R. Ross and P. Gendreau (eds.), *Effective Correctional Treatment*. Toronto: Butterworth.
- Andrews, D.A. and J.S. Wormith
1989 Personality and crime: Knowledge destruction and construction in criminology. *Justice Quarterly* 6:289-309.
- Andrews, D.A., J.J. Kiessling, D. Robinson, and S. Mickus
1986 The risk principle of case classification: An outcome evaluation with young adult probationers. *Canadian Journal of Criminology* 28:377-396.

- Andrews, D.A., J. Bonta, and R.D. Hoge
1990 Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior* 17:19-52.
- Bachara, G.H. and J.N. Zaba
1978 Learning disabilities and juvenile delinquency. *Journal of Learning Disabilities* 11:242-246.
- Baer, D.J., P.J. Jacobs, and F.E. Carr
1975 Instructors' ratings of delinquents after Outward Bound survival training and their subsequent recidivism. *Psychological Reports* 36:547-553.
- Bailey, W.C.
1966 Correctional outcome: An evaluation of 100 reports. *Journal of Criminal Law, Criminology and Police Science* 57:153-160.
- Baird, S.C., R.C. Heinz, B.J. Bemus
1979 Project Report #14: A Two Year Follow-up. Bureau of Community Corrections. Wisconsin: Department of Health and Social Services.
- Barkwell, L.J.
1976 Differential treatment of juveniles on probation: An evaluative study. *Canadian Journal of Criminology and Corrections* 18:363-378.
- Barton, C., J.F. Alexander, H. Waldron, C.W. Turner, and J. Warburton
1985 Generalizing treatment effects of functional family therapy: Three replications. *The American Journal of Family Therapy* 13:16-26.
- Basta, J.M. and W.S. Davidson
1988 Treatment of juvenile offenders: Study outcomes since 1980. *Behavioral Sciences and the Law* 6:355-384.
- Beal, D. and P. Duckro
1977 Family counseling as an alternative to legal action for the juvenile status offender. *Journal of Marriage and Family Counseling* 3:77-81.
- Berman, J.J.
1978 An experiment in parole supervision. *Evaluation Quarterly* 2:71-90.
- Buckner, J.C. and M. Chesney-Lind
1983 Dramatic cures for juvenile crime: An evaluation of a prisoner-run delinquency prevention program. *Criminal Justice and Behavior* 10:227-247.
- Buchler, R.E., G.R. Patterson, and J.M. Furniss
1966 The reinforcement of behavior in institutional settings. *Behavioral Research and Therapy* 4:157-167.
- Byles, J.A.
1981 Evaluation of an attendance center program for male juvenile probationers. *Canadian Journal of Criminology* 23:343-355.
- Byles, J.A. and A. Maurice
1979 The Juvenile Services Project: An experiment in delinquency control. *Canadian Journal of Criminology* 21:155-165.
- Clarke, R.V.G. and D.B. Cornish
1978 The effectiveness of residential treatment for delinquents. In L.A. Hersov, M. Berger, and D. Shaffer (eds.), *Aggression and Anti-social Behavior in Childhood and Adolescence*. Oxford: Pergamon.

CORRECTIONAL TREATMENT META-ANALYSIS 389

- Collingwood, T.R. and R.W. Genthner
 1980 Skills trainings as treatment for juvenile delinquents. *Professional Psychology* 11:591-598.
- Collingwood, T.R., A.F. Douds, and H. Williams
 1976 Juvenile diversion: The Dallas Police Department Youth Services Program. *Federal Probation* 40:23-27.
- Craft, M., G. Stephenson, and C. Granger
 1966 A controlled trial of authoritarian and self-governing regimes with adolescent psychopaths. *American Journal of Orthopsychiatry* 34:543-554.
- Cullen, F.T. and P. Gendreau
 1989 The effectiveness of correctional rehabilitation. In L. Goodstein and D.L. MacKenzie (eds.), *The American Prison: Issues in Research Policy*. New York: Plenum.
- Cullen, F.T. and K.E. Gilbert
 1982 *Reaffirming Rehabilitation*. Cincinnati: Anderson.
- Cullen, F.T., J.B. Cullen and J.F. Woznick
 1988 Is rehabilitation dead? The myth of the punitive public. *Journal of Criminal Justice* 16:303-317.
- Currie, E.
 1989 Confronting crime: Looking toward the twenty-first century. *Justice Quarterly* 6:5-25.
- Davidson, W.S. and T.R. Wolfred
 1977 Evaluation of a community-based behavior modification program for prevention of delinquency. *Community Mental Health Journal* 13:296-306.
- Davidson, W.S., L. Gottschalk, L. Gensheimer, and J. Mayer
 1984 *Interventions with Juvenile Delinquents: A Meta-analysis of Treatment Efficacy*. Washington, D.C.: National Institute of Juvenile Justice and Delinquency Prevention.
- Davidson, W.S., R. Redner, C. Blakely, C. Mitchell, and J. Emshoff
 1987 Diversion of juvenile offenders: An experimental comparison. *Journal of Consulting and Clinical Psychology* 55:68-75.
- Douds, A.F. and T.R. Collingwood
 1978 Management by objectives: A successful application. *Child Welfare* 57:181-185.
- Druckman, J.M.
 1979 A family-oriented policy and treatment program for female juvenile status offenders. *Journal of Marriage and the Family* 41:627-636.
- Dutton, D.G.
 1986 The outcome of court-mandated treatment for wife assault: A quasi-experimental evaluation. *Violence and Victims* 1:163-175.
- Empey, L.T. and M.L. Erickson
 1972 *The Provo Experiment: Evaluating Community Control of Delinquency*. Lexington, Mass.: Lexington Books.
- Farrington, D.P. and T. Bennett
 1981 Police cautioning of juveniles in London. *British Journal of Criminology* 21:123-135.

- Faust, D.
1965 Group counseling of juveniles by staff without professional training in group work. *Crime and Delinquency* 11:349-354.
- Finckenauer, J.O.
1982 *Scared Straight! and the Panacea Phenomenon*. Englewood Cliffs, N.J.: Prentice-Hall.
- Fo, W.S.O. and C.R. O'Donnell
1975 The buddy system: Effect of community intervention on delinquent offenses. *Behavior Therapy* 6:522-524.
- Freud, S.
1953 *A General Introduction to Psychoanalysis*. Reprint ed. New York: PermaBooks.
- Garrett, C.J.
1985 Effects of residential treatment of adjudicated delinquents: A meta-analysis. *Journal of Research in Crime and Delinquency* 22:287-308.
- Geismar, L.L. and K.M. Wood
1985 *Family and Delinquency: Resocializing the Young Offender*. New York: Human Sciences Press.
- Gendreau, P. and R.R. Ross
1979 Effectiveness of correctional treatment: Bibliotherapy for cynics. *Crime and Delinquency* 25:463-489.
1981 Correctional potency: Treatment and deterrence on trial. In R. Roesch and R.R. Corrado (eds.), *Evaluation and Criminal Justice Policy*. Beverly Hills, Calif.: Sage.
1987 Revivification of rehabilitation: Evidence from the 1980s. *Justice Quarterly* 4:349-408.
- Gensheimer, L.K., J.P. Mayer, R. Gottschalk, and W.S. Davidson
1986 Diverting youth from the juvenile justice system: A meta-analysis of intervention efficacy. In S.J. Apter and A. Goldstein (eds.), *Youth Violence: Programs and Prospects*. Elmsford, N.Y.: Pergamon.
- Gilbert, G.R.
1977 Alternate routes: A diversion project in the juvenile justice system. *Evaluation Quarterly* 1:301-318.
- Glaser, D.
1974 Remedies for the key deficiency in criminal justice evaluation research. *Journal of Research in Crime and Delinquency* 11:144-153.
- Glueck, S. and E.T. Glueck
1950 *Unraveling Juvenile Delinquency*. Cambridge, Mass.: Harvard University Press.
- Gordon, D.A., J. Arbuthnot, K.E. Gustafson, and P. McGreen
1988 Home-based behavioral systems family therapy with disadvantaged juvenile delinquents. Unpublished paper, Ohio University.
- Grant, J.D.
1965 Delinquency treatment in an institutional setting. In H.C. Quay (ed.), *Juvenile Delinquency: Research and Theory*. Princeton, N.J.: Van Nostrand.

CORRECTIONAL TREATMENT META-ANALYSIS 391

- Grant, J.D. and M.Q. Grant
1959 A group dynamics approach to the treatment of nonconformists in the navy. *Annals of the American Academy of Political and Social Science* 322:126-135.
- Greenwood, P.W.
1988 *The Role of Planned Interventions in Studying the Desistance of Criminal Behavior in Longitudinal Study*. Santa Monica, Calif.: Rand.
- Greenwood, P.W. and F.E. Zimring
1985 *One More Chance: The Pursuit of Promising Intervention Strategies for Chronic Juvenile Offenders*. Santa Monica, Calif.: Rand.
- Gruher, M.
1979 Family counseling and the status offender. *Juvenile and Family Court Journal* 30:23-27.
- Hackler, J.C. and J.L. Hagan
1975 Work and teaching machines as delinquency prevention tools: A four-year follow-up. *Social Service Review* 49:92-106.
- Harris, P.W.
1988 The interpersonal maturity level classification system: I-level. *Criminal Justice and Behavior* 15:58-77.
- Hoge, R.D. and D.A. Andrews
1986 A model for conceptualizing interventions in social service. *Canadian Psychology* 27:332-341.
- Horowitz, A. and M. Wasserman
1979 The effect of social control on delinquent behavior: A longitudinal test. *Sociological Focus* 12:53-70.
- Hunt, G.M. and N.H. Azrin
1973 A community-reinforcement approach to alcoholism. *Behavior Research and Therapy* 11:91-104.
- Izzo, R.L. and R.R. Ross
1990 Meta-analysis of rehabilitation programs for juvenile delinquents. *Criminal Justice and Behavior* 17:134-142.
- Jackson, P.C.
1983 Some effects of parole supervision on recidivism. *British Journal of Criminology* 23:17-34.
- Jesness, C.F.
1975 Comparative effectiveness of behavior modification and transactional analysis programs for delinquents. *Journal of Consulting and Clinical Psychology* 43:758-779.
- Jessor, R. and S.L. Jessor
1977 *Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth*. New York: Academic Press.
- Johnson, R.E.
1979 *Juvenile Delinquency and Its Origins: An Integrative Theoretical Approach*. New York: Cambridge University Press.

- Johnson, B.D. and R.T. Goldberg
 1983 Vocational and social rehabilitation of delinquents. *Journal of Offender Counseling, Services, and Rehabilitation* 6:43-60.
- Kassebaum, G., D. Ward, and D. Wilner
 1971 *Prison Treatment and Parole Survival: An Empirical Assessment*. New York: John Wiley & Sons.
- Kelley, T.M., A.K. Havva, and R.A. Blak
 1979 The effectiveness of college student companion therapists with predelinquent youths. *Journal of Police Science and Administration* 7:186-195.
- Kirby, B.C.
 1954 Measuring effects of treatment of criminals and delinquents. *Sociology and Social Research* 38:368-374.
- Kirigin, K.A., C.J. Braukman, J.D. Atwater, and M.W. Montrose
 1982 An evaluation of Teaching Family (Achievement Place) Group Homes for juvenile offenders. *Journal of Applied Behavior Analysis* 15:1-16.
- Klein, M.W.
 1971 *Street Gangs and Street Workers*. Englewood Cliffs, N.J.: Prentice-Hall.
 1986 Labeling theory and delinquency policy: An experimental test. *Criminal Justice and Behavior* 13:47-79.
- Klein, N.C., J.F. Alexander, and B.V. Parsons
 1977 Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology* 3:469-474.
- Kratcoski, P.C. and L.D. Kratcoski
 1982 The Phoenix Program: An educational alternative for delinquent youths. *Juvenile and Family Court Journal* 33:17-23.
- Kraus, J.
 1978 Remand in custody as a deterrent in juvenile jurisdiction. *British Journal of Criminology* 18:17-23.
 1981 Police caution of juvenile offenders: A research note. *Australian and New Zealand Journal of Criminology* 14:91-94.
- Lab, S.P. and J.T. Whitehead
 1988 An analysis of juvenile correctional treatment. *Crime and Delinquency* 34:60-83.
- Leschied, A.W., G.W. Austin, and P.G. Jaffe
 1988 Impact of the Young Offenders Act on recidivism rates of special needs youth: Clinical and policy implications. *Canadian Journal of Behavioural Science* 20:322-331.
- Lewis, R.V.
 1983 Scared straight—California style: Evaluation of the San Quentin Program. *Criminal Justice and Behavior* 10:209-226.
- Link, B.G., F.T. Cullen, E. Struening, P.E. Shrout, and B.P. Dohrenwend
 1989 A modified labeling theory approach to mental illness. *American Sociological Review* 54:400-423.

CORRECTIONAL TREATMENT META-ANALYSIS 393

- Lipsey, M.W.
1989 The efficacy of intervention for juvenile delinquency: Results from 400 studies. Paper presented at the 41st annual meeting of the American Society of Criminology, Reno, Nev.
- Lipsey, M.W., D.S. Cordray, and D.E. Berger
1981 Evaluation of a juvenile diversion program using multiple lines of evidence. *Evaluation Review* 5:283-306.
- Loeber, R. and M. Stouthamer-Loeber
1987 Prediction. In H.C. Quay (ed.), *Handbook of Juvenile Delinquency*. New York: John Wiley & Sons.
- Logan, C.H.
1972 Evaluation research in crime and delinquency: A reappraisal. *Journal of Criminal Law, Criminology and Police Science* 63:378-387.
- Martinson, R.
1974 What works? Questions and answers about prison reform. *The Public Interest* 35:22-54.
1979 New findings, new views: A note of caution regarding prison reform. *Hofstra Law Review* 7:243-258.
- Maskin, M.B.
1976 The differential impact of work vs communication-oriented juvenile correction programs upon recidivism rates in delinquent males. *Journal of Clinical Psychology* 32:432-433.
- Mayer, J.P., L.K. Gensheimer, W.S. Davidson, and R. Gottschalk
1986 Social learning treatment within juvenile justice: A meta-analysis of impact in the natural environment. In S.J. Apter and A. Goldstein (eds.), *Youth Violence: Programs and Prospects*. Elmsford, N.Y.: Pergamon.
- McPherson, S.J., L.E. McDonald, and C.W. Ryder
1983 Intensive counseling with families of juvenile offenders. *Juvenile and Family Court Journal* 34:27-33.
- Mitchell, C.M.
1983 The dissemination of a social intervention: Process and effectiveness of two types of paraprofessional change agents. *American Journal of Community Psychology* 11:723-739.
- Mott, J.
1983 Police decisions for dealing with juvenile offenders. *British Journal of Criminology* 23:249-262.
- Murphy, B.C.
1972 A Test of the Effectiveness of an Experimental Treatment Program for Delinquent Opiate Addicts. Ottawa: Information Canada.
- O'Donnell, C.R., R. Lydgate, and W.S.O. Fo
1979 The buddy system: Review and follow-up. *Child Behavior Therapy* 1:161-169.
- Ostrum, T.M., C.M. Steele, L.K. Resenblood, and H.L. Mirels
1971 Modification of delinquent behavior. *Journal of Applied Social Psychology* 1:118-136.

- Palmer, T.
1975 Martinson revisited. *Journal of Research in Crime and Delinquency* 12:133-152.
1983 The effectiveness issue today: An overview. *Federal Probation* 46:3-10.
- Palmer, T. and R.V. Lewis
1980 A differentiated approach to juvenile diversion. *Journal of Research in Crime and Delinquency* 17:209-227.
- Persons, R.
1967 Relationship between psychotherapy with institutionalized boys and subsequent adjustments. *Journal of Consulting Psychology* 31:137-141.
- Petersilia, J., S. Turner, and J. Peterson
1986 Prison versus Probation in California: Implications for Crime and Offender Recidivism. Santa Monica, Calif.: Rand.
- Phillips, E.L., E.A. Phillips, D.L. Fixen, and M.W. Wolf
1973 Achievement Place: Behavior shaping works for delinquents. *Psychology Today* 6:75-79.
- Quay, H.C. and C.T. Love
1977 The effect of a juvenile diversion program on rearrests. *Criminal Justice and Behavior* 4:377-396.
- Rausch, S.
1983 Court processing vs. diversion of status offenders: A test of deterrence and labeling theories. *Journal of Research in Crime and Delinquency* 20:39-54.
- Redfering, D.L.
1973 Durability of effects of group counseling with institutionalized females. *Journal of Abnormal Psychology* 82:85-86.
- Regoli, R., E. Wilderman, and M. Pogrebin
1985 Using an alternative evaluation measure for assessing juvenile diversion. *Children and Youth Services Review* 7:21-38.
- Reitsma-Street, M. and A.W. Leschied
1988 The conceptual level matching model in corrections. *Criminal Justice and Behavior* 15:92-108.
- Romig, D.A.
1976 Length of institutionalization, treatment program completion, and recidivism among delinquent adolescent males. *Criminal Justice Review* 1:115-119.
- Rosenbaum, D.P.
1988 Community crime prevention: A review and synthesis of the literature. *Justice Quarterly* 4:513-544.
- Ross, R.R. and E.A. Fabiano
1985 Time to Think: A Cognitive Model of Delinquency Prevention and Offender Rehabilitation. Johnson City, Tenn.: Institute of Social Sciences and Arts.
- Ross, R.R. and H.B. McKay
1976 A study of institutional treatment programs. *International Journal of Offender Therapy and Comparative Criminology* 21:165-173.
- Ross, R.R., E.A. Fabiano, and C.D. Ewles
1988 Reasoning and rehabilitation. *International Journal of Offender Therapy and Comparative Criminology* 32:29-35.

CORRECTIONAL TREATMENT META-ANALYSIS 395

- Rothman, D.J.
1980 *Conscience and Convenience: The Asylum and Its Alternatives in Progressive America*. Boston: Little, Brown.
- Sarason, I.G. and V.J. Ganzer
1973 Modeling and group discussions in the rehabilitation of juvenile delinquents. *Journal of Counseling Psychology* 20:442-449.
- Schneider, A.L. and P.R. Schneider
1984 A comparison of programmatic and ad hoc restitution in juvenile court. *Justice Quarterly* 1:529-547.
- Schur, E.M.
1973 *Radical Nonintervention: Rethinking the Delinquency Problem*. Englewood Cliffs, N.J.: Prentice-Hall.
- Sechrest, L., S.O. White, and E.D. Brown
1979 *The Rehabilitation of Criminal Offenders: Problem and Prospects*. Washington, D.C.: National Academy Press.
- Seidman, E., J. Rappaport, and W.S. Davidson
1980 Adolescents in legal jeopardy: Initial success and replication of an alternative to the criminal justice system. In R.R. Ross and P. Gendreau (eds.), *Effective Correctional Treatment*. Toronto: Butterworth.
- Shichor, D. and A. Binder
1982 Community restitution for juveniles: An approach and preliminary investigation. *Criminal Justice Review* 7:46-50.
- Shore, M.F. and J.L. Massimo
1979 Fifteen years after treatment: A follow-up study of comprehensive vocationally-oriented psychotherapy. *American Journal of Orthopsychiatry* 49:240-245.
- Shorts, I.D.
1986 Delinquency by association. *British Journal of Criminology* 26:156-163.
- Smith, D.A. and P.R. Gartin
1989 Specifying specific deterrence. *American Sociological Review* 54:94-105.
- Sorenson, J.L.
1978 Outcome evaluation of a referral system for juvenile offenders. *American Journal of Community Psychology* 6:381-388.
- Sowles, R.C. and J. Gill
1970 Institutional and community adjustment of delinquents following counseling. *Journal of Consulting and Clinical Psychology* 34:398-402.
- Stephenson, R.M. and F.R. Scarpitti
1974 *Group Interaction as Therapy: The Use of the Small Group in Corrections*. Westport, Conn.: Greenwood Press.
- Stringfield, N.
1977 The impact of family counseling in resocializing adolescent offenders within a positive peer treatment milieu. *Offender Rehabilitation* 1:349-360.
- Stuart, R.B., S. Jayaratne, and T. Tripodi
1976 Changing adolescent deviant behaviour through reprogramming the behaviour of parents and teachers: An experimental evaluation. *Canadian Journal of Behavioural Science* 8:132-143.

- Tittle, C.R., W.J. Villimez, and D.A. Smith
1978 The myth of social class and criminality: An empirical assessment of the empirical evidence. *American Sociological Review* 43:643-656.
- Viano, E.C.
1976 Growing up in an affluent society: Delinquency and recidivism in suburban America. *Journal of Criminal Justice* 3:223-236.
- Vinglis, E., E. Adlap, and L. Chung
1982 The Oshawa Impaired Drivers Programme: An evaluation of a rehabilitation programme. *Canadian Journal of Criminology* 23:93-102.
- Visher, C.A.
1987 Incapacitation and crime control: Does a "lock 'em up" strategy reduce crime? *Justice Quarterly* 4:513-544.
- Vito, G.V. and H.E. Allen
1981 Shock probation in Ohio: A comparison of outcomes. *International Journal of Offender Therapy and Comparative Criminology* 25:70-76.
- Wade, T.C., T.L. Morton, J.E. Lind, and N.R. Ferris
1977 A family crisis intervention approach to diversion from the juvenile justice system. *Juvenile Justice Journal* 28:43-51.
- Walker, S.
1989 Sense and Nonsense about Crime: A Policy Guide. Pacific Grove, Calif.: Brooks/Cole.
- Walsh, A.
1985 An evaluation of the effects of adult basic education on rearrest rates among probationers. *Journal of Offender Counseling, Services and Rehabilitation* 9:69-76.
- Walter, T.L. and C.M. Mills
1980 A behavioral-employment intervention program for reducing juvenile delinquency. In R.R. Ross and P. Gendreau (eds.), *Effective Correctional Treatment*. Toronto: Butterworth.
- Warren, M.Q.
1969 The case for differential treatment of delinquents. *Annals of the American Academy of Political and Social Science* 381:47-59.
- Wellford, C.
1975 Labelling theory and criminology: An assessment. *Social Problems* 22:332-345.
- Whitaker, J.M. and L.J. Severy
1984 Service accountability and recidivism for diverted youth: A client- and service-comparison analysis. *Criminal Justice and Behavior* 11:47-74.
- Whitehead, J.T. and S.P. Lab
1989 A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime and Delinquency* 26:276-295.
- Williams, L.
1984 A police diversion alternative for juvenile offenders. *Police Chief* (Feb), 54-56.

CORRECTIONAL TREATMENT META-ANALYSIS 397

- Willman, M.T. and J.R. Snortum
1982 A police program for employment and youth gang members. *International Journal of Offender Therapy and Comparative Criminology* 26:207-214.
- Wilson, J.Q. and R.J. Herrnstein
1985 *Crime and Human Nature*. New York: Simon & Schuster.
- Winterdyk, J. and R. Roesch
1982 A Wilderness Experimental Program as an alternative for probationers. *Canadian Journal of Criminology* 23:39-49.
- Wormith, J.S.
1984 Attitude and behavior change of correctional clientele: A three year follow-up. *Criminology* 22:595-618.
- Wright, W.E. and M.C. Dixon
1977 Community prevention and treatment of juvenile delinquency. *Journal of Research in Crime and Delinquency* 14:35-67.
- Zeisel, H.
1982 Disagreement over the evaluation of a controlled experiment. *American Journal of Sociology* 88:378-389.

D.A. Andrews is a professor of psychology at Carleton University and co-director of the Laboratory for Research on Assessment and Evaluation in the Human Services. His research interests include the assessment of risk, need, and responsivity factors among clients of correctional, youth, family, and mental health agencies, as well as the analysis and evaluation of effective treatment services. Additionally, he is interested in the social psychology of criminological knowledge.

Ivan Zinger, a recent graduate from Carleton University, is currently a student of law at the University of Ottawa. His research interests include the clinical/community psychology of crime and the social psychology of law and justice.

Robert D. Hoge is a professor of psychology at Carleton University and co-director of the Laboratory for Research on Assessment and Evaluation in the Human Services. His major areas of research interest concern problems in psychological assessment, with a particular concern for educational and criminal justice settings.

James Bonta is Chief Psychologist at the Ottawa-Carleton Detention Centre, clinical associate professor at the University of Ottawa, and research adjunct professor at Carleton University. His research interests include the assessment of offenders, the effectiveness of correctional rehabilitation, and the effects of incarceration.

Paul Gendreau is a professor of psychology at the University of New Brunswick (Saint John), consultant and past-director of research at Centracare Saint John, and consultant to Saint John Police Services. His research interests include assessment, treatment, and consultation.

Francis T. Cullen is a professor of criminal justice at the University of Cincinnati. His research interests include theories of crime and deviance, white-collar crime, correctional policy, and the attitudes of the public and professionals toward deviance and official processing.

APPENDIX: NONPARAMETRIC SUMMARY AND OVERVIEW OF THE STUDIES

Descriptions of the 154 explorations of treatment and recidivism are presented in Table A1. The major subheadings in the table identify Type of Treatment. Levels of the remainder of the variables are indicated in the columns labeled Sample, System, Design, Setting, Beh., and Phi. As noted at the bottom of the table, numeric codes reflect the levels for each Sample of Studies, Justice System, Quality of Research Design, Setting, and Behavioral Intervention. The minor subheadings in Table A1 enhance descriptions of type of treatment but did not enter into the analyses. The Comments column is intended as a guide for readers who wish to recreate the 2×2 tables that we drew from the original studies. Many of the comments will make little sense without reference to those original studies.

CRIMINAL SANCTIONS

Inspection of Table A1 provides an overview of the 30 "criminal justice" comparisons. Phi coefficients were signed positive when lower recidivism rates were found under "more" processing conditions. The first eight comparisons were culled from four studies of diversion through police cautioning versus regular processing, and only one phi estimate, a negative one, reached the .20 level (the standard of effectiveness in Whitehead and Lab was a positive phi coefficient of at least .20). The next set of 20 comparisons involved less versus more severe judicial dispositions, and six of the phi coefficients, all negative, equaled or exceeded .20. The final two studies in the criminal justice set reveal that completion of restitution contracts was only mildly associated with reduced recidivism rates. One might expect that the confound with selection factors would have had a stronger inflationary effect on the phi estimates. Overall, the findings of the 30 criminal justice comparisons were consistent with expectations: Only seven (20%) phi coefficients reached the criterion of .20 and, more consistent with labelling than deterrence theory, they were each negative in sign.

INAPPROPRIATE CORRECTIONAL SERVICE

Thirty-eight comparisons involved "inappropriate" treatments—treatments that we predicted would be either unrelated to recidivism or have a negative effect. Inspection confirms that only five phi coefficients reached the .20 level, and each was negative in sign. The mean phi coefficient was $-.06$. The only surprises in this set of comparisons were the positive phi coefficients, albeit statistically insignificant, yielded by Davidson et al.'s (1987) paraprofessional relationship-oriented program. Overall, the hypothesized ineffectiveness of inappropriate service was supported.

UNSPECIFIED CORRECTIONAL SERVICE

Table A1 provides an overview of 32 comparisons involving unspecified correctional service. The number of positive phi coefficients equaling or exceeding .20 was 10 (34%), and the mean phi was clearly positive but low (.13). In regard to our hypothesis, we now begin to uncover evidence of the effectiveness of rehabilitative service. Note, in addition, the many significant but low phi estimates. Obviously, many weak effects emerged significant statistically because of the large samples studied. Not as obvious, except upon a reading of the original papers (e.g., Palmer and Lewis, 1980), several Type of Client \times Type of Service interactions were found in this set of studies. Hence, some of the tabled effect size estimates are misleading because they reflect an averaging of what were actually positive and negative effects dependent upon type of case. For example, Palmer and Lewis (1980) reported that nonspecific family counseling for female first offenders was associated with clearly negative effects, apparently replicating the Druckman (1979) study. Unfortunately, these interactions were not reported in a manner that allowed the simple effects of treatment to be coded by type of case. Klein (1986) also reported some intriguing interactions that suggest weak or negative effects with low-risk cases. This pattern would be consistent with those tests of treatment in the inappropriate service set that involved the delivery of services to low-risk cases (studies 58-66 in Table A1).

APPROPRIATE CORRECTIONAL SERVICE

The overall pattern here reveals that 70% ($^{23/34}$) of the comparisons within the appropriate service set yielded a positive phi of at least .20, and the overall mean phi coefficient was .30. In every comparison but two, which involved token economy programs in residential settings, the phi coefficients were positive. Appropriate treatment appears to work at least moderately well. Note that many of the studies in the appropriate set involved small samples (and sample size is inversely correlated with effect size: Lipsey, 1989). Future research will explore the relative contributions of methodological, statistical, and therapeutic integrity factors to this correlation between sample size and effect size. Preliminary explorations, however, have revealed that the effect of Type of Treatment on phi estimates is found in both smaller and larger sample studies. For example, 30 of the 54 tests of appropriate service involved a control group with 30 or fewer cases, compared with only 28 of the 100 other tests of treatment. Among the small sample tests, 77% ($^{23/30}$) of the tests of appropriate treatment yielded a positive phi of at least .20 compared with 21% ($^{6/28}$) of the tests of less appropriate treatments. The corresponding figures among the tests based on larger samples were 63% ($^{15/24}$) for appropriate treatment and 7% ($^{5/72}$) for other treatments.

NONPARAMETRIC SUMMARY

The proportion of coefficients within each of the four levels of Type of Treatment reaching or exceeding the Whitehead and Lab (1989) standard of effectiveness were .00, .00, .34, and .70 for the criminal sanction, inappropriate service, unspecified service, and appropriate service sets, respectively; chi-square = 68.83, $p < .000$, Eta = .67, $r = .64$, gamma = .92.

Table A1. Summary of 154 Tests of Correctional Treatment

ID	Author (Year)	Sample	System	Design	Setting	Beh	Phi	Rec Rate: % (n)		Comments	
								Treat	Control		
TYPE OF TREATMENT: 1) CRIMINAL SANCTIONS											
Sanctioning vs Cautioning											
1	Kraus (81)	1	1	1	INSD	1	-15	41	(78)	27 (78)	
2	Klein (86)	2	1	2	INSD	1	-25*	73	(81)	49 (82) Release vs Petition	
Studies with Higher Risk Cases											
3	Mott (83)	1	1	2	INSD	1	-08	58	(167)	46 (26)	
4		1	1	2	INSD	1	19	53	(30)	80 (5) (girls)	
5	Farrington & P (81)	1	1	2	INSD	1	04	45	(11)	50 (8)	
Studies with Lower Risk Cases											
6	Mott (83)	1	1	2	INSD	1	-03	33	(57)	30 (174)	
7		1	1	2	INSD	1	-05	14	(7)	9 (75) (girls)	
8	Farrington & P (81)	1	1	2	INSD	1	10	0	(2)	12 (24)	
More vs Less Severe Disposition											
9	Viano (76)	1	1	1	1	SD	1	-08	26	(35)	19 (21)
10		1	1	1	1	SD	1	-20	26	(35)	10 (38) Informal adjustment
11		1	1	1	1	SD	1	-12	19	(21)	10 (38) Dismissal
12	Rausch (83)	1	1	1	1	SD	1	-01	47	(196)	44 (18)
13		1	1	1	1	SD	1	-02	47	(196)	45 (91) Probation
14		1	1	1	1	SD	1	-05	47	(196)	47 (45) Maximum Community
15		1	1	1	1	SD	1	04	40	(45)	44 (18) Community agent
16		1	1	1	1	SD	1	05	40	(45)	45 (91) Probation
17		1	1	1	1	SD	1	00	45	(91)	44 (18) Community agent
18	Kraus (78)	1	1	1	1	PPC	1	-28*	64	(90)	37 (90)
19	Horowitz & W (79)	1	1	2	1	PPC	1	-22	91	(196)	75 (67)
20		1	1	2	1	PPC	1	-32*	83	(29)	43 (106) Lower risk
21	Stephenson & S (74)	2	1	1	2		1	-23*	61	(44)	39 (44) Prob vs Group Home
22		2	1	1	2		1	-20*	59	(44)	39 (44) Inst vs Prob
23		2	1	1	2		1	02	59	(44)	61 (44) Inst vs Group Home
24	Phillips P F & W (73)	2	1	1	1	PPC	1	01	53	(15)	54 (13) Inst vs Prob
25	Vito & A (81)	2	2	1	1	PPC	1	-07*	17	(585)	12 (938) Shock Incar vs Prob
26	Petersilia T & P (86)	2	2	1	2		1	-07	41	(162)	34 (162) I vs P (viol offs)
27		2	2	1	2		1	-18*	61	(219)	43 (219) I vs P (prop offs)
28		2	2	1	2		1	-17*	35	(130)	20 (130) I vs P (drug offs)
Restitution (Successful Completion of)											
29	Schicor & B (82)	1	1	1	1	INSD	1	14	7	(59)	15 (55)
30	Schneider & S (84)	1	1	1	1	PPC	1	18*	60	(190)	80 (61)

CORRECTIONAL TREATMENT META-ANALYSIS 401

Table A1. (continued)

ID Author (Year)	Sample	System	Design	Setting	Beh	Phi	Rec Rate: % (n)		Comments
							Treat	Control	
TYPE OF TREATMENT: 2) INAPPROPRIATE CORRECTIONAL SERVICE									
Intensive Non-Behavioral Group Interaction (including Recreation)									
31 Byles (81)	1	1	1	1	IPPC	1 -11	71	(31) 60	(35) Attendance Centers
32	1	1	1	1	IPPC	1 08	68	(25) 76	(49)
33 Shorts (86)	2	1	1	1	INSD	1 -01	48	(43) 47	(19)
34	2	1	1	1	INSD	1 03	10	(31) 12	(17)
35 Winterdyk & R (82)	2	1	2	1	SD	1 00	20	(30) 20	(30) Wilderness Program
Non-Directive Client-Centered/Psychodynamic Counselling									
36 Klein A & P (77)	1	1	2	1	SD	1 -11	60	(30) 49	(56)
37	1	1	2	1	SD	1 -17	61	(30) 40	(10) (Sibs)
38 Adams & V (82)	1	1	2	2		1 -44*	100	(14) 69	(13) Group Psychodrama
39 Davidson R B M & E (87)	2	1	1	1	INSD	1 06	33	(12) 43	(112) Rel vs Beh/Adv/Act
40	2	1	1	1	INSD	1 16	33	(12) 58	(89) Rel vs Controls
41 Berman (78)	2	2	2	1	IPPC	1 00	25	(16) 25	(16) Non-Bev Para-prof
Non-Behavioral Milieu Therapy/Guided Group Interaction									
42 Stringfield (77)	1	1	1	1	IPPC	1 -31*	56	(32) 25	(20) Milieu vs Fam
43 Clarke & C (78)	1	1	2	2		1 -01	70	(86) 69	(87)
44 Stephenson & S (74)	2	1	1	1	IPPC	1 -20*	59	(44) 39	(44) GGI vs Prob
45	2	1	1	2		1 00	59	(44) 59	(44) GGI vs Inst
46	2	1	1	2		1 02	59	(44) 61	(44) GGI vs Group Home
47 Empey & E (72)	2	1	2	1	IPPC	1 -03	58	(71) 54	(79) GGI vs Prob
48	2	1	1	1	IPPC	1 16*	64	(44) 79	(132) GGI vs Incar
49 Craft S & G (64)	2	1	2	2		1 -13	58	(24) 46	(24) Milieu vs Auth
Non-Behavioral Weakly-Focused Academic/Vocational Approaches									
50 Willman & S (82)	1	1	1	1	SD	1 -11	71	(68) 60	(68)
51 Maskin (76)	1	1	1	2		1 -39*	50	(30) 13	(30)
52 Hackler & H (75)	1	1	2	1	INSD	1 -05	33	(85) 29	(70)
53	1	1	2	1	INSD	1 07	25	(67) 32	(131)
54 Zeisel (82)	2	2	2	1	IPPC	1 00	49	(???) 49	(???) (TARP)
Confrontational Groups (Scared Straight)									
55 Buckner & C-L (83)	1	1	1	1	SD	1 -04	41	(100) 37	(100)
56	1	1	1	1	SD	1 11	22	(50) 32	(50) (female)
57 Lewis (83)	1	1	2	1	SD	1 -16	67	(55) 81	(53)
Mismatched According to Risk or Responsivity/Need Systems									
58 Sorenson (78)	1	1	2	1	INSD	1 -35*	30	(30) 4	(45)
59 Byles & M (79)	1	1	2	1	SD	1 -12	57	(94) 43	(114)
60 Gruher (79)	1	1	2	1	SD	1 06	32	(38) 38	(40)
61 Quay & L (77)	1	1	2	1	INSD	1 07	28	(268) 36	(92)
62 O'Donnell L & F (79)	1	1	2	1	INSD	2 -07	25	(169) 19	(130)
63	1	1	2	1	INSD	2 -10	18	(116) 11	(65) (female)
64 Baird H & B (79)	2	2	2	1	IPPC	1 -13	10	(58) 3	(58)
65 Andrews & K (80)	2	2	2	1	IPPC	1 -09	17	(58) 11	(62) Para-prof prog
66 Andrews K M & R (86)	2	2	2	1	IPPC	1 -09	14	(98) 2	(28) Para-prof prog
67 Grant & G (59)	2	2	2	2		1 -14*	52	(91) 38	(144) Low maturity
68 Andrews & K (80)	2	2	2	1	IPPC	1 11	42	(23) 48	(13) Low Emp/High Risk

Table A1. (continued)

ID Author (Year)	Sample System Design	Setting	Beh	Phi	Rec Rate: % (n)			Comments		
					Treat	Control				
TYPE OF TREATMENT: 3) UNSPECIFIED CORRECTIONAL SERVICE										
Service-Oriented Diversion										
69 Regoli W & P (85)	1	1	1	1	19*	2	(52)	11	(52)	Complete prog vs pre-program controls
70	1	1	1	1	16*	8	(98)	21	(98)	
71	1	1	1	1	31*	6	(61)	32	(61)	
72	1	1	1	1	12	10	(72)	18	(72)	
73	1	1	1	1	-06	29	(119)	24	(119)	
74	1	1	1	1	-05	26	(107)	24	(107)	
75 Lipsey C & B (81)	1	1	1	1	18*	26	(776)	44	(476)	Complete vs Incomplete
76	1	1	1	1	10*	27	(870)	37	(533)	
77	1	1	1	1	10*	35	(543)	45	(333)	
78 Whitaker & S (84)	1	1	1	1	10*	33	(???)	46	(???)	More vs Less Diverse
79 Palmer & L (80)	1	1	1	1	07*	25	(1345)	31	(1192)	Unspec quasi control
80 Gilbert (77)	1	1	1	1	30*	34	(58)	65	(78)	Assign vs Preprog conts
81 Klein (86)	2	1	2	1	17*	57	(88)	73	(81)	Ref vs Petition
82	2	1	2	1	12	62	(55)	73	(81)	Ref+ vs Petition
83	2	1	2	1	-08	57	(88)	49	(82)	Ref vs Release
84	2	1	2	1	-13	62	(55)	49	(82)	Ref+ vs Release
Appropriateness Uncertain On Targets/Style										
85 Romig (76)	1	1	1	1	15*	14	(301)	27	(127)	Parole Supervision
86 Jackson (83)	1	1	2	1	03	82	(198)	84	(98)	Parole Supervision
87 Barkwell (76)	1	1	2	1	-16	88	(16)	75	(16)	Prob Serv vs Surveill
88 Druckman (79)	1	1	1	1	-17	50	(14)	33	(15)	Family Counseling
89 Seidman R & D (80)	2	1	2	1	46*	50	(12)	90	(12)	Parapro Advocacy
90 Wade M L & F (77)	2	1	1	1	51*	15	(34)	70	(77)	Family crisis
91 Romig (76)	2	1	1	1	10*	12	(177)	20	(251)	Mths served/rel order
92 Johnson & G (83)	2	1	2	1	05	3	(87)	5	(87)	State Vocational Serv
93 Sowles & G (70)	2	1	2	1	22	37	(30)	60	(15)	Ind/Group (boys)
94	2	1	2	1	38	0	(10)	20	(5)	(girls)
95 Ostrum S R & M (71)	2	1	2	1	22	26	(19)	48	(19)	Mixed socio-psych prog
96 Redfering (73)	2	1	1	1	29	35	(17)	64	(14)	Cl-Ce Group (appro tar)
97 Jesness in grant (65)	2	1	2	1	05	73	(11)	77	(13)	Small vs Large Units
98 Ross F & E (88)	2	2	2	2	23	47	(17)	70	(23)	Life Skill vs Reg Prob
99 Vinglis A & C (82)	2	2	2	1	-05	15	(58)	19	(62)	Impaired Driving
100 Walsh (85)	2	2	1	1	21*	24	(50)	44	(50)	Gen Equiva Prog

Table A1. (continued)

ID Author (Year)	Sample	System Design	Setting	Beh	Phi	Rec Rate: % (n)				Comments	
						Treat	Control				
TYPE OF TREATMENT: 4) APPROPRIATE CORRECTIONAL SERVICE											
Short-Term Behavioral/Systems Family Counseling											
101 Alexander C S P (76)	1	1	1	1	SD	2	64*	0	(12)	56	(9)
102 Klein A & P (77)	1	1	2	1	SD	2	23*	26	(46)	48	(56)
103	1	1	2	1	SD	2	18	20	(46)	40	(10) (sibs)
104	1	1	2	1	SD	2	31*	26	(46)	57	(30)
105	1	1	2	1	SD	2	41*	20	(46)	60	(30) (sibs)
106 McPherson M & R (83)	1	1	2	1	SD	1	20*	33	(15)	58	(60) (target = discipline)
107 Gordon A G & M (88)	2	1	1	1	SD	2	83*	0	(12)	75	(4) (girls)
108	2	1	1	1	SD	2	44*	20	(15)	65	(23) (boys)
109 Stuart J & T (76)	2	1	2	1	INSD	2	19	0	(30)	7	(30)
110 Barton A W T & W (85)	2	1	1	2		2	41*	60	(30)	93	(44)
Structured One-on-One Paraprofessional/Peer Program											
111 Kelly H & B (79)	1	1	1	1	INSD	1	26*	0	(65)	12	(63) (target = thinking)
112 Mitchell (83)	1	1	1	1	PPC	2	29*	14	(29)	43	(63)
113 Ross & M (77)	1	1	1	2		2	33	7	(15)	33	(15)
114	1	1	1	2		2	46*	7	(15)	60	(45)
115 Seidman R & D (80)	2	1	2	1	INSD	2	51*	48	(25)	100	(12) Beh/adv vs Controls
116	2	1	2	1	INSD	2	60*	25	(12)	92	(12) Beh vs Controls
117	2	1	2	1	INSD	2	17	25	(12)	50	(12) Beh vs Advocacy
118 Davison R B M & E (87)	2	1	1	1	INSD	2	15*	43	(112)	58	(89) Beh/Adv/Action vs C
119 Andrews (80)	2	2	2	1	PPC	2	15*	15	(72)	28	(116) Hi Emp Hi So Officers
Specialized Academics/Vocational Services											
120 Bachra & Z (78)	1	1	1	1	INSD	1	38*	7	(31)	42	(48) (specific focus)
121 Kratcoski & K (82)	1	1	1	1	PPC	2	65*	42	(38)	100	(83)
122 Walter & M (80)	2	1	1	1	PPC	2	62*	9	(53)	70	(23)
123 Shore & M (79)	2	1	2	1	INSD	1	52*	40	(10)	90	(10) (intense/individualized)
Intensive Structured Skill Training											
124 Collingwood D & W (76)	1	1	1	1	SD	2	41*	11	(813)	51	(196) Participants vs Nonpar
125 Williams (84)	1	1	1	1	SD	2	33*	21	(564)	64	(77) Participants vs Nonpar
126 Sarason & G (73)	2	1	2	2		2	18*	19	(64)	34	(64) Modeling vs Contr
127	2	1	2	2		2	24*	14	(64)	34	(64) Discuss vs Control
128 Ross F & E (88)	2	2	2	1	PPC	2	52*	18	(22)	70	(23) Cog-Beh vs Reg Prob
129	2	2	2	1	PPC	2	31*	18	(22)	47	(17) Cog-Beh vs Life Skill
130 Dutton (86)	2	2	1	1	PPC	2	43*	4	(50)	40	(50) Cog-beh (wife batter)
Introduction of Individualized Rehabilitative Regime											
131 Jesness (75)	1	1	1	2		2	10*	32	(398)	42	(499) Token Eco vs Pre-Prog

Table A1. (continued)

ID Author (Year)	Sample	System	Design	Setting	Beh	Phi	Rec Rate: % (n)		Comments
							Treat	Control	
132	1	1	1	2	1	14*	33 (453)	47 (660)	(target=ind crimino need)
133 Ross & M (76)	1	1	1	2	2	27	10 (10)	33 (15)	
134	1	1	1	2	2	38*	10 (10)	60 (45)	
Token Economy									
135 Kirigin B A & W (82)	1	1	1	2	2	21	27 (38)	47 (30)	(girls)
136	1	1	1	2	2	12	57 (102)	73 (22)	(boys)
137 Davidson & W (77)	1	1	1	2	2	-26*	?? (??)	?? (??)	
138 Ross & M (77)	1	1	1	2	2	-23	60 (45)	33 (15)	
139 Phillips P F & W (73)	2	1	1	2	2	36*	18 (16)	53 (15)	Ach Place vs Inst
140	2	1	1	2	2	37*	18 (16)	54 (13)	Ach Place vs Prob
Individual/Group Counselling									
141 Persons (67)	2	1	2	2	2	29*	32 (41)	61 (41)	Ind + Group
Appropriately Matched According to Risk or Responsivity/Need Systems									
142 Sorenson (78)	1	1	2	1NSD	1	06	25 (44)	31 (26)	
143 Byles & M (79)	1	1	2	1SD	1	27*	68 (60)	92 (37)	
144 Gruher (79)	1	1	2	1SD	1	07	56 (16)	63 (30)	
145 Quay & L (77)	1	1	2	1NSD	1	23*	36 (164)	65 (40)	
146 O'Donnell L & F (79)	1	1	2	1NSD	2	20	62 (37)	81 (21)	(boys)
147	1	1	2	1NSD	2	08	38 (13)	50 (2)	(girls)
148 Barkwell (76)	1	1	2	1PPC	1	35*	56 (16)	88 (16)	
149	1	1	2	1PPC	1	20	56 (16)	75 (16)	
150 Baird H & B (79)	2	2	2	1PPC	1	17*	16 (184)	30 (184)	
151 Andrews K M & R (86)	2	2	2	1PPC	1	31*	33 (54)	75 (12)	Para-prof prog
152 Andrews & K (80)	2	2	2	1PPC	1	82*	0 (11)	80 (10)	(Hi Emp & Risk)
153 Grant & G (59)	2	2	2	2	1	09	29 (135)	38 (141)	(High Maturity)
154 Andrews & K (80)	2	2	2	1PPC	1	27*	31 (34)	58 (23)	Para-prof prog

* $p < .05$ (Chi square)

Note: The value labels for codes "1" and "2" are as follows: Sample of Studies (Whitehead & Lab, Sample 2), Justice System (juvenile, adult), Quality of Research Design (weaker, stronger), Behavioral Intervention (no, yes) and Setting (community, institutional/residential). The letters beside code "1" for Setting refer to different types of community settings (NSD: nonsystem diversion; SD: system diversion; PPC: probation, parole, community).

The Positive Effects of Cognitive-Behavioral Programs for Offenders:
A Meta-Analysis of Factors Associated with Effective Treatment

Nana A. Landenberger & Mark W. Lipsey

Vanderbilt Institute for Public Policy Studies

Correspondence: Mark W. Lipsey, Vanderbilt Institute for Public Policy Studies, 1207 18th Ave. South, Nashville, TN 37212 or mark.lipsey@vanderbilt.edu.
Supported by grants from the National Institute for Mental Health (NIMH)

In press, *Journal of Experimental Criminology*, 2005

Abstract

A meta-analysis of 58 experimental and quasi-experimental studies of the effects of cognitive-behavioral therapy (CBT) on the recidivism of adult and juvenile offenders confirmed prior positive findings and explored a range of potential moderators to identify factors associated with variation in treatment effects. With method variables controlled, the factors independently associated with larger recidivism reductions were treatment of higher risk offenders, high quality treatment implementation, and a CBT program that included anger control and interpersonal problem solving but not victim impact or behavior modification components. With these factors accounted for, there was no difference in the effectiveness of different brand name CBT programs or generic forms of CBT.

Keywords: Cognitive-behavioral therapy, cognitive-behavioral treatment, CBT, criminal rehabilitation, treatment effectiveness, offenders, recidivism, evaluation, meta-analysis.

The Positive Effects of Cognitive-Behavioral Programs for Offenders:
A Meta-Analysis of Factors Associated with Effective Treatment

Several well conducted meta-analyses have identified cognitive-behavioral therapy (CBT) as a particularly effective intervention for reducing the recidivism of juvenile and adult offenders. Pearson, Lipton, Cleland, and Yee (2002), for instance, conducted a meta-analysis of 69 research studies covering both behavioral (e.g., contingency contracting, token economy) and cognitive-behavioral programs. They found that the cognitive-behavioral programs were more effective in reducing recidivism than the behavioral ones, with a mean recidivism reduction for treated groups of about 30%. Similarly, a meta-analysis by Wilson, Bouffard, and MacKenzie (2005) examined 20 studies of group-oriented cognitive behavioral programs for offenders and found that CBT was very effective for reducing their criminal behavior. In their analysis, representative CBT programs showed recidivism reductions of 20-30% compared to control groups.

Although these meta-analyses provide strong indications of the effectiveness of cognitive-behavioral treatment for offenders, they encompassed considerable diversity within the range of offender types, outcome variables, quality of study design, and (especially in Pearson et al., 2002) variations in what was counted as a cognitive-behavioral treatment. A more circumscribed meta-analysis conducted by Lipsey, Chapman, and Landenberger (2001) examined 14 experimental and quasi-experimental studies that emphasized cognitive change as the defining condition of CBT, considered only effects for general offender samples, and focused on reoffense recidivism as the treatment outcome. The results showed that the odds of recidivating for offenders receiving CBT were only about 55% of that for offenders in control groups. Lipsey and Landenberger (in press) then focused further on an updated and overlapping set of 14 randomized experiments and found that the mean recidivism for the treatment groups in those studies was 27% lower than that of the control groups.

Variation in Effectiveness

Meta-analysis has thus consistently indicated that CBT, on average, has significant positive effects on recidivism. However, there is also significant variation across studies in the effect sizes that contribute to those mean values that must be acknowledged. Identification of the moderator variables that describe the study characteristics associated with larger and smaller effects is another kind of contribution meta-analysis can make to understanding the effectiveness of CBT with offenders. Of particular importance is the role such moderator analysis can play in ascertaining which variants of CBT are most effective and for which offenders.

Lipsey and Landenberger (in press) identified a few factors that were related to variation in recidivism effects. They found that treatment of high risk offenders, greater levels of CBT training for treatment providers, and CBT programs set up for research or demonstration purposes (in contrast to “real world” routine practice programs) were associated with larger effects. What most characterized the research and demonstration programs, in turn, was smaller sample sizes, greater monitoring of offender attendance and adherence to the intervention plan (treatment fidelity checks), and providers with mental health backgrounds. These factors suggest that treatment effectiveness is mainly a function of the quality of the CBT provided.

That meta-analysis involved only a small number of studies, however, and did not permit much exploration of potential moderator variables. Wilson, Bouffard, and MacKenzie (2005) computed mean effect sizes separately for Moral Reconciliation Therapy (MRT), Reasoning and Rehabilitation (R&R), and “other” CBT programs. They found that R&R showed somewhat smaller mean effects, but did not examine any other moderator variables. Pearson et al. (2002), as noted above, compared the effects of behavioral and cognitive-behavioral programs and also reported larger effects for better designed studies, but did not pursue further moderator analysis.

There has thus been only limited meta-analytic investigation of factors identifiable in the body of research on CBT that are associated with variation in its effects on offender recidivism. The most likely candidates for such factors fall into categories relating to the specific type of CBT program applied, the nature and extent of its implementation, the characteristics of the offenders to whom it is provided, and the study methods used to investigate its effects.

The type of CBT program, for instance, relates, first, to the “brand name” curriculum used, such as MRT and R&R as examined by Wilson, Bouffard, and MacKenzie (2005). CBT programs also differ in the nature and mix of treatment elements included, e.g., whether oriented mainly toward cognitive restructuring exercises or cognitive skills training and whether such topics as anger management, relapse prevention, interpersonal problems solving and moral reasoning are covered. Another potentially important distinction is whether CBT constitutes virtually the full program offered or is combined with other services, e.g., educational classes, vocational training, or mental health counseling.

Factors related to program implementation that might influence the effects of CBT include, most centrally, how much treatment is provided. The duration of CBT programs in correctional settings, for instance, varies from weeks to years and may involve many meetings per week or less than one. The fidelity of the implementation to the curriculum specifications may also be important along with the degree of expertise possessed by the personnel providing the program. As mentioned earlier, Lipsey and Landenberger (in press) found that programs implemented principally for research or demonstration purposes showed larger effects than routine practice programs. In these programs, the researchers themselves generally exercise control over the various phases and facets of implementation. The treatment provided to offenders in those circumstances almost certainly differs in important ways from that provided when the program under everyday conditions in criminal justice settings.

Among the characteristics of the offenders participating in CBT that may influence the outcome are age, gender, and ethnic background as well as criminal history and other such risk indicators. The “risk principle” of Andrews et al. (1990), for instance, posits that effective treatment will have greater impact on higher-risk offenders because they have more room for improvement than lower-risk offenders.

Finally, there is good reason to believe that the methods and procedures used in the research will influence the magnitude of the observed effects. Random assignment studies are expected to yield unbiased results while findings from nonrandomized comparisons may over or understate effects. Weisburd, Lum, and Petrosino (2001), for example, found larger effects in nonrandomized studies of criminal justice programs though, for CBT, Pearson et al. (2002)

reported larger effects for studies of higher methodological quality. After assignment to conditions, attrition from outcome measurement can also bias effect estimates if, as is likely, it is not randomly distributed across conditions. The operationalization of the outcome measure is another potential source of difference. Even when the focus is on recidivism, some studies index it with the rate of rearrest, others by assessing reconvictions, incarcerations, probation or parole violations, and the like. Moreover, the timing of recidivism measurement varies, ranging from a period close to the end of treatment to months or years later.

Purpose of This Meta-Analysis

The objective of this meta-analysis is to examine the relationships of selected moderator variables to the effects of CBT on the recidivism of adult and juvenile offenders. In order to have a sufficient number of studies to permit examination of between-study differences, an especially thorough search was made of the available research. To assist in expanding the number of studies, and to ensure methodological diversity so that variation in methods could be investigated, quasi-experimental studies were included as well as randomized field experiments. Though not all potentially interesting moderator variables are reported well enough in the source studies to allow systematic comparison, a detailed coding protocol was applied to extract as much relevant information for analysis as possible from each study report.

Methods

Criteria for Inclusion of Studies

Studies were assessed and selected for this meta-analysis if they met the following criteria:

Intervention. The treatment under investigation was a variant of cognitive-behavioral therapy representing or substantially similar to such recognized “brand name” CBT programs as Reasoning and Rehabilitation (Ross & Fabiano, 1985), Moral Recondition Therapy (Little & Robinson, 1986), Aggression Replacement Training (Goldstein & Glick, 1987), the Thinking for a Change curriculum (Bush, Glick & Taymans, 1997), and the Cognitive Interventions Program (NIC, 1996). In particular, it was directed toward changing distorted or dysfunctional cognitions (cognitive restructuring) or teaching new cognitive skills and involved therapeutic techniques typically associated with CBT, i.e., structured learning experiences designed to affect such cognitive processes as interpreting social cues, monitoring one’s own thought processes, identifying and compensating for distortions and errors in thinking, reasoning about right and wrong behavior, generating alternative solutions, and making decisions about appropriate behavior. If CBT was offered in the context of a multimodal program that simultaneously provided other services, the CBT must have been provided to all participants and constitute a major component of the program.

Participants. The recipients of the intervention were criminal offenders, either juveniles or adults, treated while on probation, incarcerated/institutionalized, or during aftercare/parole. Offenders were drawn from a general offender population and not selected for, or restricted to, those committing specific types of offenses (e.g., sex offenses, DUI, drug offenses, status offenses).

Outcome measures. The study reported criminal offending subsequent to treatment as an outcome variable. Outcome results were presented in a quantitative form that permitted computation or reasonable estimation of an effect size statistic representing the difference in recidivism rates between treated and untreated offenders.

Research methods. The study used a randomized or quasi-experimental design that compared a CBT treatment condition with a control condition that did not include CBT treatment. Quasi-experimental designs were eligible only if subjects in the treatment and control conditions were matched or statistically controlled on pre-treatment risk-related variables (e.g., relevant personal, demographic, and criminal background characteristics) or if pre-treatment measures of criminal or antisocial behavior or significant risk factors for such behavior were reported in a form that permitted assessment of the initial equivalence of the treatment and control groups. To eliminate explicit self-selection as a biasing factor in group assignment, however, studies were not included if the control groups were created with individuals who began CBT but dropped out prior to completing treatment or who were offered CBT and refused. Control groups could represent placebo, wait-list, no treatment, or “treatment as usual” conditions, with the latter restricted to cases of clearly routine probation, institutional, or aftercare/parole practices.

Source. Both published and unpublished studies were eligible for inclusion, conducted in any country, and reported in any language.

Search Strategy

An initial set of eligible studies came from those assembled and analyzed for the Lipsey, Chapman, and Landenberger (2001) and Lipsey & Landenberger (in press) meta-analyses. This number was expanded through a comprehensive search using the following procedures.

Meta-analysis databases. The second author has constructed a meta-analysis database of coded studies for interventions with juvenile offenders based on a comprehensive search for studies reported in 2002 or earlier. All the studies in that database were reviewed for eligibility and an independent search was conducted for studies published after 2002. In addition, the studies in a database of interventions with adult offenders that is nearing completion were reviewed for eligibility.

Database searches. Computerized bibliography searches were conducted for studies reported from 1965 through 2005. To the best of our knowledge, the first systematic applications of CBT to offenders were developed and published in the mid-1970s (e.g., Yochelson & Samenow, 1976); searching back to 1965 was aimed at ensuring that none were missed. The keywords for searching were concatenations of words describing the population (e.g., inmates, offenders), CBT treatment (e.g., cognitive, CBT, criminal thinking), and effectiveness research (e.g., outcomes, evaluation, effectiveness). The databases searched included the Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR), Dissertation Abstracts Online, ERIC, MEDLINE, The National Criminal Justice Reference Service (NCJRS), PsychInfo/PsychLit, Sociological Abstracts, and a number of others.

Cross-referencing of bibliographies. Relevant review articles, meta-analyses, and primary studies reviewed for eligibility were scanned for citations to potentially eligible studies.

Internet searches. Relevant government websites (e.g., NIJ, NIC, OJJDP, Home Office) as well as foundation, professional associations and policy research firm websites were searched. In addition, keyword searches were conducted using search engines such as google.com.

Journals. Vanderbilt University subscribes to a large number of electronic journals and the full text of those judged relevant was searched with selected keywords. Major journals publishing empirical studies related to crime and delinquency were also hand searched for eligible studies.

Informal sources. Unpublished results from evaluations of two CBT programs were available from the first author, and several colleagues alerted us to eligible studies that were not accessible through the above channels.

The search for CBT studies on adult offenders produced 2947 study citations with 771 reports judged promising enough to retrieve for closer examination. The search for juvenile offender studies produced 1487 study citations with 299 reports retrieved. Review of the retrieved studies ultimately identified 58 studies meeting the criteria for inclusion in the present meta-analysis.

Data Management and Extraction

Descriptive and outcome data were coded for each of the 58 eligible studies using a coding protocol developed specifically for this purpose. Table 1, presented later, shows the major coding categories used for descriptive information. Recidivism outcomes were reported in several different forms but, in virtually all instances, either the proportions of offenders in each research condition that recidivated were specified or information was provided from which the proportions could be estimated. When more than one recidivism outcome was reported, only one was selected for analysis using criteria that maximized cross-study similarity on the variables and times of measurement. This procedure favored rearrest recidivism, then reconviction and incarceration in that order, and the measure taken closest to 12 months post-treatment.

The selected recidivism outcomes were coded as odds ratios representing the odds of “success” (not recidivating) for treatment group participants relative to the odds for control participants. For binary outcomes, the odds ratio provides an effect size statistic that has favorable properties and yields readily interpretable results (Haddock, Rindskopf, & Shadish, 1998). Statistical analysis with odds ratios is facilitated if they are represented by their log, so the logged odds ratios were used in all analyses. Otherwise, the statistical analysis was conducted using conventional meta-analysis techniques (Lipsey & Wilson, 2001) with each effect size weighted by its inverse variance in random effects analyses.

Results

Table 1 summarizes the characteristics of the 58 studies included in the meta-analysis. Several features of this research are notable. Randomized designs, matched designs, and group comparisons using neither of these procedures are represented in roughly equal numbers and

involve a wide range of sample sizes. Attrition from outcome measurement is virtually zero in a majority of the studies but ranges over 30% in some of the remaining ones. About half the programs studied were implemented as routine practice with the other half set up and implemented by researchers as either demonstration or research programs, with demonstration programs defined as those mounted mainly for research purposes but at a scale and in a manner somewhat more representative of actual practice than those categorized as research programs. More studies were conducted with adult than juvenile offenders and most used only or predominately male offenders. Treatment was administered while the offenders were incarcerated in a correctional institution in nearly half the studies and generally lasted less than 20 weeks. In most instances, the treatment providers had little or no evident mental health background and had received relatively minimal training in cognitive behavioral therapy. The treatment was typically one of the “brand name” manualized CBT programs and incorporated multiple treatment elements.

Effect Size Variation Associated with Study Methods

The mean odds ratio representing the average effect of intervention was 1.53 ($p < .001$), indicating that the odds of success (no recidivism in the post-intervention interval of approximately 12 months) for individuals in the treatment group were more than one and a half times as great as those for individuals in the control group. In relation to the mean recidivism rate for the control groups of about .40, this odds ratio indicates a recidivism reduction of 25% to .30. There was also significant variation across studies in the odds ratio for intervention effects ($Q = 214.02$, $df = 57$, $p < .001$). We turn now to an examination of the study characteristics associated with that variation.

The recidivism effects observed in the studies in this meta-analysis are potentially influenced by both the methodological characteristics of the studies and the substantive attributes of the treatments and the recipients. One of the first steps in the analysis, therefore, was to determine which methodological features were correlated with the effect sizes so they could be controlled while examining relationships with substantive attributes. The method variables available from the study coding and considered relevant for this purpose were as follows:

- (a) Design, categorized as randomized, matched, or neither, each dummy coded to produce three design variables.
- (b) Design problem-- indications of initial nonequivalence between groups on pretreatment variables, or problems during or after the intervention that could have led to nonequivalence of the treatment and control group, rated by the coder on a 3-point scale (1=favors control group; 2=favors neither or insufficient evidence; 3=favors treatment group).
- (c) Attrition proportion—the proportion of the total initial sample (treatment plus control group) for which recidivism outcome data were not available.
- (d) Intent to treat, coded yes/no for whether treatment dropouts were retained in the treatment group for the recidivism outcome data reported in the study.
- (e) Type of recidivism, categorized as rearrests, reconvictions, incarcerations, or other with each dummy coded to produce four recidivism variables.
- (f) Recidivism interval represented by the number of months posttreatment over which recidivism was measured. Because of the possibility of more frequent recidivism in early months than later ones, the log of this variable was also used in the analysis.

Table 2 shows the zero-order correlation of each of the method variables with the recidivism effect sizes (represented as logged odds ratios). These are inverse-variance weighted, random effects analyses with the random effects component estimated using maximum likelihood techniques (Raudenbush, 1994).

As Table 2 reports, there was no significant relationship overall between the effect sizes and the study design. In particular, the effects observed in randomized studies did not differ significantly from those observed in matched studies or those with comparison groups that were neither matched nor randomly assigned. Nor was a significant relationship observed for the coder's rating of whether there was an evident design problem favoring the control or treatment group, that is, indications of nonequivalence that might affect recidivism outcomes. Similarly, there were no significant correlations with the attrition proportion, the way recidivism was measured, or the interval over which it was measured.

The only methodological variable that showed a significant ($p < .10$) relationship with the effect size was whether the study presented the treatment-control contrast as an intent-to-treat analysis. When the treatment dropouts were included in the outcome recidivism, the effect sizes were smaller than when they were excluded, as would be expected. In light of this indication that the intent-to-treat variable might influence effect sizes, it was carried forward as a control variable for the analysis of the relationships between effect sizes and substantive factors relating to the treatment and recipients. As a further precaution against confounds with methodological characteristics, the other three method variables with zero-order correlations of .10 or greater were also included as method controls (design problem, attrition proportion, and rearrest recidivism).

Effect Size Variation Associated with the Treatment and its Recipients

The relationship between the recidivism effect sizes and each of the descriptive variables for CBT treatment and its recipients (see listing in Table 1) was next examined with the four selected method variables included as controls. These analyses were conducted with a set of random effects multiple regressions that included a descriptive variable and the four control variables. These were run separately for each descriptive variable in this initial analysis to ensure that any having potentially important relationships with effect size were identified despite whatever correlations they had with other variables in the set. Because of the modest number of studies and the broad confidence intervals associated with random effects analysis, $\alpha = .10$ was set as the threshold for statistical significance. Table 3 presents the results.

The variables in Table 3 are grouped into categories that represent different aspects of the studies and the nature of the CBT treatment studied. The most general study characteristics (country, type of publication, and date of publication) showed no significant relationships with effect size. The other candidate moderator variables are grouped according to a simple model that assumes that, with method variables controlled, treatment effects will be a function of the characteristics of the participants, the amount of treatment received, the quality of the treatment implementation, and the specific type of treatment.

Participant characteristics. Of the characteristics of the treatment recipients that could be coded from most studies, only the recidivism risk rating was significantly related to the effect sizes. This rating was made by the coder on the basis of the description in the study of the criminal history of the treated offenders and the recidivism rate of the control group. That rating, in turn, was not significantly correlated with any of the other variables describing the participants shown in Table 3. It is worth noting that there was no relationship between effect size and whether the treated offenders were juveniles or adults. The gender mix of the CBT recipients also showed no relationship to effect size but, as Table 1 shows, most of the samples were all male or predominately male so there was little variation on this measure.

Amount of CBT. Dosage variables were coded as the number of sessions per week, the number of hours of treatment per week, the total hours of treatment, and the number of weeks of treatment from beginning to end (see Table 1). The distributions for the latter three had long tails and the logged values of these variables were used in the analysis (and showed stronger relationships to effect size than the unlogged versions). As Table 3 shows, all these variables except length of treatment were significantly related to effect size. Total hours, which showed the largest relationship, however, is a function of both the number of hours per week and the number of weeks. The study-level correlations among these variables showed that length of treatment was significantly related to total hours ($r=.51$) as were the number of sessions per week ($r=.58$) and number of hours per week ($r=.75$), with the latter two being highly correlated with each other ($r=.81$).

From this pattern of relationships we concluded that the best representation of the amount of treatment should distinguish the number of sessions or hours per week from the length of the treatment. That approach allows further examination of the finding in Table 3 that number of sessions and hours per week are related to the effect size but, apparently, the duration of treatment is not. Between the number of sessions per week and the number of hours, sessions showed the stronger relationship to effect size. Table 3 also shows the interaction between number of sessions and length of treatment, but it was not significantly related to effect size.

Quality of CBT Implementation. In this category we include the practice-research dimension that distinguishes between CBT treatments implemented on a routine basis in real-world criminal justice contexts, demonstration programs in similar circumstances but with significant influence by the researcher, and research programs implemented by the researchers largely for research purposes. Our assumption is that the progressively greater involvement of researchers translates into better implementation and more fidelity to the treatment protocol.

Table 4 shows that the study-level correlations were all significant between the practice-research variable and the other variables assumed related to implementation quality listed in Table 3-- proportion of dropouts from treatment, extent of implementation monitoring reported, amount of CBT training indicated for providers, and the mental health background of the providers. Table 3 shows relationships in the expected direction with effect size for all these variables except providers' mental health background, though only proportion of treatment dropouts and the practice-research dimension reached statistical significance. To summarize the relationship of these implementation quality variables to effect size, a composite variable was created in the form of a factor score from a principal components analysis. As shown in Table 3, that

composite implementation factor is more strongly related to the effect sizes than any of the component variables.

Other program characteristics. Table 3 also shows the relationship between effect size and two other program characteristics. One is the setting within which CBT was provided, differentiated between treatment while incarcerated and treatment in the community (e.g., for probationers and parolees); this variable was not significantly related to effect size. The extent to which CBT was emphasized in the treatment program, on the other hand, did show a significant relationship. That variable ranged across categories of CBT supplemented by other services, CBT with some other treatment elements, and CBT alone. As the negative sign on the coefficient in Table 3 shows, the effects are significantly larger when CBT is combined with other services. Examples of such components include mental health counseling, employment and vocational training, and educational programs.

Specific nature of the CBT treatment. The last two sections of Table 3 show two alternative ways of representing the specific nature of the CBT treatment provided. One set of dummy-coded items differentiates the various major named types of CBT along with a somewhat more generic category of CBT programs focusing on substance abuse and two residual categories of less common but manualized treatments and a few that do not appear to be manualized. None of these program variables is significantly related to effect size, meaning that no brand of CBT produces effects that stand out from the average of the other brands.

The other way we coded CBT treatment was in terms of the specific treatment elements identified in the descriptions provided in the study reports. Those descriptions varied in detail and extensiveness but when they mentioned a distinct treatment element, we coded it as present using a dummy code. The elements that appeared with sufficient frequency to support analysis are shown in Table 3, defined briefly as follows:

- Cognitive skills: Training on general thinking and decision-making skills such as to stop and think before acting, generate alternative solutions, evaluate consequences, and make decisions about appropriate behavior.
- Cognitive restructuring: Activities and exercises aimed at recognizing and modifying the distortions and errors that characterize criminogenic thinking.
- Interpersonal problem solving: Training in problem-solving skills for dealing with interpersonal conflict and peer pressure.
- Social skills: Training in prosocial behaviors, interpreting social cues, taking other persons' feelings into account, and the like.
- Anger control: Training in techniques for identifying triggers and cues that arouse anger and maintaining self-control.
- Moral reasoning: Activities designed to improve the ability to reason about right and wrong behavior and raise the level of moral development.
- Victim impact: Activities aimed and getting offenders to consider the impact of their behavior on their victims.
- Substance abuse: Application of any of the typical CBT techniques specifically to the issue of substance abuse.
- Behavior modification: Behavioral contracts and/or reward and penalty schemes designed to reinforce appropriate behavior.

- Relapse prevention: Training on strategies to recognize and cope with high-risk situations and halt the relapse cycle before lapses turn into full relapses.
- Individual attention: Any individualized one-on-one treatment element that supplements CBT group sessions, e.g., individual counseling.

As Table 3 shows, the presence of some of these treatment elements in the CBT program was significantly related to effect sizes. The strongest relationship appeared for individual attention, followed by anger control and cognitive restructuring.

The Relative Influence of Different Moderator Variables

The results presented in Table 3 identify a number of variables describing the participants and the CBT interventions that are related to treatment effects with key method variables controlled. Each of these moderator variables represents a way to differentiate the circumstances of CBT treatment that yield larger and smaller effects on recidivism. The variable-by-variable results in Table 3, however, do not tell us about the relative influence of the different moderators. To examine the independent relationships of these variables with the others taken into account, two summary random effects regression analyses were conducted. These were configured to model the treatment effect sizes as a function of participant characteristics, the amount of CBT, the quality of the CBT, and the specific type of CBT, with method differences controlled.

Drawing on the results in Table 3, the relevant participant characteristics were represented by recidivism risk, the only variable in that set significantly related to effect size. The amount of CBT was represented by the combination of variables previously designated for that purpose—sessions per week, length in weeks, and their interaction. The quality of the CBT implementation was represented by the composite implementation factor, also described earlier. The type of CBT was represented in the first analysis as the set of brand name categories (with the two “other” categories omitted as a reference set). In the second it was represented in terms of the specific treatment elements identified as present in the intervention. In both analyses, the CBT emphasis variable was also included to add information about the primacy of CBT in the overall intervention.

Table 5 shows the results when the CBT was represented in brand name categories. Once again, no specific type of CBT program had effects significantly different from the mean of all the other types. Only two moderator variables were individually significant in this analysis—recidivism risk (higher risk was associated with larger effects) and the composite implementation factor (higher quality implementation was associated with larger effects).

Table 6 shows the parallel analysis with the CBT intervention represented in terms of treatment elements. As in the previous analysis, recidivism risk and high quality implementation were associated with better outcomes. In addition, however, four of the individual treatment elements showed significant relationships with effect size. Interpersonal problem solving and anger control were positively related; their presence was associated with larger effects on recidivism. Victim impact and behavior modification were negatively related; they were associated with worse outcomes.

Effects of “Best Practice” CBT on Recidivism

We can use the multiple regression analysis in Table 6 to explore optimal CBT treatment circumstances by predicting the effect size expected in a favorable scenario. For this purpose we assumed the best quality study method and measurement characteristics (no design problems, zero attrition, intent-to-treat analysis, and an arrest recidivism outcome). We also assumed the subject sample was comprised of moderately high risk offenders who received the median number of sessions per week (two) with high quality implementation over the median number of weeks (16). The CBT treatment assumed was any one of the brand name programs alone (not supplemented with other services), but with anger control and interpersonal problem-solving components included.

When the corresponding variable values are entered into the prediction equation represented in Table 6, the predicted effect size is a logged odds ratio of 1.05, corresponding to an odds ratio of 2.86. Compared to a control group recidivism of .40 (the overall mean), this represents a decrease to a recidivism rate of .19 in the treatment group, that is, a 52% decrease overall. This impressive effect is not a mathematical projection beyond what appears in the data. An odds ratio of 2.86 is at the 82nd percentile of the distribution of effects for the 58 studies in this meta-analysis.

Discussion

This meta-analysis confirmed the findings of positive CBT effects on the recidivism of offenders that have been reported in other recent meta-analyses (Lipsey, Chapman, & Landenberger, 2001; Lipsey & Landenberger, in press; Pearson et al., 2002; Wilson, Bouffard, & MacKenzie, 2005). The mean odds ratio indicated that the odds of not recidivating in the 12 months after intervention for individuals in the treatment group were 1.53 times as great as those for individuals in the control group. This represents a reduction from the .40 mean recidivism rate of the control groups to a mean rate of .30 for the treatment groups, a 25% decrease. The most effective configurations of CBT produced odds ratios nearly twice as large as the mean, corresponding to recidivism rates of around .19 in the treatment groups, more than a 50% decrease from the .40 rate of the average control group.

The main emphasis of this meta-analysis, however, was the search for key moderator variables that would distinguish situations in which CBT produced larger effects from those in which it produced smaller ones. On this issue, there are two themes in the findings. First, a number of variables characterizing the subject samples, amount and implementation of CBT, and the CBT treatment elements were significantly correlated with the effect sizes for recidivism outcomes. In this regard, there are numerous moderators of the treatment effects. These are not all independent relationships, however. Intervention studies tend to come with bundles of co-occurring characteristics that are, therefore, correlated with each other across studies. This confounding of moderator variables with each other makes it difficult to identify those most critical to the outcome (Lipsey, 2003).

Application of multiple regression analysis to identify the moderator variables with the strongest independent relationships to effect size led to the second theme in our findings. Of the many

study characteristics that showed significant relationships with effect size, relatively few remained significant when the influence of the others was taken into account. The net result was that much of the variation in recidivism effects could be explained by a small number of moderator variables. The only factors independently related to the effect sizes were (a) the risk level of the participating offenders, (b) how well the treatment was implemented, and (c) the presence or absence of a few treatment elements. In the latter category, inclusion of anger control and interpersonal problem solving components in the treatment program were associated with larger effects; inclusion of victim impact and behavior modification were associated with smaller effects. Most striking was that, controlled for other moderators, none of the major CBT brand name programs produced effects on recidivism that were significantly larger than the average effects of the other programs.

Though not informative for purposes of identifying the most effective treatment conditions, the relationships between characteristics of the study methods and the effects sizes were nonetheless interesting. The aspect of method that is usually of greatest concern for intervention studies is whether a randomized design was used. For the studies included in this meta-analysis, however, there were no significant effect size differences between randomized and nonrandomized designs. Only the intent-to-treat variable, indicating whether treatment dropouts were included in the outcome measures, was significantly related to effect size and that relationship dissipated when other moderators were included in the analysis.

Implications for Practice

With the key participant and general implementation characteristics controlled, no significant differences were found in the effectiveness of the different types or “brand names” of CBT. It thus appears to be the general CBT approach, and not any specific version, that is responsible for the overall positive effects on recidivism. Within that framework, inclusion of distinct anger control and interpersonal problem solving components in the CBT program enhance the effects while victim impact and behavior modification components appear to diminish it.

What seems to most strongly characterize effective CBT programs is high quality implementation as represented by low proportions of treatment dropouts, close monitoring of the quality and fidelity of the treatment implementation, and adequate CBT training for the providers. These characteristics are more closely associated with research and demonstration programs than with those implemented in routine practice. This is an encouraging picture from the standpoint of practice. It suggests that any representative CBT program that is well-implemented might have results in practice that approach the very positive effects on recidivism produced by the most effective programs documented in the available research studies.

It is also encouraging that the effects of CBT were greater for offenders with higher risk of recidivism than those with lower risk ones, contrary to any presumption that higher risk offenders might be less amenable to treatment. The effectiveness of CBT with higher risk offenders is consistent with the principles of effective correctional treatment developed by Andrews et al. (e.g., Andrews & Bonta, 2002; Andrews et al., 1990). They argue that the best results occur when higher-risk offenders receive more intensive services that target criminogenic

needs (e.g., criminal thinking patterns) using cognitive behavioral and social learning approaches.

From a practical standpoint, it is also worth highlighting a couple of variables that were not related to treatment effects once other relevant program characteristics were controlled. In particular, CBT was as effective for juveniles as adults, other things equal, and thus should be useful in both juvenile justice and criminal justice settings. The treatment setting was also not related to treatment effects. Offenders treated in prison (generally close to the end of their sentences) showed recidivism decreases comparable to those of offenders treated in the community (e.g., while on probation, parole or in transitional aftercare).

Implications for Research

Of the 58 studies that met the inclusion criteria for this review, only 19 used random assignment designs and, of those, only 13 maintained sufficiently low attrition from outcome measurement to yield results with high internal validity. Moreover, only six of the random assignment studies were conducted on “real world” CBT practice; the others were research and demonstration programs. The amount of high quality research on CBT in representative correctional practice is not yet large enough to determine whether the impressive effects on recidivism found in this meta-analysis can be routinely attained under everyday circumstances.

Though generalization to routine practice cannot be assured, the consistency and magnitude of the effects found in the research to date leave little doubt that CBT is capable of producing significant reductions in the recidivism of even high risk offenders under favorable conditions. However, much remains to be learned about the optimal configuration of CBT and the conditions under which it is most effective. In this meta-analysis we coded as much detail as possible about the program characteristics and context from the descriptions provided in the research reports. At best, those descriptions were limited and fell well short of providing full information about critical program details. An important direction for future research is to better differentiate and document the dimensions along which CBT varies in different applications and to identify the characteristics most critical for attaining optimal effects. The central issue for research on CBT with offender populations at this juncture is not to determine if it has positive effects, but to determine when and why it has the most positive effects.

References

* References marked with an asterisk are reports of studies included in the meta-analysis.

*Anderson, J. (2002). Overview of the Illinois DOC high-risk parolee re-entry program and 3-year recidivism outcomes of program participants. *Cognitive-Behavioral Treatment Review*, 11(1-2), 4-6.

Andrews, D. A., & Bonta, J. (2002). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically-relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.

- *Arbuthnot, J. & Gordon, D. A. (1986). Behavioral and cognitive effects of a moral reasoning development intervention for high-risk behavior-disordered adolescents. *Journal of Consulting and Clinical Psychology, 54*, 208-216.
- *Armstrong, T. A. (2003). The Effect of Moral Reconciliation Therapy on the Recidivism of Youthful Offenders: A Randomized Experiment. *Criminal Justice & Behavior, 30*(6), 668-687.
- *Barnoski, R. (2002). *Preliminary findings for the Juvenile Rehabilitation Administration's dialectic behavior therapy program* (Research Rep. No. 02-07-1203). Olympia, WA: The Evergreen State College, Washington State Institute for Public Policy.
- *Barnoski, R. (2002). *Washington State's implementation of aggression replacement training for juvenile offenders: Preliminary findings*. Olympia, WA: The Evergreen State College, Washington State Institute for Public Policy.
- *Berry, S. (1998). *An evaluation of the Montgomery House Violence Prevention Program*. Hamilton, NZ: Psychological Service, Department of Corrections.
- *Bonta, J.; Wallace-Capretta, S. & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior, 27*(3), 312-329.
- *Bottcher, J. (1985). *The Athena program: An evaluation of a girl's treatment program at the Fresno County Probation Department's juvenile hall*. Sacramento, CA: California Youth Authority.
- *Burnett, W. L. (1997). Treating post-incarcerated offenders with moral reconnection therapy(r): A one-year recidivism study. *Cognitive-Behavioral Treatment Review, 6*, 2.
- *Bush, J. (1995). Teaching self-risk management to violent offenders. In McGuire, J. (Ed.), *What works: Reducing reoffending* (pp. 139-154). New York, NY: John Wiley & Sons.
- Bush, J., Glick, B., & Taymans, J. (1997, revised 1998). *Thinking for a change: Integrated cognitive behavior change program*. National Institute of Corrections. Washington D.C.: U.S. Department of Justice.
- *Cann, J., Falshaw, L., Nugent, F., & Friendship, C. (2003). Understanding what works: accredited cognitive skills programmes for adult men and young offenders. *Research Finding 226*. London, UK: Home Office.
- *Coughlin, R. A., Cosby, J. & Landenberger, N. A. (2003). *Thinking for a change: Cognitive-behavioral treatment with high-risk offenders on intensive supervision probation*. Preliminary Report. Nashville, TN: Public Safety Collaborative, Tennessee Board of Probation and Parole.
- *Culver, H. E. (1993). *Intentional skill development as an interventional tool (interpersonal problem-solving)* (Doctoral dissertation, University of Massachusetts, 1993). Dissertation Abstracts International, A 54(06), 2053. (University Microfilms No. AAT 9329590).
- *Curulla, V. L. (1991). *Aggression replacement training in the community for adult learning disabled offenders* (Doctoral dissertation, University of Washington, 1991). Dissertation Abstracts International, 53(02-A), 627. (University Microfilms No. AAD92-16113).
- *Deschamps, T. (1998). *MRT: is it effective in decreasing recidivism rates with young offenders?* (Master's thesis, University of Windsor, Ontario, Canada). Master's Abstracts International.
- *Dowden, C., Blanchette, K., & Serin, R. (1999). *Anger management programming for federal male inmates: An effective intervention* (Research Report R-82). Ottawa, Canada: Research Branch, Correctional Service Canada.
- *Falshaw, L., Friendship, C., Travers, R. & Nugent, F. (2003). *Searching for 'what works': an evaluation of cognitive skills programmes* (Research Findings 206). London, UK: Home Office.
- *Finn, P. (1998). *The Delaware Department of Correction life skills program*. Washington, DC: U. S. Department of Justice, Office of Justice Programs.
- *Friedman, A. S., Terras, A., & Glassman, K. (2002). Multimodel substance use intervention program for male delinquents. *Journal of Child and Adolescent Substance Abuse, 11*(4), 43-65.

- *Friendship, C., Blud, L., Erikson, M., Travers, R., & Thornton, D. (2003). Cognitive-behavioral treatment for imprisoned offenders: An evaluation of HM prison service's cognitive skills programmes. *Legal and Criminological Psychology, 8*(1), 103-114.
- *Golden, L. S. (2003). *Evaluation of the efficacy of a cognitive behavioral program for offenders on probation: Thinking for a Change* (Doctoral dissertation, The University of Texas Southwestern Medical Center at Dallas, 2002). Dissertation Abstracts International, 63(10), 4902. (University Microfilms No. (not avail. fr. UMI))
- Goldstein, A. P., & Glick, B. (1987). *Aggression Replacement Training: A comprehensive intervention for aggressive youth*. Champaign, IL: Research Press.
- Goldstein, A. P., & Glick, B. (1994). *The prosocial gang: Implementing Aggression Replacement Training*. Thousand Oaks, CA: Sage.
- *Goldstein, A. P., Glick, B., & Irwin, M. J., Pask-McCartney, C., & Rubama, I. (1989). *Reducing delinquency: Intervention in the community* (1st ed.). New York: Pergamon Press.
- *Grandberry, G. (1998). *Moral Reconciliation Therapy evaluation, final report*. Olympia, WA: Washington State Department of Correction.
- *Greenwood, P. W. & Turner, S. (1993). Evaluation of the Paint Creek Youth Center: A residential program for serious delinquents. *Criminology, 31*(2), 263-279.
- *Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: 2. Intervention. *Developmental Psychology, 26*, 269-277.
- Haddock, C. K., Rindskopf, D., Shadish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods, 3*(3), 339-353.
- *Hall, E., Prendergast, M. L., Wellisch, J., Patten, M., & Cao, Y. (2004). Treating drug-abusing women prisoners: An outcome evaluation of the Forever Free program. *The Prison Journal, 84*(1), 81-105.
- *Hanson, G. (2000). *Pine Lodge Intensive Inpatient Treatment Program*. Olympia, WA: Planning and Research Section, WA State Department of Correction.
- *Henning, K. R., & Frueh, B. C. (1996). Cognitive-behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Corrections' cognitive self-change programme. *Criminal Justice and Behavior, 23*, 523-541.
- *Hughey, R., & Klemke, L. W. (1996). Evaluation of a jail-based substance abuse treatment program. *Federal Probation, 60*(4), 40-44.
- *Johnson, G. & Hunter, R. M. (1995). Evaluation of the specialized drug offender program. In R. R. Ross, & R. D. Ross (Eds.), *Thinking straight* (pp. 215-234). Ottawa, Canada: AIR Publications.
- *Kirkpatrick, B. L. H. (1996). *Cognitive restructuring: Effects on recidivism* (Doctoral dissertation, Ball State U.). Dissertation Abstracts International, 57(6A), 2680. (University Microfilms No. AAM9632826).
- *Knott, C. (1995). The STOP Programme: Reasoning and rehabilitation in a British setting. In J. McGuire (Ed.), *What works: Reducing reoffending* (pp. 115-126). Chichester, UK: John Wiley and Sons.
- *Kownacki, R. J. (1995). The effectiveness of a brief cognitive-behavioral program on the reduction of antisocial behavior in high-risk adult probationers in a Texas community. In Ross, R. R. & R. D. Ross (Eds.), *Thinking straight: The reasoning and rehabilitation program for delinquency prevention and offender rehabilitation* (pp. 249-257). Ottawa, Canada: Air Training and Publications.
- *Landenberger (2004). *Evaluation of the Tennessee Bridges Program*. Report to the Reentry Steering Committee. Nashville, TN: Tennessee Department of Correction.

- *Larson, K. A. (1989). Problem-solving training and parole adjustment in high-risk young adult offenders. In S. Duguid (Ed.), *Yearbook of correctional education: 1989* (pp. 279-299). Burnaby, BC, Canada: Simon Fraser University.
- *Leeman, L. W., Gibbs, J., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. *Aggressive Behavior, 19*, 281-292.
- Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science, 587*, 69-81.
- Lipsey, M. W., Chapman, G., & Landenberger, N. A. (2001). Cognitive-behavioral programs for offenders. *The Annals of the American Academy of Political and Social Science, 578*, 144-157.
- Lipsey, M. W. & Landenberger, N. A. (in press). Cognitive-behavioral interventions: A meta-analysis of randomized controlled studies. In B. C. Welsh & D. P. Farrington (eds.), *Preventing crime: What works for children, offenders, victims, and places*. New York: Springer.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Little, G. L., & Robinson, K. D. (1986). *How to escape your prison*. Memphis, TN: Eagle Wing Books.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1993). Cognitive behavioral treatment of felony drug offenders: A five-year recidivism report. *Psychological Reports, 73*(3), 1089-1090.
- *McCracken, L., Hearn, C., & Stuckey, S. (2003). Juvenile DWI/Drug Court Albuquerque, NM *Cognitive-Behavioral Treatment Review, 12*(1), 8-9.
- *Motiuk, L., Smiley, C., & Blanchette, K. (1996). Intensive programming for violent offenders: A comparative investigation. *Forum on Corrections Research, 8*(3), 10-12.
- *Myers, W. C., Burton, P. R. S., Sanders, P. D., Donat, K. M., Cheney, J., Fitzpatrick, T. M., & Monaco, L. (2000). Project Back-on-Track at 1 year: A delinquency treatment program for early-career juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(9), 1127-1134.
- NIC (National Institute of Corrections) (1996). *Cognitive interventions program: Think*. National Institute of Corrections Information Center. Washington D.C.: U.S. Department of Justice.
- Pearson, F. S., Lipton, D. S., Cleland, C. M., & Yee, D. S. (2002). The effects of behavioral/cognitive-behavioral programs on recidivism. *Crime and Delinquency, 48*(3), 476-496.
- *Pelissier, B. M., Rhodes, W., Saylor, W. G., Gaes, G. G., Camp, S.D., Vanyur, S.D., and Wallace, S. (2001). TRIAD drug treatment evaluation project. *Federal Probation, 65*(3), 3-7.
- *Porporino, F. J., Fabiano, E. A., & Robinson, D. (1991). *Focusing on successful reintegration: Cognitive skills training for offenders* (Research Rep. No. 19). Ottawa, Canada: Correctional Service of Canada.
- *Pullen, S. (1996). *Evaluation of the Reasoning and Rehabilitation cognitive skills development program as implemented in juvenile ISP in Colorado*. Boulder, Colorado: Division of Criminal Justice.
- Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301-321). New York: Russell Sage Foundation.
- *Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders* (Research Rep. No. R-41). Ottawa, Canada: Correctional Service Canada.
- *Robinson, D., Grossman, M., & Porporino, F. (1991). *Effectiveness of the cognitive skills training program. From pilot project to national implementation*. (Research Report No. B-07). Ottawa: Correctional Services of Canada, Research and Statistics Branch.
- *Robinson, S. C. (1995). *Implementation of the cognitive model of offender rehabilitation and delinquency prevention* (Doctoral dissertation, University of Utah, 1994). Dissertation Abstracts International, 55(8-A), 2582. (University Microfilms No. AAT9502199).

- Ross, R. R., & Fabiano, E. A. (1985). *Time to think: A cognitive model of delinquency prevention and offender rehabilitation*. Johnson City: Institute of Social Sciences and Arts.
- *Ross, R. R., Fabiano, E. A., & Ewles, C. D. (1988). Reasoning and rehabilitation. *International Journal of Offender Therapy and Comparative Criminology*, 32, 29-35.
- *Shivrattan, J. L. (1988). Social interactional training and incarcerated juvenile delinquents. *Canadian Journal of Criminology*, 30, 145-163.
- *Spencer, E. G., Jr. (1986). *An outcome study of social-behavioral skill development in male adolescent delinquency* (Doctoral dissertation, The California School of Professional Psychology, San Diego, 1986). Dissertation Abstracts International, 47(5), 1891A. (University Microfilms No. 8618153).
- *State of Washington, Department of Corrections (1998, Jan.). *Evaluation of the Work Ethic Camp*. Report to the Legislature. Olympia, WA: Washington State Department of Corrections.
- *T3 Associates under contract (1999). *An Outcome Evaluation of CSC Substance Abuse Programs: OSAPP, ALTO and Choices*. Final Report. Volume 1. Correctional Services Canada. Ottawa, Canada.
- *Thambidurai, G. A. (1980). *A comparative outcome study of a contract parole program for individuals committed to the youth correctional complex in the state of New Jersey* (Doctoral dissertation, Rutgers University, The State University of New Jersey, 1980). Dissertation Abstracts International, 41(01), 371B. (University Microfilms No. 80-16503).
- *Van Voorhis, P., Spruance, L. M., Ritchey, P. N., Johnson-Listwan, S., & Seabrook, R. (2004). The Georgia cognitive skills experiment: A replication of Reasoning and Rehabilitation. *Criminal Justice and Behavior*, 31(3), 282-305.
- *Van Voorhis, P., Spruance, L. M., Ritchie, P. N., Johnson-Listwan S. Seabrook, R., & Pealer, J. (unknown year). *The Georgia cognitive skills experiment. Outcome evaluation phase II*. Atlanta, Georgia: Georgia Board of Pardons and Paroles and National Institute of Justice. (NCJRS Document Reproduction Service No. Grant #98-CE-VX-0013).
- *Walters, G. D. (1999). Short-term outcome of inmates participating in the lifestyle change program. *Criminal Justice and Behavior*, 26(4), 322-337.
- 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.
- Wilson, D. B., Bouffard, L. A., & MacKenzie, D. L. (2005). A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Journal of Criminal Justice and Behavior*. 32(2), 172-204.
- Yochelson, S., & Samenow, S. E. (1976). *The criminal personality: Vol I. A profile for change*. New York, NY: Aronson.

Table 1: Characteristics of the Studies Included in the Meta-Analysis

	<i>N</i>	<i>%</i>		<i>N</i>	<i>%</i>
Publication type			Program studied		
Journal	19	33	Practice	31	53
Chapter	7	12	Demonstration	18	31
Technical report	25	43	Research	9	16
Thesis	7	12	Treatment setting		
Year of publication			Correctional institution	27	47
1980-1990	10	17	Community	31	53
1991-2000	31	53	Treatment sessions/week		
2001-2004	17	29	1	18	31
Country			2	17	29
USA	42	72	3	8	14
Canada	10	17	4-5	10	17
UK	5	9	6-10	5	9
New Zealand	1	2	Treatment length		
Design			5-10 wks	12	21
Randomized	19	33	11-20 wks	26	45
Matched	23	40	21-40 wks	13	22
Neither	16	28	41-104 wks	7	12
Design problem			Proportion of treatment dropouts		
Yes, favors control	13	22	.00	13	22
No or not noted	41	71	.01-.10	6	10
Yes, favors treatment	4	7	.11-.20	18	31
Attrition from posttest			.21-.30	8	14
.00	37	64	> .30	13	22
.01-.10	7	12	CBT treatment type		
.11-.30	8	14	Reasoning & Rehabilitation	15	26
> .30	6	10	Moral Reconciliation Therapy	11	19
Intent to treat			Aggression Replacement Therapy	6	10
Yes, Tx dropouts included	49	84	Interpersonal Problem Solving Therapy	4	7
Cannot tell	4	7	Thinking for a Change	5	9
No, Tx dropouts not included	5	9	Substance abuse focus	5	9
Type of recidivism			Other manualized	9	16
Rearrest	29	50	All other	3	5
Reconviction	20	34	CBT Emphasis		
Incarceration	8	14	CBT with other services	11	19
Other	1	2	CBT with some other Tx elements	11	19
			CBT only	36	62

*Continued on next page**Continued on next page*

Recidivism interval			CBT treatment elements indicated*		
1-5 mo	2	3	Cognitive skills	45	78
6 mo	9	16	Interpersonal problem solving	45	78
7-11 mo	5	9	Social skills	43	74
12 mo	29	50	Cognitive restructuring	37	64
13-24 mo	9	16	Anger control	20	35
25-36 mo	4	7	Substance abuse	19	33
Sample size			Moral reasoning	17	29
14-50	10	17	Relapse prevention	15	26
51-100	8	14	Behavior modification	11	19
101-200	14	24	Individual attention	10	17
201-500	11	19	Victim impact	7	12
501-3000	15	26	<i>* multiple elements, not mutually exclusive</i>		
Sample age			Implementation monitoring		
Juvenile	17	29	None indicated	17	29
Adult	41	71	Minimal	20	35
Percent male			Good	17	29
0	3	5	Very good	4	7
50	2	3	CBT training for providers		
70-98	11	19	Minimal	31	53
100	36	62	Moderate	14	24
Not reported	6	10	Extensive	13	22
Percent minority			Mental health background of		
0-25	12	21	None or minimal		40
26-50	9	16	Moderate	7	12
51-75	12	21	Extensive	11	19
76-100	4	7			
Not reported	21	36			
Recidivism risk rating					
Low	18	31			
Low-medium	9	16			
Medium	18	31			
Medium-high	7	12			
High	6	10			

Table 2: Correlations between Study Method Characteristics and Recidivism Effect Sizes (N=58)

Method Variable	Correlation	<i>p</i>
Design		
Randomized (no/yes)	.04	.77
Matched (no/yes)	-.03	.80
Neither (no/yes)	.00	.98
Design problem (favors control/no/favors Tx)	.19	.14
Attrition proportion	.12	.35
Intent to treat (yes/no)	-.24*	.06
Type of recidivism		
Rearrest (no/yes)	.10	.44
Reconviction (no/yes)	-.04	.77
Incarceration (no/yes)	-.08	.57
Other (no/yes)	-.02	.90
Recidivism interval		
Linear	-.01	.93
Log	-.04	.74

Note: weighted random effects analysis * $p < .10$ ** $p < .05$

Table 3: Relationships of Participant and Intervention Characteristics to Effect Size with Selected Method Variables Controlled

Study Characteristic	Beta with Method Controls^a
General Study Characteristics	
Country: U.S.(1) vs Canada/UK/NZ(2)	-.03
Publication type: report/thesis (1) vs journal/chapter (2)	.13
Year of publication	-.11
Participant Characteristics	
Juveniles(1) /adults(2)	-.03
% male	-.07
% minority	.16
Recidivism risk rating	.27**
CBT Amount	
Sessions per week	.34**
Hours per week (logged)	.23*
Total hours of treatment (logged)	.38**
Length in weeks (logged)	-.03
Sessions per week x Length in weeks (logged)	-.08
Quality of CBT Implementation	
Proportion of Tx dropouts	-.28**
Implementation monitoring	.20
CBT training for providers	.21
Mental health background of providers	-.07
Practice(1)/demonstration(2)/research(3) program	.31**
Composite implementation factor	.40**
Other Program Characteristics	
Treatment setting: prison(1) /community(2)	.20
CBT emphasis: with other components (1)/ CBT alone (3)	-.30**

Table 3 continued on next page

Table 3, continued

Specific CBT Program	
Reasoning & Rehabilitation	-.21
Moral Reconciliation Therapy	.04
Aggression Replacement Therapy	.16
Interpersonal Problem Solving Therapy	-.09
Thinking for a Change	.12
Substance abuse focus	.00
Other manualized	.02
All other	.01
CBT Treatment Elements	
Cognitive skills	.02
Cognitive restructuring	.27**
Interpersonal problem solving	.04
Social skills	.02
Anger control	.32**
Moral reasoning	.11
Victim impact	-.14
Substance abuse	.11
Behavior modification	.03
Relapse prevention	.12
Individual attention (in addition to group sessions)	.39**

Note: Beta values from random effects multiple regression.

(a) controlling for design problems, attrition proportion, intent-to-treat comparison, and arrest recidivism.

* $p < .10$ ** $p < .05$

Table 4: Correlations Between Potential Moderator Variables Related to the Quality of CBT Implementation (N=58)

	Proportion of treatment dropouts	Implementation monitoring	CBT training for providers	Mental health background of providers
Implementation monitoring	-.17			
CBT training for providers	-.17	.40**		
Mental health background of providers	.08	-.07	.13	
Practice- demonstration- research program	-.29**	.44*	.23*	.24*

* $p < .10$ ** $p < .05$

Table 5: Regression Model for Effect Size Moderators Using Specific Type of CBT**Program**

Variables in the Model^a	B	z	p	Beta
Method Controls				
Design problem	.11	1.02	.31	.14
Attrition proportion	-.13	-.21	.83	-.03
Intent to treat	-.13	-1.21	.23	-.19
Arrest recidivism	.13	1.04	.30	.15
Participant Characteristics				
Recidivism risk rating **	.19	1.99	.05	.26
CBT Amount				
Sessions per week	.05	1.21	.23	.22
Length in weeks (logged)	.04	.36	.72	.06
Sessions x length	.03	.73	.46	.12
Quality of Implementation				
Composite implementation factor **	.26	2.93	.00	.45
Other Program Characteristics				
CBT emphasis	-.10	-.90	.37	-.19
Specific CBT Program				
Reasoning & Rehabilitation	-.01	-.10	.92	-.02
Moral Reconciliation Therapy	.16	.99	.32	.15
Aggression Replacement Therapy	-.09	-.35	.73	-.05
Interpersonal Problem Solving	-.31	-.82	.41	-.10
Thinking for Change	.00	.02	.99	.00
Substance abuse focus	-.19	-.93	.35	-.15

a. Weighted, random effects multiple regression analysis with inverse-variance weights.

* $p < .10$ ** $p < .05$

Table 6: Regression Model for Effect Size Moderators Using CBT Treatment Elements

Variables in the Model^a	B	z	p	Beta
Method Controls				
Design problem	-.02	-.27	.79	-.03
Attrition proportion	.08	.12	.90	.01
Intent to treat	.03	.30	.77	.05
Arrest recidivism	.01	.08	.94	.01
Participant Characteristics				
Recidivism risk rating **	.20	2.83	.00	.27
CBT Amount				
Sessions per week	.01	.37	.71	.07
Length in weeks (logged)	-.03	-.35	.72	-.05
Sessions x length	.04	.74	.46	.13
Quality of Implementation				
Composite implementation factor *	.14	1.82	.07	.23
Other Program Characteristics				
CBT emphasis *	-.20	-1.84	.07	-.41
CBT Treatment Elements				
Cognitive skills	-.26	-1.23	.22	-.26
Cognitive restructuring	.13	.84	.40	.16
Interpersonal problem solving **	.28	2.16	.03	.32
Social skills	.19	1.23	.22	.19
Anger control **	.32	2.23	.03	.36
Moral reasoning	-.03	-.17	.87	-.03
Victim impact **	-.45	-2.36	.02	-.31
Substance abuse	.13	.87	.39	.16
Behavior modification *	-.29	-1.70	.09	-.31
Relapse prevention	-.19	-1.32	.19	-.19
Individual attention	.07	.37	.71	.06

a. Weighted, random effects multiple regression analysis with inverse-variance weights.

* $p < .10$ ** $p < .05$

Adhering to the Risk and Need Principles: Does It Matter for Supervision-Based Programs?

Christopher T. Lowenkamp, Ph.D.

Center for Criminal Justice Research, University of Cincinnati

Jennifer Pealer

Department of Political Science & Criminal Justice

Northern Kentucky University

Paula Smith

Division of Criminal Justice, University of Cincinnati

Edward J. Latessa

Division of Criminal Justice, University of Cincinnati

IN THE PAST 20 YEARS, there has been a re-emergence of interest in the effectiveness of correctional treatment programs for offenders. This interest has led to the development of the principles of effective interventions (Gendreau, 1996; Gendreau, French, & Taylor, 2002). Research has now shown a link between these program characteristics and effectiveness (Andrews & Dowden, 1999; Lipsey & Wilson, 1995; Gendreau, 1996; Lowenkamp, 2004; Lowenkamp, Latessa, and Smith, 2006). However, most of these studies have examined traditional residential treatment programs. Therefore, the question remains: Do these principles apply to community non-residential programs such as intensive supervision probation? The current study examines the effects of program characteristics on recidivism using a sample drawn from community non-residential programs to determine if the risk and need principles apply to traditional supervision-oriented programs such as intensive supervision probation, electronic monitoring, day reporting, and work release.

Risk, Need, and Treatment Principles

In 1996, Gendreau introduced several principles of effective interventions. These prin-

ciples may be collapsed into risk, need, responsivity, and treatment. While each is equally important to the provision of sound correctional interventions, we focus on the risk and need principles in this paper. As such, only the risk and need principles are reviewed below; however, readers are encouraged to review other principles related to effective correctional interventions (for a review see Gendreau, 1996; Gendreau, et al., 2002).

The risk principle states that programming should be matched to the risk level of the offenders (Andrews, Bonta, & Hoge, 1990), and higher-risk offenders should receive more intensive programming for longer periods of time to reduce their risk of re-offending. Moreover, and equally important, applying intensive treatment to low-risk offenders may actually serve to increase their risk of recidivism (Andrews, Bonta, and Hoge, 1990 and Lowenkamp & Latessa, 2005). Much research has found support for the risk principle. For example, a meta-analysis conducted by Andrews and Dowden (1999) found that programs that adhere to the risk principle reduced recidivism by 19 percent but programs that violated the risk principle increased recidivism by 4 percent. Similarly, a study of intensive rehabilitation

supervision by Bonta, Wallace-Capretta, and Rooney (2000) found a 20 percent reduction in recidivism for higher-risk offenders that received more intensive supervision, but a 17 percent increase for lower-risk offenders. A more recent examination of the risk principle was conducted by Lowenkamp and Latessa (2005) using a sample of adult halfway house participants. Lowenkamp and Latessa found that these intensive programs worked for higher-risk offenders and led to reductions in recidivism from 10 to 30 percent. However, most of these same programs increased recidivism for lower-risk offenders. While the type of offender placed in a correctional program is certainly related to program effectiveness, what a program targets while the offender is in the program is equally important. The need principle, discussed below, gives programs strong guidance regarding what offender needs should be targeted to reduce the propensity of criminal behavior.

Simply put, the need principle identifies appropriate needs to be targeted by correctional interventions in attempting to reduce offender recidivism (Andrews, et al., 1990; Gendreau, 1996). Research has consistently identified certain dynamic correlates of criminal behavior (also known as crimi-

nogenic needs) such as antisocial attitudes, antisocial peers, antisocial personality, poor familial relationships, and low educational or vocational achievement (Gendreau, et al., 1996; Simourd and Andrews, 1994). Research has also indicated that if a correctional intervention or program targets these dynamic risk factors, the reductions in recidivism follow (Dowden & Andrews, 1999a). In a more recent study, Gendreau, et al. (2002) found that the density of criminogenic needs targeted was strongly related to program effectiveness in reducing offender recidivism. Specifically, programs that targeted 4 to 6 more criminogenic than non-criminogenic needs reduced recidivism, on average, by about 30 percent. Programs that targeted 1 to 3 more criminogenic than non-criminogenic needs were associated with a slight increase in recidivism.

Hence, the research on the risk and need principles indicates that these principles are important to correctional treatment interventions. Intensive treatment programs were more successful in reducing recidivism with higher-risk offenders (Andrews, et al., 1990; Lipsey & Wilson, 1998; Andrews & Dowden, 1999; Lowenkamp & Latessa, 2005). Furthermore, when programs targeted more criminogenic needs, recidivism declined more there (Dowden & Andrews, 1999b; Gendreau, et al., 2002). However, the question remains: "Are the risk and need principles related to the effectiveness of supervision-based correctional interventions in reducing recidivism?"

Research on Supervision-Oriented Programming

There has been some research that indirectly tests the relationship between the characteristics of supervision-based interventions and effectiveness. This research, in summary, did find support for the relationship between treatment and effectiveness for supervision-oriented programs (Petersilia & Turner, 1993; Fulton, Gendreau, Paparozzi, 1996; Bonta et al, 2000; Fulton, Stone & Gendreau, 1994; Aos, Miller & Drake, 2006). For example, in a review of three types of programs within a probation department in Colorado, Johnson and Hunter (1992) found that offenders who received ISP with the cognitive component had lower recidivism rates than offenders who participated in only the supervision probation component. Furthermore, in a multi-site evaluation of ISPs conducted by the RAND Corporation, Petersilia and Turner (1993) found that high-

er levels of program participation (measured as any employment, any counseling sessions, any community service, and any restitution paid) were associated with a 10 to 20 percent reduction in recidivism.

A recent meta-analysis conducted by Aos, Miller, and Drake (2006) examined the effectiveness of various correctional programs and supervision. They systematically reviewed 34 studies of intensive supervision probation programs that have been conducted within the last 35 years. The analysis revealed that ISPs that incorporated some treatment resulted in an average reduction of 21.9 percent, whereas ISPs that were surveillance-oriented had no impact on recidivism. Accordingly, while research has found that non-residential programs such as ISPs may be effective in reducing recidivism if they incorporate treatment into the services delivered, the exact characteristics that are necessary to reduce recidivism have not yet been tested empirically.

Method

The current study examined 66 community-based correctional programs to determine if adherence to the risk and need principle enhanced effectiveness in reducing recidivism. These programs were jail and prison diversion programs funded by the Community Corrections Act (CCA) in the state of Ohio (for a description of the Community Corrections Act and the programs see <http://www.drc.state.oh.us/web/BCS.HTM>). The participants were offenders sentenced to community-based correctional programs serving 52 counties during the fiscal year 1999. Offenders served by the CCA programs were compared to offenders that were processed as usual in jail, municipal probation, or prison. Offenders from the treatment group were matched to offenders from the comparison group on sex, risk,¹ and county of supervision. Recidivism data was collected on all offenders, with the follow-

¹ Risk level was determined using a risk measure developed in previous research (Lowenkamp and Latessa, 2002) and includes 13 measures including measures of criminal history, current offense, substance abuse, alcohol abuse, marital status, employment, age, and educational attainment. Recidivism rates for the varying categories of risk, based on a two-year follow up, and using incarceration as the outcome measure were: Low risk—7 percent; Low-Moderate risk—22 percent; Moderate risk—38 percent; and High risk—53 percent. For more details and analysis using arrest as the outcome measure see Lowenkamp and Latessa, 2005.

TABLE 1:
Distribution of Sample

Group	N	percent
Prison Diversion		
Day Reporting	6	10.1
Electronic Monitoring	2	3.6
ISP	42	76.4
Work Release	2	3.6
Substance Abuse	3	5.5
Jail Diversion		
Day Reporting	3	27.3
Work Release	1	9.1
Residential Treatment	1	9.1
Domestic Violence	1	9.1
ISP	5	45.5

up time being two years from the date of placement in a CCA program, placement on municipal probation, release from jail, or release from prison.

Programs

Table 1 reviews the different sites that were examined for this study. Two types of programs were used in the current study—prison diversion and jail diversion programs that were funded by the Community Corrections Act.

The prison diversion programs included those offenders that were referred by the local court to a CCA-funded program and participated in the CCA programs for at least 30 days. These offenders are sentenced to a term in prison. That sentence is then suspended and the offenders participate in one or more community-based programs. Of the 66 sites examined, 55 (83.3 percent) were prison diversion programs. Of these programs, the predominant program type was intensive supervision probation (42 programs), followed by day reporting (10.1 percent), substance abuse programs (5.5 percent), electronic monitoring (3.6 percent), and work release (3.6).

The jail diversion programs included those offenders that were placed in programming in lieu of serving time in a jail or as part of their sentence to a jail. Across the various jail diversion programs, the majority of the programs were again intensive supervision probation (5 programs), followed by day reporting (27.3 percent), and then work release, residential treatment, and domestic violence (9.1 percent each).

Offenders

The prison diversion cases were compared to a matched sample of parolees.² A total of 5,781 prison diversion cases were compared to an equal number of parolees. While attempts were made to develop comparison groups from regular felony probation caseloads, this was not always possible. We therefore decided to use parole cases since they provided comparison cases for every program.³ The matched jail diversion cases were compared to jail releases or regular municipal probation cases, depending on the data available within each jurisdiction. We were able to develop jail comparison cases for only three programs (one county). Regular municipal probation cases were used as comparison cases in eight other sites. In total, 707 comparison cases were used as a matched sample for the jail diversion programs (n = 707). Three sites were compared to jail releases, while eight other jail diversion sites were compared to regular municipal probationers.

Table 2 reports the descriptive statistics for the two treatment groups and the comparison cases. For the prison diversion sample, the two groups were relatively similar in racial composition and gender. However, the treatment group was more likely to be single (73 percent) when compared to the comparison group. Furthermore, the comparison group was more likely to have been incarcerated three or more times and was more likely to be under supervision for an offense against a person. When examining the risk category for the offenders, a clear majority of offenders (73 percent) were classified as moderate risk or higher.

When examining the jail diversion sample, we again found the groups similar in regards to race and gender. Sixty-two percent of the treatment group was white compared to 65 percent of the comparison group. Nineteen percent of both groups was female. The groups differ significantly in marital status, prior arrests, prior incarcerations, and offense type. Sixty-one percent of the treatment group was single, with a slightly higher percentage of the comparison group being single (70 percent). Approximately 35 percent of each group had three or more prior

arrests, while roughly 20 percent of each group had at least one prior incarceration. In terms of risk, 78 percent of each group is low to low-moderate, with 20 percent being classified as moderate risk.

Review of Program Level Measures

The current study used four measures of program content. Three measures relate to adherence to the risk principle: higher-risk sample, risk supervision, and risk treatment. One additional measure relates to the need principle: referral ratio. All of these measures were developed from data gathered from a database maintained by the State of Ohio Department of Rehabilitation and Correction.

Higher-risk sample was defined as present for a particular program if 75 percent or more of the sample was moderate or high risk. This measure was included to determine if the program was targeting higher-risk offenders, as is indicated by the risk principle.

The next two measures, risk supervision and risk treatment, were developed to determine, if advised by the risk principle,

if programs were varying the duration of and services received by risk level. Risk supervision was determined to be present if higher-risk offenders were in the program, on average, longer than lower-risk offenders. For the purposes of the risk supervision factor, any difference where the higher-risk group received longer periods of supervision than the lower-risk group was considered to be evidence of meeting this factor. Programs where the lower- and higher-risk groups had equal lengths of supervision or where the lower-risk group had a longer period of supervision did not meet this factor.

Risk treatment was determined to be present for a particular program if, on average, higher-risk offenders received at least one-half more referrals for services than lower-risk offenders. For example, if the higher-risk offenders, on average, were referred to 2.5 programs and the lower-risk offenders were referred to 2.0 or fewer programs, this criterion was considered to be met by the program.

Finally, we included a measure relating to the need principle, which tapped the density of services targeting criminogenic needs. This measure was a ratio of referrals target-

TABLE 2:
Descriptive Statistics for Treatment and Comparison Cases

Variable	Prison Diversion Cases				Jail Diversion Cases			
	Treatment		Parole		Treatment		Jail/Probation	
	N	percent	N	percent	N	percent	N	percent
Variable								
White	2,454	48	2,300	45	438	62	460	65
Female	358	7	358	7	134	19	134	19
Single	3,732	73	3,323	65	431	61	495	70
Prior arrest								
0	272	5	736	14	169	24	260	37
1-2	1,679	33	1,241	24	262	37	212	30
3+	3,161	62	3,135	61	276	39	235	33
Prior incarceration								
0	3,219	63	2,336	46	583	83	534	75
1-2	1,629	32	1,724	34	101	14	134	19
3+	264	5	1,052	21	23	3	39	6
Offense type								
Person	730	14	1,318	26	176	25	155	22
Sex	153	3	153	3	1	0	1	0
Drug	1,647	32	1,444	29	22	4	57	8
Property	1,847	36	1,746	35	70	14	62	9
Other	735	14	379	8	438	62	432	61
Risk category								
Low	235	5	235	5	186	26	186	26
Low-moderate	1,192	23	1,192	23	374	52	374	52
Moderate	3,147	62	3,147	62	142	20	142	20
High	538	11	538	11	5	1	5	1

² Comparison cases were matched to the treatment cases on gender, county of supervision, and risk category.

³ Alternate analyses using regular felony probation cases were conducted and are reported in the original report by Lowenkamp and Latessa, 2005.

ing criminogenic needs to referrals targeting non-criminogenic needs. For this measure to be considered present, a program had to make three referrals targeting criminogenic needs for every one referral targeting non-criminogenic needs. For example, a program that referred offenders to substance abuse treatment, employment placement, and cognitive behavioral programming and community service would have met this principle, since the first three referrals listed target criminogenic needs while only one, community service, targets non-criminogenic needs.

Outcome measures included any new arrest for jail diversion cases and any new period of incarceration in prison (for a technical violation or new criminal behavior) for prison diversion cases. The outcome measures differed due to differences in the populations served. Jail diversion cases tend to be lower-level offenders that are not subject to prison for the current offense and often lack a history of incarceration. The base rate of return to prison for this group was fairly low. Therefore, we selected an alternate measure to use for the jail diversion cases. The follow-up time period was consistent across all groups and lasted for two years.

Analysis

For each site, a correlation co-efficient, or r-value, was calculated that represented the magnitude of the relationship between program participation and recidivism. The r-value can be interpreted as the percentage difference in recidivism rates between the treatment (offenders participating in the CCA program) and comparison (offenders on parole, released from jail, or on municipal probation) groups (see Rosenthal, 1991 and Gendreau, Goggin, and Paparozzi, 1996). For example, if the treatment group from hypothetical program A had a 40 percent recidivism rate and the matched comparison group had a 50 percent recidivism rate, an r-value of .10 would be generated (since 50 percent or .50 minus 40 percent or .40 equals .10). Positive r-values indicate recidivism rates that favor the treatment group—that is, where the recidivism rate of the treatment group was lower than that of the comparison group. The opposite is true for negative r-values. Negative r-values favor the comparison group or indicate programs where the treatment group participants had higher recidivism rates than the comparison group. For example, a -.10 would indicate a program

where the program participants (treatment group) had a 60 percent recidivism rate (or .60) and the comparison group had a 50 percent recidivism rate (or .50).

We categorized each program based on whether it met the factors listed in the measures section which related to the risk and need principles (high-risk sample, risk treatment, risk supervision, and referral ratio). We then calculated the average correlation coefficient for the programs based on that categorization.

Results

Figure 1 reveals the r-values for the programs categorized by whether they met the risk and need program factors described earlier. The first set of bars represents the average r-values by whether the program met the criterion "higher-risk sample," which again indicated that 75 percent or more of the sample was higher (moderate or high) risk. Only 15 programs met the criteria for higher-risk sample. Programs that met this factor, our proxy measure for targeting higher-risk offenders, resulted in an average decrease in recidivism of 5 percent across the 15 programs. Comparatively, programs that did not adhere to this criterion were associated with a 2 percent increase in recidivism on average.

Our second measure relating to the risk principle was risk supervision. The 19 programs that met this measure were associated with a four percentage point decrease in recidivism. Programs that did not meet this criterion, that is, where the program length

did not vary by risk level, had no impact on recidivism.

The third set of bars represents the average reductions in recidivism based on the "risk treatment" measure. On average, programs where higher-risk offenders received more referrals than lower-risk offenders reduced recidivism by 7 percent. Programs that did not meet this criterion (i.e., lower-risk offenders received more referrals or there was no difference in referrals among risk levels) only saw a 1 percent reduction in recidivism.

Finally, our last measure, referral ratio, which related to the need principle, was associated with program effectiveness. Programs (n = 16) where 75 percent of the referrals were treatment-oriented and targeted criminogenic needs reduced recidivism, on average, by 11 percent. Programs that did not have a 3 to 1 referral ratio favoring services targeting criminogenic needs increased recidivism, on average, by 3 percent.

Prior research has shown that program characteristics have cumulative properties, indicating that as program content and capacity increases, reductions in recidivism are greater (Lowenkamp & Latessa, 2002). Therefore we calculated the average r-value across the four-point factor score. There were 9 sites that did not meet any of the criteria. The average r-value for these sites was -0.13, indicating that these programs were associated with an increase in recidivism rates of 13 percent. When programs (35 sites) met one or two factors, there was a decrease in recidivism of 3 percent. Finally, when programs (n = 4) adhered to three or

FIGURE 1
Average r-value by Risk and Need Principles Program Factors

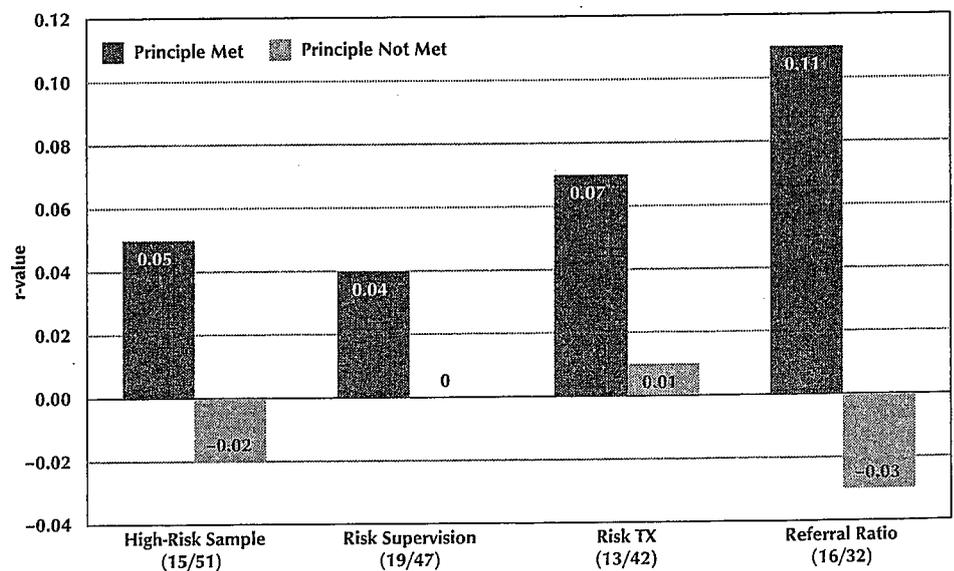
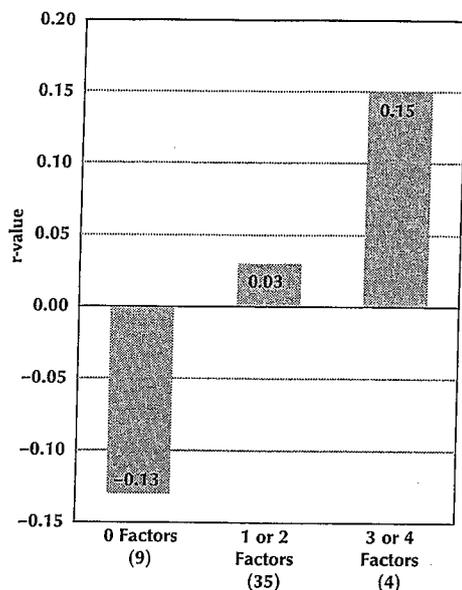


FIGURE 2
Average *r*-value by summed 4 point factor score



more factors, there was a 15 percentage point reduction in recidivism.

Summary

A recent report from the U.S. Department of Justice indicated that the number of offenders under correctional supervision reached an all-time high at the end of 2003 (Glaze, 2004). This continued growth in the offender population causes concern for many agencies, especially given the fact that some recidivism estimates for probation samples are as high as 65 percent (Petersilia, 1985). However, unlike 25–30 years ago, research has identified certain program characteristics that work to reduce the probability of re-offending. While many studies have examined the relationship between programming and recidivism, most of these studies focused on programs that were residential and/or were traditional treatment programs. The current study is one of the first to examine the relationship between program characteristics and effectiveness using community non-residential programs such as intensive supervision probation. The analyses yielded by the current study provide support for the relationship between program characteristics, relating to the risk and need principles, and a program's effectiveness in reducing recidivism. All of the programs in this study were supervision-based programs that differentially adhered to the risk and need principles. The analyses revealed that these intensive programs were more successful for the higher-risk offenders. When at least

75 percent of the population was classified as high risk, there was a 5 percent decrease in recidivism compared to a slight increase in recidivism for programs that incorporated more low-risk offenders. Furthermore, when examining the relationship between risk level and supervision, programs that required higher-risk offenders to be in the programs for a longer period of time saw a 4 percent reduction in recidivism, while those that had a one-size-fits-all approach had no effect on recidivism. Programs that had more referrals for higher-risk offenders reduced recidivism by 7 percent, whereas programs that did not have more referrals for this population only saw a marginal reduction in recidivism. Finally, programs in which 75 percent or more of the referrals were for treatment programming had an 11 percent reduction in returns to prison. Programs in which more than 25 percent of their referrals were non-treatment increased recidivism by 3 percent.

Overall, when examining the cumulative nature of the measures, we found that the more factors a program adhered to the more effective it was in reducing recidivism. Programs that did not meet any of the four criteria increased recidivism by 13 percent, programs that met one to two factors decreased recidivism slightly, and programs that met at least 3 factors decreased recidivism by 15 percent. None of the programs met all four factors.

Based on these findings it appears that the risk and need principles are important factors to consider when developing and/or operating a correctional intervention that is non-residential and traditionally based on supervision. These findings can assist programs in increasing effectiveness and, when taken in the aggregate, public safety. Implementing such strategies is no simple task and would require the adoption and use of a sound risk and need assessment, training of staff, and the availability of relevant and validated treatment programs. While this research does not resolve these issues or tackle these barriers, it does underscore the importance of meeting the risk and need principle when our correctional goal is to reduce recidivism.

References

Andrews, D.A., & Dowden, C. (1999). A meta-analytic investigation into effective correctional intervention for female offenders. *Forum on Corrections Research*, 11, 18–21.

- Andrews, D.A., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19–52.
- Andrews, D.A., Zinger, I., Hoge, R., Gendreau, P., & Cullen, F. (1990b). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369–404.
- Aos, S., Miller, M., & Drake, E. (2006). *Evidence-based adult corrections programs: What works and what does not*. Olympia: Washington State Institute for Public Policy.
- Bonta, J., S. Wallace-Capretta, and J. Rooney, (2000). A Quasi-Experimental Evaluation of an Intensive Rehabilitation Supervision Program. *Criminal Justice and Behavior*, 27, 312–329.
- Community Corrections Act. (Retrieved 5/30/2006). www.drc.state.oh.us/web/bcs.htm
- Dowden, C., & Andrews, D.A. (1999). What works for female offenders: A meta-analytic review. *Crime and Delinquency*, 45, 438–452.
- Fulton, B., Gendreau, P., & Paparozzi, M. (1996). APPA's prototypical intensive supervision program: ISP as it was meant to be. *Perspectives*, 19(2), 25–41.
- Fulton, B., Stone, S. & Gendreau, P. (1994). *Restructuring intensive supervision programs: Applying "what works."* Lexington, KY: American Probation and Parole Association.
- Gendreau, P. (1996). The principles of effective interventions with offenders. In A.T. Harland (ed.), *Choosing Correctional Options that Work: Defining Demand and Evaluating the Supply*. (pp. 117–130). Thousand Oaks, CA: Sage.
- Gendreau, P., French, S., & Taylor, A. (2002). *What works (what doesn't) revised 2002: The principles of effective correctional treatment*. Unpublished manuscript, University of New Brunswick, St. John, New Brunswick, Canada.
- Gendreau, P., Goggin, C., & Paparozzi, M. (1996). Principles of effective assessment for community corrections. *Federal Probation*, 60(3), 64–70.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, 34, 575–607.
- Glaze, L.E. (2004). *Probation and parole in the United States, 2003*. Washington DC: U.S. Department of Justice.
- Johnson, G., & Hunter, R. (1992). *Evaluation of the specialized drug offender program for the Colorado Judicial Department*. Boulder, CO: University of Colorado at Boulder, Center for Action Research.

- Lipsey, M.W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T.D. Cook, H. Cooper, D.S. Cordray, H. Hartmann, L.V. Hedges, R.J. Light (eds.) *Meta-analysis for explanation: A Casebook* (pp. 83-127). New York: Russell Sage.
- Lipsey, M.W., & Wilson, D.B. (1995). Reply to comments on Lipsey and Wilson (1993). *American Psychologist*, 50, 113-115.
- Lipsey, M.W., & Wilson, D.B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D.P. Farrington (eds.), *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions*. (pp. 313-345). Thousand Oaks, CA: Sage.
- Lowenkamp, C.T. (2004). *A program level analysis of the relationship between correctional program integrity and treatment effectiveness*. Unpublished doctoral dissertation, University of Cincinnati.
- Lowenkamp, C.T. & Latessa, E.J. (2002). *Evaluation of Ohio's community based correctional facilities and halfway house programs*. Technical Report. University of Cincinnati.
- Lowenkamp, C.T. & Latessa, E.J. (2005). Increasing the effectiveness of correctional programming through the risk principle: Identifying offenders for residential placement. *Criminology and Public Policy*, 4, 501-528.
- Lowenkamp, C.T., Latessa, E.J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective interventions. Submitted to *Criminology and Public Policy*.
- Petersilia, J. (1985). Probation and felony offenders. *Federal Probation*, 49, 4-9.
- Petersilia, J. & Turner, S. (1993). Evaluating intensive supervision probation/parole: Results of a nationwide experiment. *Research in Brief*. Washington DC: National Institute of Justice.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research (revised)*. Newbury Park: Sage.
- Simourd, L. & Andrews, D.A. (1994). Correlates of delinquency: A look at gender differences. *Forum on Corrections Research*, 6, 26-31.

January 2006

EVIDENCE-BASED ADULT CORRECTIONS PROGRAMS: WHAT WORKS AND WHAT DOES NOT[‡]

In recent years, public policy decision-makers throughout the United States have expressed interest in adopting “evidence-based” criminal justice programs. Similar to the pursuit of evidence-based medicine, the goal is to improve the criminal justice system by implementing programs and policies that have been shown to work. Just as important, research findings can be used to eliminate programs that have failed to produce desired outcomes. Whether for medicine, criminal justice, or other areas, the watchwords of the evidence-based approach to public policy include: outcome-based performance, rigorous evaluation, and a positive return on taxpayer investment.

This report to the Washington State Legislature summarizes our latest review of evidence-based adult corrections programs. We previously published a review on this topic in 2001.¹ In this study, we update and significantly extend our earlier effort.

The overall goal of this research is to provide Washington State policymakers with a comprehensive assessment of adult corrections programs and policies that have a proven ability to affect crime rates.

We are publishing our findings in two installments. In this preliminary report, we provide a systematic review of the evidence on what works (and what does not) to reduce crime. In a subsequent final report, to be published in October 2006, we will extend this analysis to include a benefit-cost estimate for each option.

[‡] Suggested citation: Steve Aos, Marna Miller, and Elizabeth Drake. (2006). *Evidence-Based Adult Corrections Programs: What Works and What Does Not*. Olympia: Washington State Institute for Public Policy.

¹ S. Aos, P. Phipps, R. Barnoski, and R. Lieb (2001). *The Comparative Costs and Benefits of Programs to Reduce Crime*, Olympia: Washington State Institute for Public Policy.

Summary

This study provides a comprehensive review of evidence-based programs for adult offenders. We asked a simple question: What works, if anything, to lower the criminal recidivism rates of adult offenders? To provide an answer, we systematically reviewed the evidence from 291 rigorous evaluations conducted throughout the United States and other English-speaking countries during the last 35 years.

We find that some types of adult corrections programs have a demonstrated ability to reduce crime, but other types do not. The implication is clear: Washington’s adult corrections system will be more successful in reducing recidivism rates if policy focuses on proven evidence-based approaches.

Washington’s Offender Accountability Act

This research was undertaken as part of our evaluation of Washington’s Offender Accountability Act (OAA). Passed in 1999, the OAA affects how the state provides community supervision to adult felony offenders. In broad terms, the OAA directs the Washington State Department of Corrections to do two things:

- 1) Classify felony offenders according to their risk for future offending as well as the amount of harm they have caused society in the past; and
- 2) Deploy more staff and rehabilitative resources to higher-classified offenders and—because budgets are limited—spend correspondingly fewer dollars on lower-classified offenders.

When the Legislature enacted the OAA, it defined a straight-forward goal for the Act: to “reduce the risk of reoffending by offenders in the community.”² To determine whether the OAA results in lower recidivism rates, the Legislature also directed the Washington State Institute for Public Policy (Institute) to evaluate the impact of the Act.³

Whether the OAA is able to affect crime rates will depend, in part, on the policy and programming choices made to implement the Act. As we show in this report, there are some adult corrections programs that have a demonstrated ability to reduce crime, but there are other types of programs that fail to affect crime rates. Given these mixed results, it is reasonable to conclude that the OAA (or any other adult corrections policy initiative) will be successful in reducing crime only if it encourages the implementation of effective approaches and discourages the use of ineffective programs. The purpose of this report is to assist policymakers in sorting through the many evidence-based choices.

The Evidence-Based Review: The Basic Question

The goal of the present study is to answer a simple question: Are there any adult corrections programs that work? Additionally, in order to estimate costs and benefits, we seek to estimate the magnitude of the crime reduction effect of each option.

To answer these fundamental questions, we conducted a comprehensive statistical review of all program evaluations conducted over the last 40 years in the United States and other English-speaking countries. As we describe, we found 291 evaluations of individual adult corrections programs with sufficiently rigorous research to be included in our analysis. These evaluations were of many types of programs—drug courts, boot camps, sex offender treatment programs, and correctional industries employment programs, to name a few.

It is important to note that only a few of these 291 evaluations were of Washington State adult

corrections programs; rather, almost all of the evaluations in our review were of programs conducted in other locations. A primary purpose of our study is to take advantage of all these rigorous evaluations and, thereby, learn whether there are conclusions that can allow policymakers in Washington to improve this state’s adult criminal justice system.

Research Methods

The research approach we employ in this report is called a “systematic” review of the evidence. In a systematic review, the results of *all* rigorous evaluation studies are analyzed to determine if, on average, it can be stated scientifically that a program achieves an outcome. A systematic review can be contrasted with a so-called “narrative” review of the literature where a writer selectively cites studies to tell a story about a topic, such as crime prevention. Both types of reviews have their place, but systematic reviews are generally regarded as more rigorous and, because they assess all available studies and employ statistical hypotheses tests, they have less potential for drawing biased or inaccurate conclusions. Systematic reviews are being used with increased frequency in medicine, education, criminal justice, and many other policy areas.⁴

For this report, the outcome of legislative interest is crime reduction. In particular, since the programs we consider in this review are intended for adult offenders already in the criminal justice system, the specific outcome of interest is reduction in recidivism rates. Therefore, the research question is straightforward: *What works, if anything, to lower the recidivism rates of adult offenders?*

As we describe in the Appendix, we only include rigorous evaluation studies in our review. To be included, an evaluation must have a non-treatment comparison group that is well matched to the treatment group.

² RCW 9.94A.010.

³ The Institute’s first five publications on the Offender Accountability Act are available for downloading at the Institute’s website: www.wsipp.wa.gov. The final OAA report is due in 2010.

⁴ An international effort aimed at organizing systematic reviews is the Campbell Collaborative—a non-profit organization that supports systematic reviews in the social, behavioral, and educational arenas. See: <http://www.campbellcollaboration.org>.

Researchers have developed a set of statistical tools to facilitate systematic reviews of the evidence. The set of procedures is called “meta-analysis,” and we employ that methodology in this study.⁵ In the Technical Appendix to this report (beginning on page 9) we list the specific coding rules and statistical formulas we use to conduct the analysis—technical readers can find a full description of our methods and detailed results.

Findings

The findings from our systematic review of the adult corrections evaluation literature are summarized on Exhibit 1.⁶ We show the expected percentage change in recidivism rates for many types of evaluated adult corrections programs. A zero percent change means that, based on our review, a program does not achieve a statistically significant change in recidivism rates compared with treatment as usual.

We found a number of adult corrections programs that have a demonstrated ability to achieve reductions in recidivism rates. We also found other approaches that do not reduce recidivism. Thus, the first basic lesson from our evidence-based review is that some adult corrections programs work and some do not. A direct implication from these mixed findings is that a corrections policy that reduces recidivism will be one that focuses resources on effective evidence-based programming and avoids ineffective approaches.

As an example of the information on Exhibit 1, we analyzed the findings from 25 well-researched cognitive-

⁵ We follow the meta-analytic methods described in: M. W. Lipsey and D. Wilson (2001). *Practical meta-analysis*. Thousand Oaks: Sage Publications.
⁶ Technical meta-analytical results are presented in Exhibit 2.

Exhibit 1		
Adult Corrections: What Works?		
Estimated Percentage Change in Recidivism Rates (and the number of studies on which the estimate is based)		
Example of how to read the table: an analysis of 56 adult drug court evaluations indicates that drug courts achieve, on average, a statistically significant 10.7 percent reduction in the recidivism rates of program participants compared with a treatment-as-usual group.		
Programs for Drug-Involved Offenders		
Adult drug courts	-10.7%	(56)
In-prison “therapeutic communities” with community aftercare	-6.9%	(6)
In-prison “therapeutic communities” without community aftercare	-5.3%	(7)
Cognitive-behavioral drug treatment in prison	-6.8%	(8)
Drug treatment in the community	-12.4%	(5)
Drug treatment in jail	-6.0%	(9)
Programs for Offenders With Co-Occurring Disorders		
Jail diversion (pre- and post-booking programs)	0.0%	(11)
Programs for the General Offender Population		
General and specific cognitive-behavioral treatment programs	-8.2%	(25)
Programs for Domestic Violence Offenders		
Education/cognitive-behavioral treatment	0.0%	(9)
Programs for Sex Offenders		
Psychotherapy for sex offenders	0.0%	(3)
Cognitive-behavioral treatment in prison	-14.9%	(5)
Cognitive-behavioral treatment for low-risk offenders on probation	-31.2%	(6)
Behavioral therapy for sex offenders	0.0%	(2)
Intermediate Sanctions		
Intensive supervision: surveillance-oriented programs	0.0%	(24)
Intensive supervision: treatment-oriented programs	-21.9%	(10)
Adult boot camps	0.0%	(22)
Electronic monitoring	0.0%	(12)
Restorative justice programs for lower-risk adult offenders	0.0%	(6)
Work and Education Programs for the General Offender Population		
Correctional industries programs in prison	-7.8%	(4)
Basic adult education programs in prison	-5.1%	(7)
Employment training and job assistance in the community	-4.8%	(16)
Vocational education in prison	-12.6%	(3)
Program Areas in Need of Additional Research & Development		
<i>(The following types of programs require additional research before it can be concluded that they do or do not reduce adult recidivism rates)</i>		
Case management in the community for drug offenders	0.0%	(12)
“Therapeutic community” programs for mentally ill offenders	-27.4%	(2)
Faith-based programs	0.0%	(5)
Domestic violence courts	0.0%	(2)
Intensive supervision of sex offenders in the community	0.0%	(4)
Mixed treatment of sex offenders in the community	0.0%	(2)
Medical treatment of sex offenders	0.0%	(1)
COSA (Faith-based supervision of sex offenders)	-31.6%	(1)
Regular parole supervision vs. no parole supervision	0.0%	(1)
Day fines (compared to standard probation)	0.0%	(1)
Work release programs	-5.6%	(4)

behavioral treatment programs for general adult offenders. We found that, on average, these programs can be expected to reduce recidivism rates by 8.2 percent. That is, without a cognitive-behavioral program we expect that about 49 percent of these offenders will recidivate with a new felony conviction after an eight-year follow-up. With a cognitive-behavioral treatment program, we expect the recidivism probability to drop four points to 45 percent—an 8.2 percent reduction in recidivism rates.

It is important to note that even relatively small reductions in recidivism rates can be quite cost-beneficial. For example, a 5 percent reduction in the reconviction rates of high risk offenders can generate significant benefits for taxpayers and crime victims. Moreover, a program that has no statistically significant effect on recidivism rates can be cost-beneficial if the cost of the program is less than the cost of the alternative. Jail diversion programs are examples of this; even if research demonstrates that diversion programs have no effect on recidivism, the programs may still be economically attractive if they cost less than avoided jail costs. In the final version of this report, to be delivered to the Legislature in October 2006, we will present full benefit-cost estimates for each of the programs shown in Exhibit 1.⁷

Findings by Type of Program

We organized our review of the adult corrections evidence base into eight categories of correctional programming (as shown in Exhibit 1). A brief discussion of our findings for each of these categories follows.

Programs for Drug-Involved Offenders. We analyzed 92 rigorous evaluations of drug treatment programs. These programs are for drug-involved adult offenders in a variety of prison and community settings. We found that, on average, drug treatment leads to a statistically significant reduction in criminal recidivism rates. We examined adult drug courts, in-prison therapeutic communities, and other types of drug

treatment including cognitive-behavioral approaches.

Adult Drug Courts. Specialized courts for drug-involved offenders have proliferated throughout the United States, and there are several adult drug courts in Washington. We found 56 evaluations with sufficient rigor to be included in our statistical review. We conclude that drug courts achieve, on average, a statistically significant 10.7 percent reduction in the recidivism rates of program participants relative to treatment-as-usual comparison groups.

In-Prison Therapeutic Communities. Programs for drug offenders in a prison or jail setting are typically called “therapeutic communities” when they contain separate residential units for the offenders and when they follow group-run principles of organizing and operating the drug-free unit. Some evaluations of the effectiveness of in-prison therapeutic community programs have also included community-based aftercare for offenders once they leave incarceration. Based on our review of the evaluation literature, we found that the average therapeutic community reduces recidivism by 5.3 percent. The community aftercare component, however, produces only a modest additional boost to program effectiveness—to a 6.9 percent reduction. Thus, most of the recidivism reduction effect appears to stem from the prison-based therapeutic community experience for these offenders.

Other Types of Drug Treatment. As shown in Exhibit 1, we also studied the effects of three other types of drug treatment modalities: prison-based drug treatment that employs a cognitive-behavioral approach, general drug treatment approaches in the community, and general drug treatment programs in local jails. We found that each of these approaches achieve, on average, a statistically significant reduction in recidivism.

Jail Diversion Programs for Offenders With Mental Illness and Co-Occurring Disorders. There is young but growing research literature testing the effectiveness of jail diversion programs for mentally ill adults and for offenders with co-occurring mental health and substance abuse disorders. Some of these are pre-booking programs implemented by the police, and some are post-booking programs implemented by court personnel, such as mental health courts. We found 11 evaluations with sufficient research rigor

⁷ An overview of what will be included in the October 2006 report can be found at www.wsipp.wa.gov/ Steve Aos (2006). *Options to Stabilize Prison Populations in Washington State, Interim Report*, Olympia: Washington State Institute for Public Policy.

to be included in our review. Eight of these programs were part of a recent federally-funded effort (Broner et al., 2004). On average, these approaches have not demonstrated a statistically significant reduction in the recidivism rates of program participants. This null finding does not mean the programs are not valuable; since they are typically designed to divert offenders from costly sentences in local jails, they may save more money than the programs cost. As mentioned earlier, we will review the economics of all programs in the present study in our October 2006 final report.

Treatment Programs for the General Offender Population.

Cognitive-Behavioral Treatment. We found 25 rigorous evaluations of programs for the general offender population that employ cognitive-behavioral treatment. This type of group therapy addresses the irrational thoughts and beliefs that lead to anti-social behavior. The programs are designed to help offenders correct their thinking and provide opportunities to model and practice problem-solving and pro-social skills. On average, we found these programs significantly reduce recidivism by 8.2 percent. We identified three well-defined programs that provide manuals and staff training regimens: *Reasoning and Rehabilitation (R&R)*, *Moral Reconation Therapy (MRT)*, and *Thinking for a Change (T4C)*. Effects of R&R and MRT are significant and similar to each other and to the other cognitive-behavioral treatment programs in our review. Only a single evaluation of T4C is currently available. Since, on average, all of these programs produce similar results, we recommend the state choose any of the three well-defined programs for implementation in Washington.

Programs for Domestic-Violence Offenders

Education/Cognitive-Behavioral Treatment. Treatment programs for domestic violence offenders most frequently involve an educational component focusing on the historical oppression of women and cognitive-behavioral treatment emphasizing alternatives to violence. Treatment is commonly mandated by the court. Based on our review of nine rigorous evaluations, domestic violence treatment programs have yet, on average, to demonstrate reductions in recidivism.

Programs for Sex Offenders.⁸ We found 18 well-designed evaluations of treatment programs for sex offenders. Some of these programs are located in a prison setting and some are in the community. Sex offenders sentenced to prison are typically convicted of more serious crimes than those sentenced to probation. We found that cognitive-behavioral treatments are, on average, effective at reducing recidivism, but other types of sex offender treatment fail to demonstrate significant effects on further criminal behavior.

*Psychotherapy/Counseling for Sex Offenders.*⁹ These programs involve insight-oriented individual or group therapy or counseling. We found only three rigorous studies of this approach to treatment. The results indicate that this approach does not reduce recidivism in sex offenders.

Cognitive-Behavioral Treatment of Sex Offenders in Prison. Sex offenders sentenced to prison are typically convicted of more serious crimes than those sentenced to probation. We examined five rigorous studies of these specialized cognitive-behavioral programs that may also include behavioral reconditioning to discourage deviant arousal, and modules addressing relapse prevention. Among the five programs in this category was a randomized trial¹⁰ with an eight-year follow-up showing small but non-significant effects on recidivism. On average across all five studies, however, we found that cognitive-behavioral therapy for sex offenders in prison significantly reduces recidivism by 14.9 percent.

Cognitive-Behavioral Treatment of Low-Risk Sex Offenders on Probation. Offenders sentenced to probation have usually been convicted of less serious crimes than sex offenders sentenced to prison. Cognitive-behavioral programs for sex offenders on probation are similar to the programs in prisons, and may also incorporate behavioral reconditioning and relapse prevention. We found six rigorous studies and conclude that cognitive-

⁸ The categories of sex offender treatment listed here are based on those outlined in two recent reviews of sex offender treatment literature: R. K. Hanson, A. Gordon, A. J. Harris, J. K. Marques, W. Murphy, V. L. Quinsey, and M. C. Seto (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders, *Sexual Abuse: A Journal of Research and Treatment*, 14(2): 169-194; F. Losel, and M. Schmucker (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis, *Journal of Experimental Criminology*, 1: 117-146

⁹ Psychotherapy and counseling are not currently used as stand-alone treatment for sex offenders (Hanson, et al., 2002).

¹⁰ J. K. Marques, M. Wiederanders, D. M. Day, C. Nelson, and A. van Ommeren (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP), *Sexual Abuse: A Journal of Research and Treatment*, 17(1): 79-107.

behavioral therapy for sex offenders on probation significantly reduces recidivism. As a group, these programs demonstrated the largest effects observed in our analysis.

Behavioral Treatment of Sex Offenders. Behavioral treatments focus on reducing deviant arousal (using biofeedback or other conditioning) and increasing skills necessary for social interaction with age appropriate individuals. The two rigorous studies of programs using only behavioral treatment failed to show reductions in recidivism.

Intermediate Sanctions. In the 1980s and 1990s a number of sanctioning and sentencing alternatives were proposed and evaluated. Interest in developing additional alternatives continues. We found studies that center on five types of these “intermediate” sanctions.

Intensive Supervision With and Without a Focus on Treatment. We found 24 evaluations of intensive community supervision programs where the focus was on offender monitoring and surveillance. These programs are usually implemented by lowering the caseload size of the community supervision officer. This approach to offender management has not, on average, produced statistically significant reductions in recidivism rates. On the other hand, intensive supervision programs where the focus is on providing treatment services for the offenders have produced significant reductions; we found 10 well-researched evaluations of treatment-oriented intensive supervision programs that on average produced considerable recidivism reductions. The lesson from this research is that it is the treatment—not the intensive monitoring—that results in recidivism reduction.

Adult Boot Camps. Boot camps are intensive regimens of training, drilling, and some treatment. We found 24 rigorous evaluations of adult boot camps and, on average, they do not produce a statistically significant reduction in re-offense rates. As with our comment on jail diversion programs, however, it is possible that boot camps are economically attractive if they cost less to run than the alternative. Our October 2006 report will analyze the economics of adult boot camps.

Electronic Monitoring. Supervision of offenders in the community that is aided with electronic monitoring devices has been the focus of some rigorous evaluation efforts. We found 12 control-group studies; on average they indicate that electronic monitoring does not reduce recidivism.

Restorative Justice for Lower-Risk Adult Offenders. Restorative justice approaches have been tried for both juvenile and adult offenders. Offenders placed in restorative justice programs are often, but not always, lower risk compared with offenders processed through the usual court procedures. Restorative justice typically involves a form of victim-offender mediation, family group conferences, or restitution. We found six rigorous evaluations of these programs for adult offenders. On average, they did not result in lower recidivism rates. Our October 2006 report will also report on restorative justice programs for juvenile offenders. Unlike our findings for the restorative justice programs for adult offenders, our preliminary findings indicate that restorative justice programs do achieve significant reductions in recidivism rates of lower-risk juvenile offenders.

Work and Education Programs for General Offenders. We found 30 rigorous evaluations of programs that attempt to augment the educational, vocational, and job skills of adult offenders. Some of these programs are for offenders in prison and some are in community settings. On average, we found that employment and education-related programs lead to modest but statistically significant reductions in criminal recidivism rates. We examined the following five categories of these programs.

In-prison Correctional Industries Program. Most states run in-prison correctional industries programs, yet only a few have been evaluated rigorously. We located only four outcome evaluations of correctional industries programs. On average, these programs produce a statistically significant reduction in recidivism rates. Our updated economic analysis of this finding will be presented in October 2006.

Basic Adult Education Programs in Prison. We found seven rigorous evaluations of programs that teach remedial educational skills to adult offenders when they are in prison. On average, these programs reduce the recidivism rates of program participants.

Employment Training and Job Assistance Programs in the Community. We analyzed the results of 16 rigorous evaluations of community-based employment training, job search, and job assistance programs for adult offenders. These programs produce a modest but statistically significant reduction in recidivism.

Vocational Education Programs in Prison. We found only three quality studies of vocational training programs for offenders while they are in prison. On average, the programs appear to reduce recidivism, but additional tests of this tentative finding is necessary.

Programs Requiring Further Study. In our review of the adult corrections literature, we were unable to draw conclusions about recidivism reduction for a number of programs. In Exhibit 1, we list these inconclusive findings at the bottom of the table. For each of these approaches, further research is required before even tentative conclusions can be drawn.¹¹

Case Management in the Community for Drug Offenders. These types of programs typically involve an outside third-party agency that provides case coordination services and drug testing. The goal is to provide the coordination of other existing monitoring and treatment services for offenders in the community. We found 12 rigorous tests of this approach. Our statistical tests reveal that while, on average, these programs have no significant effect on recidivism, some case management programs do have an effect and some do not. This inconclusive result means that additional research is required on this class of programming in order to identify the aspects of case management that are effective or ineffective. In other words, additional research may indicate that some forms of case management reduce recidivism.¹²

“Therapeutic Community” Programs for Mentally Ill Offenders. A relatively new approach to providing treatment to mentally-ill offenders follows a modified version of the therapeutic community approach to drug offenders described earlier. This approach appears to show promise in reducing recidivism rates.

¹¹ Technical Note. As we explain in the technical appendix, we employ “fixed effects” and “random effects” modeling to derive meta-analytic estimates of program effectiveness. Sometimes, a collection of evaluations of similar programs has significant recidivism when judged with fixed effects modeling, but the same set of programs has insignificant findings when a random effects model is used. This situation provides an indication that additional meta-analytic research is needed to identify the factors that produced the heterogeneity in the outcomes. Several of the programs listed here fall into this category. For more information, see the technical appendices.

¹² As a technical note, Exhibit 2 shows that case management services produce a marginally significant ($p=.114$) effect on recidivism in a fixed effects model but the model indicates significant ($p=.000$) heterogeneity. The random effects model indicates non significance ($p=.48$). Thus, a multivariate meta-analysis of this literature may isolate the factors that were associated with successful approaches among the 12 studies.

However, this is based on only two rigorous studies, and they involved small samples of offenders. Thus, this is an approach that requires additional research.

Faith-Based Programs. These Christian-based programs provide religious ministry, including bible study, to offenders in prison and/or when offenders re-enter the community. The faith-based offender programs that have been evaluated to date do not significantly reduce recidivism.¹³ Rigorous evaluations of faith-based programs are still relatively rare—we found only five thorough evaluations—and future studies may provide evidence of better outcomes.

Domestic Violence Courts. These specialized courts are designed to provide effective coordinated response to domestic violence. Domestic violence courts commonly bring together criminal justice and social service agencies and may mandate treatment for offenders. The two courts included here differed—one was exclusively for felony cases and the other for misdemeanors. In the misdemeanor court, recidivism was lowered, while the felony court observed increased recidivism. Thus, this is an area that requires additional research.

Intensive Supervision of Sex Offenders in the Community. The programs included in the analysis were all developed in Illinois and varied by county. All involve a specialized probation caseload, frequent face-to-face meetings with offenders, and home visits and inspections. Supervision programs may also include treatment. The recidivism results in the four counties vary widely, suggesting that some of the programs may be effective while others are not. Additional research is needed to identify these characteristics.

Mixed Treatment of Sex Offenders. Two rigorous studies evaluated community sex offender treatments employed across geographic areas (Washington State and British Columbia). In each case, the individual treatment programs varied widely. On average, these mixtures of treatments significantly reduced recidivism; however, while the treatments in Washington were significant and large, those in British Columbia were very small and non-significant. Controlling for the variation, the overall effect was zero.

¹³ Similar findings were recently published in a review of faith-based prison programs: J. Burnside, N. Loucks, J. R. Addler, and G. Rose (2005). *My brother's keeper: Faith-based units in prison*, Cullompton, Devon, U.K.: Willan Publishing, p. 314.

Medical Treatment of Sex Offenders. Several medical approaches to treating sex offenders have been tried. These include castration and two types of hormonal therapy. Ethical considerations have made it difficult to conduct rigorous evaluations of these types of treatment. The single study we used in our analysis compared men who volunteered for castration to another group who volunteered but did not receive the surgery. Recidivism was significantly less among castrated offenders.

Circles of Support and Accountability (COSA/ Faith-Based Supervision of Sex Offenders). This program originated among members of the Mennonite church in Canada. Volunteers provide support to sex offenders being released from prison. Five lay volunteers visit or contact the offender every week. The volunteers are supported by community-based professionals, typically psychologists, law enforcement, correctional officers, or social service workers; the full circle meets weekly. The single evaluation of this program showed a significant reduction in recidivism of 31.6 percent.

Regular Parole Supervision vs. No Parole Supervision. The Urban Institute recently reported the results of a study that compared the recidivism rates of adult prisoners released from prison with parole to those released from prison without parole. The study used a large national database covering 15 states. It found no statistically significant effect of parole on recidivism. This null result is consistent with our results for surveillance-oriented intensive supervision programs versus regular levels of supervision (reported above). We would like to see additional treatment and comparison group tests of the parole vs. no-parole question before drawing firm conclusions.

Day Fines (compared with standard probation). We found one rigorous study of “day fines.” These fines, which are more common in Europe than the United States, allow judges to impose fines that are commensurate with an offender’s ability to pay and the seriousness of the offence. This approach has been evaluated for low-risk felony offenders and was used to divert these offenders from regular parole supervision. The approach had no effect on recidivism rates but additional research is needed to estimate whether this sentencing alternative is cost-beneficial.

Work Release Programs. We found only four quality studies of work release programs. While, on average, these programs appear to reduce recidivism, more rigorous outcome research is needed on this type of adult corrections program.

Technical Appendices

Appendix 1: Meta-Analysis Coding Criteria

Appendix 2: Procedures for Calculating Effect Sizes

Appendix 3: Institute Adjustments to Effect Sizes for Methodological Quality, Outcome Measure Relevance, and Researcher Involvement

Appendix 4: Meta-Analytic Results—Estimated Effect Sizes and Citations to Studies Used in the Analyses

Appendix 1: Meta-Analysis Coding Criteria

A meta-analysis is only as good as the selection and coding criteria used to conduct the study. The following are the key choices we made and implemented for this meta-analysis of adult corrections programs.

- 1. Study Search and Identification Procedures.** We searched for all adult corrections evaluation studies conducted since 1970. The studies had to be written in English. We used three primary means to identify and locate these studies: a) we consulted the study lists of other systematic and narrative reviews of the adult corrections research literature—there have been a number of recent reviews on particular topics; b) we examined the citations in the individual studies; and c) we conducted independent literature searches of research databases using search engines such as Google, Proquest, Ebsco, ERIC, and SAGE. As we describe, the most important inclusion criteria in our study was that an evaluation have a control or comparison group. Therefore, after first identifying all possible studies using these search methods, we attempted to determine whether the study was an outcome evaluation that had a comparison group. If a study met these criteria, we then secured a paper copy of the study for our review.
- 2. Peer-Reviewed and Other Studies.** We examined all program evaluation studies we could locate with these search procedures. Many of these studies were published in peer-reviewed academic journals, while many others were from government reports obtained from the agencies themselves. It is important to include non-peer reviewed studies, because it has been suggested that peer-reviewed publications may be biased to show positive program effects. Therefore, our meta-analysis included all available studies regardless of published source.
- 3. Control and Comparison Group Studies.** We only included studies in our analysis if they had a control or comparison group. That is, we did not include studies with a single-group, pre-post research design. This choice was made because we believe that it is only through rigorous comparison group studies that average treatment effects can be reliably estimated.
- 4. Exclusion of Studies of Program Completers Only.** We did not include a comparison study in our meta-analytic review if the treatment group was made up solely of program completers. We adopted this rule, because we believe there are too many significant unobserved self-selection factors that distinguish a program completer from a program dropout, and that these unobserved factors are likely to significantly bias estimated treatment effects. Some comparison group studies of program completers, however, contain information on program dropouts in addition to a comparison group. In these situations, we included the study if sufficient information was provided to allow us to reconstruct an intent-to-treat group that included both completers and non-completers, or if the demonstrated rate of program non-completion was very small (e.g. under 10 percent). In these cases, the study still needed to meet the other inclusion requirements listed here.
- 5. Random Assignment and Quasi- Experiments.** Random assignment studies were preferred for inclusion in our review, but we also included non-randomly assigned control groups. We only included quasi-experimental studies if, and only if, sufficient information was provided to demonstrate comparability between the treatment and comparison groups on important pre-existing conditions such as age, gender, and prior criminal history. Of the 291 individual studies in our review, about 20 percent were effects estimated from well implemented random assignment studies.
- 6. Enough information to Calculate an Effect Size.** Following the statistical procedures in Lipsey and Wilson (2001), a study had to provide the necessary information to calculate an effect size. If the necessary information was not provided, the study was not included in our review.
- 7. Mean-Difference Effect Sizes.** For this study we coded mean-difference effect sizes following the procedures in Lipsey and Wilson (2001). For dichotomous crime measures, we used the arcsine transformation to approximate the mean difference effect size, again following Lipsey and Wilson. We chose to use the mean-difference effect size rather than the odds ratio effect size because we frequently coded both dichotomous and continuous outcomes (odds ratio effect sizes could also have been used with appropriate transformations).
- 8. Unit of Analysis.** Our unit of analysis for this study was an independent test of a treatment in a particular site. Some studies reported outcome evaluation information for multiple sites; we included each site as an independent observation if a unique and independent comparison group was also used at each site.

9. **Multivariate Results Preferred.** Some studies presented two types of analyses: raw outcomes that were not adjusted for covariates such as age, gender, criminal history; and those that had been adjusted with multivariate statistical methods. In these situations, we coded the multivariate outcomes.
10. **Broadest Measure of Criminal Activity.** Some studies presented several types of crime-related outcomes. For example, studies frequently measured one or more of the following outcomes: total arrests, total convictions, felony arrests, misdemeanor arrests, violent arrests, and so on. In these situations, we coded the broadest crime outcome measure. Thus, most of the crime outcome measures that we coded in this analysis were total arrests and total convictions.
11. **Averaging Effect Sizes for Arrests and Convictions.** When a study reported both total arrests and total convictions, we calculated an effect size for each measure then took a simple average of the two effect sizes.
12. **Dichotomous Measures Preferred Over Continuous Measures.** Some studies included two types of measures for the same outcome: a dichotomous (yes/no) outcome and a continuous (mean number) measure. In these situations, we coded an effect size for the dichotomous measure. Our rationale for this choice is that in small or relatively small sample studies, continuous measures of crime outcomes can be unduly influenced by a small number of outliers, while dichotomous measures can avoid this problem. Of course, if a study only presented a continuous measure, then we coded the continuous measure.
13. **Longest Follow-Up Times.** When a study presented outcomes with varying follow-up periods, we generally coded the effect size for the longest follow-up period. The reason for this is that our intention for this analysis is to compute the long-run benefits and costs of different programs. The longest follow-up period allows us to gain the most insight into the long-run effect of these programs on criminality. Occasionally, we did not use the longest follow-up period if it was clear that a longer reported follow-up period adversely affected the attrition rate of the treatment and comparison group samples.
14. **Measures of New Criminal Activity.** Whenever possible, we excluded outcome measures that did not report on new criminal activity. For example, we avoided coding measure of technical violations of probation or parole. We do not think that technical violations are unimportant, but our purpose in this meta-analysis is to ascertain whether these programs affect new criminal activity.
15. **Some Special Coding Rules for Effect Sizes.** Most studies in our review had sufficient information to code exact mean-difference effect sizes. Some studies, however, reported some, but not all of the information required. The rules we followed for these situations are these:
 - a. **Two-Tail P-Values.** Some studies only reported p-values for significance testing of program outcomes. When we had to rely on these results, if the study reported a one-tail p-value, we converted it to a two-tail test.
 - b. **Declaration of Significance by Category.** Some studies reported results of statistical significance tests in terms of categories of p-values. Examples include: $p \leq .01$, $p \leq .05$, or “non-significant at the $p = .05$ level.” We calculated effect sizes for these categories by using the highest p-value in the category. Thus if a study reported significance at “ $p \leq .05$,” we calculated the effect size at $p = .05$. This is the most conservative strategy. If the study simply stated a result was “non-significant,” we computed the effect size assuming a p-value of .50 (i.e. $p = .50$).

Appendix 2: Procedures for Calculating Effect Sizes

Effect sizes measure the degree to which a program has been shown to change an outcome for program participants relative to a comparison group. There are several methods used by meta-analysts to calculate effect sizes, as described in Lipsey and Wilson (2001). In this, we use statistical procedures to calculate the *mean difference effect sizes* of programs. We did not use the odds-ratio effect size because many of the outcomes measured in this study are continuously measured. Thus, the mean difference effect size was a natural choice.

Many of the outcomes we record, however, are measured as dichotomies. For these yes/no outcomes, Lipsey and Wilson (2001) show that the mean difference effect size calculation can be approximated using the arcsine transformation of the difference between proportions.¹⁴

$$(A1) \quad ES_{m(p)} = 2 \times \arcsin \sqrt{P_e} - 2 \times \arcsin \sqrt{P_c}$$

In this formula, $ES_{m(p)}$ is the estimated effect size for the difference between proportions from the research information; P_e is the percentage of the population that had an outcome such as re-arrest rates for the experimental or treatment group; and P_c is the percentage of the population that was re-arrested for the control or comparison group.

A second effect size calculation involves continuous data where the differences are in the means of an outcome. When an evaluation reports this type of information, we use the standard mean difference effect size statistic.¹⁵

¹⁴ Lipsey and Wilson, *Practical meta-analysis*, Table B10, formula (22).

¹⁵ Ibid., Table B10, formula (1).

$$(A2) \quad ES_m = \frac{M_e - M_c}{\sqrt{\frac{SD_e^2 + SD_c^2}{2}}}$$

In this formula, ES_m is the estimated effect size for the difference between means from the research information; M_e is the mean number of an outcome for the experimental group; M_c is the mean number of an outcome for the control group; SD_e is the standard deviation of the mean number for the experimental group; and SD_c is the standard deviation of the mean number for the control group.

Often, research studies report the mean values needed to compute ES_m in (A2), but they fail to report the standard deviations. Sometimes, however, the research will report information about statistical tests or confidence intervals that can then allow the pooled standard deviation to be estimated. These procedures are also described in Lipsey and Wilson (2001).

Adjusting Effect Sizes for Small Sample Sizes

Since some studies have very small sample sizes, we follow the recommendation of many meta-analysts and adjust for this. Small sample sizes have been shown to upwardly bias effect sizes, especially when samples are less than 20. Following Hedges (1981),¹⁶ Lipsey and Wilson (2001)¹⁷ report the “Hedges correction factor,” which we use to adjust all mean difference effect sizes (N is the total sample size of the combined treatment and comparison groups):

$$(A3) \quad ES'_m = \left[1 - \frac{3}{4N - 9}\right] \times [ES_{m, or}, ES_{m(p)}]$$

Computing Weighted Average Effect Sizes, Confidence Intervals, and Homogeneity Tests

Once effect sizes are calculated for each program effect, the individual measures are summed to produce a weighted average effect size for a program area. We calculate the inverse variance weight for each program effect, and these weights are used to compute the average. These calculations involve three steps. First, the standard error, SE_m of each mean effect size is computed with:¹⁸

$$(A4) \quad SE_m = \sqrt{\frac{n_e + n_c}{n_e n_c} + \frac{(ES'_m)^2}{2(n_e + n_c)}}$$

In equation (A4), n_e and n_c are the number of participants in the experimental and control groups and ES'_m is from equation (A3).

Next, the inverse variance weight w_m is computed for each mean effect size with:¹⁹

$$(A5) \quad w_m = \frac{1}{SE_m^2}$$

The weighted mean effect size for a group of studies in program area i is then computed with:²⁰

$$(A6) \quad \overline{ES} = \frac{\sum (w_{m_i} ES'_{m_i})}{\sum w_{m_i}}$$

Confidence intervals around this mean are then computed by first calculating the standard error of the mean with:²¹

$$(A7) \quad SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_{m_i}}}$$

Next, the lower, ES_L , and upper limits, ES_U , of the confidence interval are computed with:²²

$$(A8) \quad \overline{ES}_L = \overline{ES} - z_{(1-\alpha)}(SE_{\overline{ES}})$$

$$(A9) \quad \overline{ES}_U = \overline{ES} + z_{(1-\alpha)}(SE_{\overline{ES}})$$

In equations (A8) and (A9), $z_{(1-\alpha)}$ is the critical value for the z -distribution (1.96 for $\alpha = .05$).

The test for homogeneity, which provides a measure of the dispersion of the effect sizes around their mean, is given by:²³

$$(A10) \quad Q_i = \left(\sum w_i ES_i^2\right) - \frac{\left(\sum w_i ES_i\right)^2}{\sum w_i}$$

The Q -test is distributed as a chi-square with $k-1$ degrees of freedom (where k is the number of effect sizes).

Computing Random Effects Weighted Average Effect Sizes and Confidence Intervals

When the p -value on the Q -test indicates significance at values of p less than or equal to .05, a random effects model is performed to calculate the weighted average effect size. This is accomplished by first calculating the random effects variance component, v .²⁴

$$(A11) \quad v = \frac{Q_i - (k - 1)}{\sum w_i - \left(\sum w_i q_i / \sum w_i\right)}$$

This random variance factor is then added to the variance of each effect size and then all inverse variance weights are recomputed, as are the other meta-analytic test statistics.

¹⁹ Ibid., 49, equation 3.24.

²⁰ Ibid., 114.

²¹ Ibid., 114.

²² Ibid., 114.

²³ Ibid., 116.

²⁴ Ibid., 134.

¹⁶ L. V. Hedges (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6: 107-128.

¹⁷ Lipsey and Wilson, *Practical meta-analysis*, 49, formula 3.22.

¹⁸ Ibid., 49, equation 3.23.

Appendix 3: Institute Adjustments to Effect Sizes for Methodological Quality, Outcome Measure Relevance, and Researcher Involvement

In Exhibit 2 we show the results of our meta-analyses calculated with the standard meta-analytic formulas described in Appendix 2. In the last column in Exhibit 2, however, we list “Adjusted Effect Sizes” that we actually use in our benefit-cost analysis of each of the programs we review. These adjusted effect sizes, which are derived from the unadjusted results, are always smaller than or equal to the unadjusted effect sizes we report in the other columns in Exhibit 2.

In Appendix 3, we describe our rationale for making these downward adjustments. In particular, we make three types of adjustments that we believe are necessary to better estimate the results that we think each program is likely to actually achieve in real-world settings. We make adjustments for: a) the methodological quality of each of the studies we include in the meta-analyses; b) the relevance or quality of the outcome measure that individual studies use; and c) the degree to which the researcher(s) who conducted a study were invested in the program’s design and implementation.

3a. Methodological Quality. Not all research is of equal quality, and this, we believe, greatly influences the confidence that can be placed in the results from a study. Some studies are well designed and implemented, and the results can be viewed as accurate representations of whether the program itself worked. Other studies are not designed as well and less confidence can be placed in any reported differences. In particular, studies of inferior research design cannot completely control for sample selection bias or other unobserved threats to the validity of reported research results. This does not mean that results from these studies are of no value, but it does mean that less confidence can be placed in any cause-and-effect conclusions drawn from the results.

To account for the differences in the quality of research designs, we use a 5-point scale as a way to adjust the reported results. The scale is based closely on the 5-point scale developed by researchers at the University of Maryland.²⁵ On this 5-point scale, a rating of “5” reflects an evaluation in which the most confidence can be placed. As the evaluation ranking gets lower, less confidence can be placed in any reported differences (or lack of differences) between the program and comparison or control groups.

On the 5-point scale, as interpreted by the Institute, each study is rated with the following numerical ratings.

- A “5” is assigned to an evaluation with well-implemented random assignment of subjects to a treatment group and a control group that does not receive the treatment/program. A good random assignment study should also indicate how well the random assignment actually occurred by reporting

values for pre-existing characteristics for the program and control groups.

- A “4” is assigned to a study that employs a rigorous quasi-experimental research design with a program and matched comparison group, controlling with statistical methods for self-selection bias that might otherwise influence outcomes. These quasi-experimental methods may include estimates made with a convincing instrumental variables modeling approach, or a Heckman approach to modeling self-selection.²⁶ A level 4 study may also be used to “downgrade” an experimental random assignment design that had problems in implementation, perhaps with significant attrition rates.
- A “3” indicates a non-experimental evaluation where the program and comparison groups were reasonably well matched on pre-existing differences in key variables. There must be evidence presented in the evaluation that indicates few, if any, significant differences were observed in these salient pre-existing variables. Alternatively, if an evaluation employs sound multivariate statistical techniques (e.g. logistic regression) to control for pre-existing differences, and if the analysis is successfully completed, then a study with some differences in pre-existing variables can qualify as a level 3.
- A “2” involves a study with a program and matched comparison group where the two groups lack comparability on pre-existing variables and no attempt was made to control for these differences in the study.
- A “1” involves a study where no comparison group is utilized. Instead, the relationship between a program and an outcome, i.e., recidivism, is analyzed before and after the program.

We do not use the results from program evaluations rated as a “1” on this scale, because they do not include a comparison group and we believe that there is no context to judge program effectiveness. We also regard evaluations with a rating of “2” as highly problematic and, as a result, we do not consider their findings in the calculations of effect. In this study, we only consider evaluations that rate at least a 3 on this 5-point scale.

An explicit adjustment factor is assigned to the results of individual effect sizes based on the Institute’s judgment concerning research design quality. We believe this adjustment is critical and is the only practical way to combine the results of a high quality study (i.e., a level 5 study) with those of lesser design quality. The specific adjustments made for these studies depend on the topic area being considered. In some areas, such as criminal justice program evaluations, there is strong evidence that less-than-random assignment studies (i.e., less than level 5 studies) have, on average, smaller effect

²⁵ L. W. Sherman, D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, and S. Bushway (1998). *Preventing crime: What works, what doesn't, what's promising*. Prepared for the National Institute of Justice. Department of Criminology and Criminal Justice, University of Maryland. Chapter 2.

²⁶ For a discussion of these methods, see W. Rhodes, B. Pelissier, G. Gaes, W. Saylor, S. Camp, and S. Wallace (2001). *Alternative solutions to the problem of selection bias in an analysis of federal residential drug treatment programs*. *Evaluation Review*, 25(3): 331-369.

sizes than weaker-designed studies.²⁷ Thus, for the typical criminal justice evaluation, we use the following “default” adjustments to account for studies of different research design quality:

- A level 5 study carries a factor of 1.0 (that is, there is no discounting of the study’s evaluation outcomes).
- A level 4 study carries a factor of .75 (effect sizes discounted by 25 percent).
- A level 3 study carries a factor of .50 (effect sizes discounted by 50 percent).
- We do not include level 2 and level 1 studies in our analyses.

These factors are subjective to a degree; they are based on the Institute’s general impressions of the confidence that can be placed in the predictive power of criminal justice studies of different quality.

The effect of the adjustment is to multiply the effect size for any study, ES'_m , in equation (A3) by the appropriate research design factor. For example, if a study has an effect size of -.20 and it is deemed a level 4 study, then the -.20 effect size would be multiplied by .75 to produce a -.15 adjusted effect size for use in the benefit-cost analysis.

3b. Adjusting Effect Sizes for Relevance or Quality of the Outcome Measure. As noted in Appendix 1, our focus in this analysis is whether adult corrections programs reduce new criminal activity. We prefer measures such as arrests or convictions and avoid measures such as technical violations of parole or probation, since these may or may not be related to the commission of new crimes. In addition, we require that all studies have at least a six-month follow up period. For those studies that had a follow-up period of under 12 months, but greater than six months, and for those studies that only reported weak measures of new criminal activity, we reduced effects sizes by 25 percent. This adjustment multiplies the effect size for any study with a short follow-up or weak measure by .75.

²⁷ M. W. Lipsey (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science*, 587(1): 69-81. Lipsey found that, for juvenile delinquency evaluations, random assignment studies produced effect sizes only 56 percent as large as nonrandom assignment studies.

3c. Adjusting Effect Sizes for Research Involvement in the Program’s Design and Implementation. The purpose of the Institute’s work is to identify and evaluate programs that can make cost-beneficial improvements to Washington’s actual service delivery system. There is some evidence that programs that are closely controlled by researchers or program developers have better results than those that operate in “real world” administrative structures.²⁸ In our own evaluation of a real-world implementation of a research-based juvenile justice program in Washington, we found that the actual results were considerably lower than the results obtained when the intervention was conducted by the originators of the program.²⁹ Therefore, we make an adjustment to effect sizes ES_m to reflect this distinction. As a parameter for all studies deemed not to be “real world” trials, the Institute discounts ES'_m by .5, although this can be modified on a study-by-study basis.

Appendix 4: Meta-Analytic Results—Estimated Effect Sizes and Citations to Studies Used in the Analyses

Exhibit 2 provides technical meta-analytic results for the effect sizes computed for these groupings of programs, including the results of the adjustments described above. Exhibit 3 lists the citations for all the studies used in the meta-analyses, arranged by program area.

²⁸ Ibid. Lipsey found that, for juvenile delinquency evaluations, programs in routine practice (i.e., “real world” programs) produced effect sizes only 61 percent as large as research/demonstration projects. See also: A. Petrosino, & H. Soydan (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1(4): 435-450.

²⁹ R. Barnoski (2004). *Outcome evaluation of Washington State’s research-based programs for juvenile offenders*. Olympia: Washington State Institute for Public Policy, available at <<http://www.wsipp.wa.gov/rptfiles/04-01-1201.pdf>>.

Exhibit 2

Estimated Effect Sizes on Crime Outcomes

(A Negative Effect Size Indicates the Program Achieves Less Crime)

Program listed in <i>italics</i> require, in our judgment, additional research for it can be concluded that they do or do not reduce recidivism.	Number of Studies Included in the Review (total number of subjects in the treatment groups in the studies in parentheses)	Meta-Analytic Results Before Applying Institute Adjustments					Adjusted Effect Size Used in the Benefit-Cost Analysis (estimated effect after downward adjustments for the methodological quality of the evidence, outcome measurement relevance, and researcher involvement)
		Fixed Effects Model			Random Effects Model		
		Weighted Mean Effect Size	Homogeneity Test		Weighted Mean Effect Size		
		ES	p-value	p-value	ES	p-value	
Adult Offenders							
Programs for Drug-Involved Offenders							
Adult drug courts	56 (18957)	-.160	.000	.000	-.183	.000	-.094
In-prison therapeutic communities with community aftercare	6 (1989)	-.152	.000	.735	na	na	-.077
In-prison therapeutic communities without community aftercare	7 (1582)	-.119	.001	.079	na	na	-.059
Cognitive-behavioral therapy in prison	8 (3788)	-.130	.000	.905	na	na	-.077
<i>Case management in the community</i>	12 (2572)	-.046	.114	.000	-.039	.480	.000
Drug treatment in the community	5 (54334)	-.137	.000	.000	-.221	.007	-.109
Drug treatment in jail	9 (1436)	-.110	.008	.025	-.106	.094	-.052
Programs for Mentally Ill and Co-Occurring Offenders							
Jail diversion (pre & post booking programs)	11 (1243)	.060	.141	.682	na	na	.000
<i>Therapeutic community programs</i>	2 (145)	-.361	.004	.542	na	na	-.230
Treatment Programs for General Offenders							
Cognitive-behavioral for the general population	25 (6546)	-.147	.000	.000	-.164	.000	-.081
<i>Faith-based programs</i>	5 (630)	-.015	.767	.043	-.028	.728	.000
Programs for Domestic Violence Offenders							
Education/cognitive-behavioral treatment	9 (1254)	-.025	.523	.120	na	na	.000
<i>Domestic violence courts</i>	2 (327)	-.086	.309	.009	-.013	.956	.000
Programs for Sex Offenders							
Psychotherapy, sex offenders	3 (313)	.134	.179	.038	.027	.892	.000
Cognitive-behavioral treatment in prison	5 (894)	-.144	.005	.173	na	na	-.087
Cognitive-behavioral treatment in the community	6 (359)	-.391	.000	.438	na	na	-.195
Cognitive-behavioral treatment in prison (sex offense outcomes)	4 (705)	-.119	.027	.080	na	na	-.069
Cognitive-behavioral treatment in the community (sex off. outcomes)	5 (262)	-.357	.001	.846	na	na	-.177
<i>Intensive supervision of sex offenders in the community</i>	4 (392)	.207	.003	.000	.202	.359	.000
Behavioral Therapy - Sex Offenders.	2 (130)	-.190	.126	.635	na	na	.000
<i>Mixed Treatment-Sex Offenders in the Community</i>	2 (724)	-.176	.001	.015	-.184	.169	.000
<i>Circles of Support & Accountability (Faith-based supervision of sex offenders)</i>	1 (60)	-.388	.035	na	na	na	-.193
<i>Medical Treatment of Sex Offenders</i>	1 (99)	-.372	.060	na	na	na	-.185
Intermediate Sanctions							
Intensive supervision: surveillance-oriented approaches	24 (2699)	-.033	.244	.146	na	na	.000
Intensive supervision: treatment-oriented approaches	10 (2156)	-.287	.000	.000	-.291	.041	-.190
<i>Regular supervision compared to no supervision</i>	1 (22016)	-.010	.591	na	na	na	.000
<i>Day fines (compared to standard probation)</i>	1 (191)	-.084	.411	na	na	na	.000
Adult boot camps	22 (5910)	-.030	.103	.000	-.017	.632	.000
Electronic monitoring	12 (2175)	.025	.411	.025	.015	.765	.000
Restorative justice programs for lower risk adult offenders	6 (783)	-.077	.130	.013	-.125	.165	.000
Work and Education Programs for General Offenders							
Correctional industries programs in prison	4 (7178)	-.119	.000	.174	na	na	-.077
Basic adult education programs in prison	7 (2399)	-.094	.001	.006	-.114	.034	-.050
Employment training & job assistance programs in the community	16 (9217)	-.047	.003	.017	-.061	.021	-.047
<i>Work release programs from prison</i>	4 (621)	-.122	.045	.285	na	na	-.055
Vocational education in prison	3 (1950)	-.189	.000	.868	na	na	-.124

Notes to the Table:

Appendices 1, 2, and 3 describe the meta-analytic methods and decision criteria used to produce these estimates. Briefly, to be included in this review: 1) a study had to be published in English between 1970 and 2005; 2) the study could be published in any format—peer-reviewed journals, government reports, or other unpublished results; 3) the study had to have a randomly-assigned or demonstrably well-matched comparison group; 4) the study had to have intent-to-treat groups that included both completers and program dropouts, or sufficient information that the combined effects could be tallied; 5) the study had to provide sufficient information to code effect sizes; and 6) the study had to have at least a six-month follow-up period and include a measure of criminal recidivism as an outcome.

Exhibit 3

Citations to the Studies Used in the Meta-Analyses (Some studies contributed independent effect sizes from more than one location)

Program Grouping	Study
Adult Boot Camps	<p>Austin, J., Jones, M., & Bolyard, M. (1993). <i>Assessing the impact of a county operated boot camp: Evaluation of the Los Angeles County regimented inmate diversion program</i>. San Francisco: National Council on Crime and Delinquency.</p> <p>Burns, J. C., & Vito, G. F. (1995). An impact analysis of the Alabama boot camp program. <i>Federal Probation</i>, 59(1): 63-67.</p> <p>Camp, D. A., & Sandhu, H. S. (1995). Evaluation of female offender regimented treatment program (FORT). <i>Journal of the Oklahoma Criminal Justice Research Consortium</i>, 2: 50-77.</p> <p>Colorado Department of Corrections. (1993). <i>Colorado regimented inmate training program: A legislative report</i>.</p> <p>Farrington, D. P., Ditchfield, J., Hancock, G., Howard, P., Jolliffe, D., Livingston, M. S., & Painter, K. (2002). <i>Evaluation of two intensive regimes for young offenders</i>. Home Office Research Study 239. London, UK: Home Office</p> <p>Gransky, L. A. & Jones, R. J. (1995). <i>Evaluation of the post-release status of substance abuse program participants: The impact incarceration program at Dixon Springs and the Gateway substance abuse program at Dwight Correctional Center</i>. Chicago: Illinois Criminal Justice Authority Report.</p> <p>Harer, M. D., & Klein-Saffran, J. (1996). <i>Lewisburg ICC evaluation</i>. Washington DC: Bureau of Prisons, Office of Research and Evaluation, memo.</p> <p>Jones, M., & Ross, D. L. (1997). Is less better? Boot camp, regular probation and rearrest in North Carolina. <i>American Journal of Criminal Justice</i>, 21(2): 147-161.</p> <p>Kempinen, C. A., & Kurlychek, M. C. (2003). An outcome evaluation of Pennsylvania's boot camp: Does rehabilitative programming within a disciplinary setting reduce recidivism? <i>Crime and Delinquency</i>, 49(4): 581-602.</p> <p>MacKenzie, D. L. & Souryal, C. (1994). <i>Multisite evaluation of shock incarceration: Executive summary</i>. Washington, DC: U.S. Department of Justice/NIJ.</p> <p>Smith, R. P. (1998). Evaluation of the work ethic camp. Olympia: Washington State Department of Corrections.</p> <p>Stinchcomb, J. B., & Terry, W. C. (2001). Predicting the likelihood of rearrest among shock incarceration graduates: Moving beyond another nail in the boot camp coffin. <i>Crime and Delinquency</i>, 47(2): 221-242.</p> <p>Wright, D. T., & Mays, G. L. (1998). Correctional boot camps, attitudes, and recidivism: The Oklahoma experience. <i>Journal of Offender Rehabilitation</i>, 28(1/2): 71-87.</p>
Adult Drug Courts	<p>Barnoski, R., & Aos, S., (2003). <i>Washington State's drug courts for adult defendants: Outcome evaluation and cost-benefit analysis</i> (Document No. 03-03-1201). Olympia: Washington State Institute for Public Policy.</p> <p>Bavon, A. (2001). The effect of the Tarrant County drug court project on recidivism. <i>Evaluation and Program Planning</i>, 24: 13-24.</p> <p>Bell, M. M. (1998). <i>King County drug court evaluation: Final report</i>. Seattle, WA: M. M. Bell, Inc.</p> <p>Breckenridge, J. F., Winfree, Jr., L. T., Maupin, J. R., & Clason, D. L. (2000). Drunk drivers, DWI 'drug court' treatment, and recidivism: Who fails? <i>Justice Research and Policy</i>, 2(1): 87-105.</p> <p>Brewster, M. P. (2001). An evaluation of the Chester County (PA) drug court program. <i>Journal of Drug Issues</i>, 31(1): 177-206.</p> <p>Carey, S. M., & Finigan, M. W. (2004). A detailed cost-analysis in a mature drug court setting: A cost-benefit evaluation of the Multnomah County drug court. <i>Journal of Contemporary Criminal Justice</i>, 20(3): 315-338.</p> <p>Craddock, A. (2002). <i>North Carolina drug treatment court evaluation: Final report</i>. Raleigh: North Carolina Court System.</p> <p>Crumpton, D., Brekhus, J., Weller, J., & Finigan, M. (2003). <i>Cost analysis of Baltimore City, Maryland drug treatment court</i>. Portland, OR: NPC Research, Inc.</p> <p>Deschenes, E. P., Cresswell, L., Emami, V., Moreno, K., Klein, Z., & Condon, C. (2001). <i>Success of drug courts: Process and outcome evaluations in Orange County, California, final report</i>. Submitted to the Superior Court of Orange County, CA.</p> <p>Ericson, R., Welter, S., & Johnson, T. L. (1999). <i>Evaluation of the Hennepin County drug court</i>. Minneapolis: Minnesota Citizens Council on Crime and Justice.</p> <p>Spokane County Drug Court. (1999). <i>Evaluation: Spokane County drug court program</i>. Spokane, WA: Spokane County Drug Court.</p> <p>Fielding, J. E., Tye, G., Ogawa, P. L., Imam, I. J., & Long, A. M. (2002). Los Angeles County drug court programs: Initial results. <i>Journal of Substance Abuse Treatment</i>, 23(3): 217-224.</p> <p>Finigan, M. W. (1998). <i>An outcome program evaluation of the Multnomah County S.T.O.P. drug diversion program</i>. Portland, OR: NPC Research, Inc.</p> <p>Godley, M. D., Dennis, M. L., Funk, R., Siekmann, M., & Weisheit, R. (1998). <i>An evaluation of the Madison County assessment and treatment alternative court</i>. Chicago: Illinois Criminal Justice Information Authority.</p> <p>Goldkamp, J. S. & Weiland, D. (1993). <i>Assessing the impact of Dade County's felony drug court. Final report</i>. Philadelphia: Crime and Justice Research Institute.</p> <p>Goldkamp, J. S., Weiland, D., & Moore, J. (2001). <i>The Philadelphia treatment court, its development and impact: The second phase (1998-2000)</i>. Philadelphia: Crime and Justice Research Institute.</p> <p>Goldkamp, J. S., White, M. D., & Robinson, J. B. (2001). Do drug courts work? Getting inside the drug court black box. <i>Journal of Drug Issues</i>, 31(1): 27-72.</p> <p>Gottfredson, D. C., Najaka, S. S., & Kearley, B. (2002 November). <i>A randomized study of the Baltimore City drug treatment court: Results from the three-year follow-up</i>. Paper presented at the annual meeting of the American Society of Criminology, Chicago.</p> <p>Gottfredson, D. C., Coblenz, K., & Harmon, M. A. (1997). A short-term outcome evaluation of the Baltimore City drug treatment court program. <i>Perspectives</i>, Winter: 33-38.</p> <p>Granfield, R., Eby, C., & Brewster, T. (1998). An examination of the Denver drug court: The impact of a treatment-oriented drug-offender system. <i>Law & Policy</i>, 20: 183-202.</p> <p>Harrell, A., Roman, J., & Sack, E. (2001). <i>Drug court services for female offenders, 1996-1999: Evaluation of the Brooklyn treatment court</i>. Washington, DC: Urban Institute.</p> <p>Johnson, G. D., Formichella, C. M., & Bowers D. J. (1998). Do drug courts work? An outcome evaluation of a promising program. <i>Journal of Applied Sociology</i>, 15(1): 44-62.</p> <p>Latessa, E. J., Shaffer, D. K., & Lowenkamp C. (2002). <i>Outcome evaluation of Ohio's drug court efforts: Final report</i>. Cincinnati: Center for Criminal Justice Research, Division of Criminal Justice, University of Cincinnati.</p> <p>Listwan, S. J., & Latessa, E. J. (2003). <i>The Kootenai and Ada County drug courts: Outcome evaluation findings, final report</i>. Cincinnati: Center for Criminal Justice Research, University of Cincinnati.</p> <p>Listwan, S. J., Shaffer, D. K., & Latessa, E. J. (2001). <i>The Akron municipal drug court: Outcome evaluation findings</i>. Cincinnati: Center for Criminal Justice Research, University of Cincinnati.</p> <p>Listwan, S. J., Sundt, J. L., Holsinger, A. M., & Latessa, E. J. (2003). The effect of drug court programming on recidivism: The Cincinnati experience. <i>Crime and Delinquency</i>, 49(3): 389-411.</p> <p>Listwan, S. J., Shaffer, D. K., & Latessa, E. J. (2001). <i>The Erie County drug court: Outcome evaluation findings</i>. Cincinnati: Center for Criminal Justice Research, University of Cincinnati.</p> <p>Logan, T., Hoyt, W., & Leukefeld, C. (2001). <i>Kentucky drug court outcome evaluation: Behaviors, costs, and avoided costs to society</i>. Lexington: Center on Drug and Alcohol Research, University of Kentucky.</p>

Program Grouping	Study
Adult Drug Courts, continued	<p>Martin, T. J., Spohn, C. C., Piper, R. K., & Frenzel-Davis, E. (2001). <i>Phase III Douglas County drug court evaluation: Final report</i>. Washington, DC: Institute for Social and Economic Development.</p> <p>Martinez, A. I., & Eisenberg, M. (2003). <i>Initial process and outcome evaluation of drug courts in Texas</i>. Austin: Criminal Justice Policy Council.</p> <p>McNeece, C. A. & Byers, J. B. (1995). <i>Hillsborough County drug court: Two-year (1995) follow-up study</i>. Tallahassee: Institute for Health and Human Services Research, School of Social Work, Florida State University.</p> <p>Miethe, T. D., Lu, H., & Reese, E. (2000). Reintegrative shaming and recidivism risks in drug court: Explanations for some unexpected findings. <i>Crime and Delinquency</i>, 46(4): 522-541.</p> <p>Peters, R. H. & Murrin, M. R. (2000). Effectiveness of treatment-based drug courts in reducing criminal recidivism. <i>Criminal Justice and Behavior</i>, 27(1): 72-96.</p> <p>Rempel, M., Fox-Kralstein, D., Cissner, A., Cohen, R., Labriola, M., Farole, D., Bader, A., & Magnani, M. (2003). <i>The New York State adult drug court evaluation: Policies, participants and impacts</i>. New York, NY: Center for Court Innovation.</p> <p>Shanahan, M., Lancsar, E., Haas, M., Lind, B., Weatherburn, D., & Chen, S. (2004). Cost-effectiveness analysis of the New South Wales adult drug court program. <i>Evaluation Review</i>, 28(1): 3-27.</p> <p>Spohn, C., Piper, R. K., Martin, T., & Frenzel, E. D. (2001). Drug courts and recidivism: The results of an evaluation using two comparison groups and multiple indicators of recidivism. <i>Journal of Drug Issues</i>, 31(1): 149-176.</p> <p>Stageberg, P., Wilson, B., & Moore, R. G. (2001). <i>Final report on the Polk County adult drug court</i>. Iowa Department of Human Rights, Division of Criminal and Juvenile Justice Planning.</p> <p>Tjaden, C. D., Diana, A., Feldman, D., Dietrich, W., & Jackson, K. (2002). <i>Denver drug court: Second year report, outcome evaluation</i>. Vail, CO: Toucan Research and Computer Solutions.</p> <p>Truitt, L., Rhodes, W. M., Seeherman, A. M., Carrigan, K., & Finn, P. (2000). <i>Phase I: Case studies and impact evaluations of Escambia County, Florida and Jackson County, Missouri drug courts</i>. Cambridge, MA: Abt Associates. Some results also reported in Belenko, S. (2001). <i>Research on drug courts: A critical review, 2001 update</i>. New York: The National Center on Addiction and Substance Abuse at Columbia University.</p> <p>Turner, S., Greenwood, P., Fain, T., & Deschenes, E. (1999). Perceptions of drug court: How offenders view ease of program completion, strengths and weaknesses, and the impact on their lives. <i>National Drug Court Institute Review</i>, 2(1): 61-86.</p> <p>Utah Substance Abuse and Anti-Violence Coordinating Council. (2001). <i>Salt Lake County drug court outcome evaluation</i>. Salt Lake City: Utah Substance Abuse and Anti-Violence Coordinating Council.</p> <p>Vito, G. F., & Tewksbury, R. A. (1998). The impact of treatment: The Jefferson County (Kentucky) drug court program. <i>Federal Probation</i>, 62(2): 46-51.</p> <p>Wolfe E., Guydish J., & Termond J. (2002). A drug court outcome evaluation comparing arrests in a two year follow-up period. <i>Journal of Drug Issues</i>, 32(4): 1155-1171.</p>
Basic Adult Education Programs in Prison	<p>Drake, E. (2006). <i>Correctional education and its impacts on post-prison employment patterns and recidivism</i>. Draft report. Olympia: Washington State Institute for Public Policy and Washington State Department of Corrections.</p> <p>Harer, M. D. (1995). <i>Prison education program participation and recidivism: A test of the normalization hypotheses</i>. Washington, DC: Federal Bureau of Prisons, Office of Research and Evaluation.</p> <p>Mitchell, O. (2002). <i>Statistical analysis of the three state CEA data</i>. University of Maryland. Unpublished study.</p> <p>Piehl, A. M. (1994). <i>Learning while doing time</i>. Kennedy School Working Paper #R94-25. Cambridge, MA: John F. Kennedy School of Government, Harvard University.</p> <p>Walsh, A. (1985). An evaluation of the effects of adult basic education on rearrest rates among probationers. <i>Journal of Offender Counseling, Services, and Rehabilitation</i>, 9(4): 69-76.</p>
Behavioral Treatment for Sex Offenders	<p>Rice, M. E., Quinsey, V. L., Harris, G. T. (1991). Sexual recidivism among child molesters released from a maximum security psychiatric institution. <i>Journal of Consulting and Clinical Psychology</i>, 59: 381-386.</p> <p>Davidson, P. R. (1984 March). <i>Behavioral treatment for incarcerated sex offenders: Post-release outcome</i>. Paper presented at 1984 Conference on Sex Offender Assessment and Treatment, Kingston, Ontario, Canada.</p>
Case Management in the Community for Drug Involved Offenders	<p>Anglin, M. D., Longshore, D., & Turner, S. (1999). Treatment alternatives to street crime: An evaluation of five programs. <i>Criminal Justice and Behavior</i>, 26(2): 168-195.</p> <p>California Department of Corrections. (1996). <i>Parolee partnership program: A parole outcome evaluation</i>. Sacramento: California Department of Corrections.</p> <p>Hanlon, T. E., Nurco, D. N., Bateman, R. W., & O'Grady, K. E. (1999). The relative effects of three approaches to the parole supervision of narcotic addicts and cocaine abusers. <i>The Prison Journal</i>, 79(2): 163-181.</p> <p>Longshore, D., Turner, S., & Fain, T. (2005) Effects of case management on parolee misconduct. <i>Criminal Justice and Behavior</i>, 32(2): 205-222.</p> <p>Owens, S., Klebe, K., Arens, S., Durham, R., Hughes, J., Moor, C., O'Keefe, M., Phillips, J., Sarno, J., & Stommel, J. (1997). The effectiveness of Colorado's TASC programs. <i>Journal of Offender Rehabilitation</i>, 26: 161-176.</p> <p>Rhodes, W., & Gross, M. (1997). <i>Case management reduces drug use and criminality among drug-involved arrestees: An experimental study of an HIV prevention intervention</i>. Final report to the National Institute of Justice/National Institute on Drug Abuse. Cambridge, MA: Abt Associates Inc.</p>
Circles of Support and Accountability (faith-based supervision of sex offenders)	<p>Wilson, R. J., Picheca, J. E., & Prinzo, M. (2005). <i>Circles of support & accountability: An evaluation of the pilot project in South Central Ontario</i>. Draft report to Correctional Service of Canada, R-168, e-mailed to M. Miller Oct 20, 2005.</p>
Cognitive-Behavioral Therapy for General Population	<p>Armstrong, T. (2003). The effect of moral reconnection therapy on the recidivism of youthful offenders. <i>Criminal Justice and Behavior</i>, 30(6): 668-687.</p> <p>Burnett, W. (1997). Treating post-incarcerated offenders with moral reconnection therapy: A one-year recidivism study. <i>Cognitive Behavioral Treatment Review</i>, 6(3/4): 2.</p> <p>Culver, H. E. (1993). Intentional skill development as an intervention tool. (Doctoral dissertation. University of Massachusetts, 1993, UMI No. 9329590).</p> <p>Falshaw, L., Friendship, C., Travers, R., & Nugent, F. (2004). Searching for 'what works': HM Prison Service accredited cognitive skills programmes. <i>British Journal of Forensic Practice</i>, 6(2): 3-13.</p> <p>Friendship, C., Blud, L., Erikson, M., Travers, R., Thornton, D. (2003). Cognitive-behavioural treatment for imprisoned offenders: An evaluation of HM Prison Service's cognitive skills programmes. <i>Legal and Criminological Psychology</i>, 8: 103-114.</p> <p>Golden, L. (2002). <i>Evaluation of the efficacy of a cognitive behavioral program for offenders on probation: Thinking for a change</i>. Dallas: University of Texas Southwestern Medical Center at Dallas. Retrieved on December 22, 2005 from http://www.nicic.org/pubs/2002/018190.pdf.</p> <p>Grandberry, G. (1998). <i>Moral reconnection therapy evaluation, final report</i>. Olympia: Washington State Department of Corrections.</p> <p>Henning, K. R., & Frueh, B. C. (1996). Cognitive-behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Corrections' cognitive self-change program. <i>Criminal Justice and Behavior</i>, 23(4): 523-541.</p>

Program Grouping	Study
Cognitive-Behavioral Therapy for General Population, continued	<p>Hubbard, D. J., & Latessa, E. J. (2004). <i>Evaluation of cognitive-behavioral programs for offenders: A look at outcome and responsibility in five treatment programs, final report</i>. Cincinnati: Division of Criminal Justice, University of Cincinnati.</p> <p>Johnson, G. & Hunter, R. M. (1995). Evaluation of the specialized drug offender program. In R. R. Ross & R. D. Ross (Eds.), <i>Thinking straight: The reasoning and rehabilitation program for delinquency prevention and offender rehabilitation</i> (pp. 214-234). Ottawa, Canada: Air Training and Publications.</p> <p>Larson, K. A. (1989). Problem-solving training and parole adjustment in high-risk young adult offenders. <i>The Yearbook of Correctional Education</i> (1989):279-299.</p> <p>Little, G. L., Robinson, K. D., & Burnette, K. D. (1993). Cognitive behavioral treatment of felony drug offenders: A five-year recidivism report. <i>Psychological Reports</i>, 73: 1089-1090.</p> <p>Little, G. L., Robinson, K. D., & Burnette, K. D. (1993). 5-year recidivism results on MRT-treated DWI offenders released. <i>Cognitive Behavioral Treatment Review</i>, 2(4): 2.</p> <p>Little, G. L., Robinson, K. D., Burnette, K. D., & Swan, E. S. (1998). Nine-year reincarceration study on MRT-treated felony offenders: Treated offenders show significantly lower reincarceration. <i>Cognitive Behavioral Treatment Review</i>, 7(1): 2-3.</p> <p>Ortmann, R. (2000). The effectiveness of a social therapy in prison: A randomized design. <i>Crime and Delinquency</i>, 46(2): 214-232.</p> <p>Porporino, F. J. & Robinson, D. (1995). An evaluation of the reasoning and rehabilitation program with Canadian federal offenders. In R. R. Ross & R. D. Ross (Eds.), <i>Thinking straight: The reasoning and rehabilitation program for delinquency prevention and offender rehabilitation</i> (pp. 214-234). Ottawa: Air Training and Publications.</p> <p>Raynor, P. & Vanstone, M. (1996). Reasoning and rehabilitation in Britain: The results of the straight thinking on probation (STOP) programme. <i>International Journal of Offender Therapy and Comparative Criminology</i>, 40(4): 272-284.</p> <p>Robinson, D. (1995). <i>The impact of cognitive skills training on post-release recidivism among Canadian federal offenders</i>. Ottawa, Ontario: Correctional Research and Development, Correctional Service Canada.</p> <p>Ross, R. R., Fabiano, E. A., & Ewles, C. D. (1988). Reasoning and rehabilitation. <i>International Journal of Offender Therapy and Comparative Criminology</i>, 32: 29-36.</p> <p>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. <i>Criminal Justice and Behavior</i>, 31(3): 282-305.</p> <p>Van Voorhis, P., Spruance, L. M., Ritchey, P. N., Johnson-Listwan, S., Seabrook, R., & Pealer, J. (2002). <i>The Georgia cognitive skills experiment outcome evaluation phase II</i>. Cincinnati, OH: University of Cincinnati, Center for Criminal Justice Research. Retrieved December 22, 2005, from http://www.uc.edu/criminaljustice/ProjectReports/Georgia_Phase_II_final.report.pdf</p> <p>Wilkinson, J. (2005). Evaluating evidence for the effectiveness of the reasoning and rehabilitation programme. <i>The Howard Journal of Criminal Justice</i>, 44(1): 70-85.</p> <p>Yessine, A. K., & Kroner, D. G. (2004). <i>Altering antisocial attitudes among federal male offenders on release: A preliminary analysis of the counter-point community program</i> (Research Report No. R-152). Ottawa, Ontario: Correctional Research and Development, Correctional Service Canada.</p>
Cognitive-Behavioral Therapy in Prison for Drug Involved Offenders	<p>Aos, S., Phipps, P., Barnoski, R. (2004). <i>Washington's drug offender sentencing alternative: An evaluation of benefits and costs</i>. Olympia: Washington State Institute for Public Policy.</p> <p>Daley M., Love C. T., Shepard D. S., Petersen C. B., White K. L., & Hall, F. B. (2004). Cost-effectiveness of Connecticut's in-prison substance abuse treatment. <i>Journal of Offender Rehabilitation</i>, 39(3): 69-92.</p> <p>Hall, E. A., Prendergast, M. L., Wellisch, J., Patten, M., & Cao, Y. (2004). Treating drug-abusing women prisoners: An outcomes evaluation of the forever free program. <i>The Prison Journal</i>, 84(1): 81-105.</p> <p>Hanson, G. (2000). <i>Pine Lodge intensive inpatient treatment program</i>. Olympia: Washington State Department of Corrections.</p> <p>Pelissier, B., Rhodes, W., Saylor, W., Gaes, G., Camp, S. D., Vanyur, S. D., & Wallace, S. (2000). <i>TRIAD drug treatment evaluation project: Final report of three-year outcomes, Part 1</i>. Washington, DC: Federal Bureau of Prisons, Office of Research and Evaluation. Retrieved December 28, 2005, from http://www.bop.gov/news/PDFs/TRIAD/TRIAD_pref.pdf</p> <p>Porporino, F. J., Robinson, D., Millson, B., & Weekes, J. R. (2002). An outcome evaluation of prison-based treatment programming for substance users. <i>Substance Use & Misuse</i>, 37(8-10): 1047-1077.</p> <p>Washington State Department of Corrections. (1998). <i>Substance abuse treatment program evaluation of outcomes and management report</i>. Olympia: Washington State Department of Corrections, Division of Management and Budget, Planning and Research Section</p> <p>Wexler, H. K., Falkin, G. P., Lipton, D. S., & Rosenblum, A. B. (1992). Outcome evaluation of a prison therapeutic community for substance abuse treatment. In C. G. Leukefeld & F. M. Tims (Eds.), <i>Drug abuse treatment in prisons and jails</i>. NIDA research Monograph 118, Rockville, MD: NIDA (pp. 156-174).</p>
Cognitive-Behavioral Treatment in Prison for Sex Offenders	<p>Bakker, L., Hudson, S. W., & Riley, D. (1999). <i>...And there was light: An evaluation of the Kia Marama treatment programme for New Zealand sex offenders against children</i>. Unpublished report.</p> <p>Looman, J., Abracen, J., & Nicholaichuk, T. P. (2000). Recidivism among treated sexual offenders and matched controls: Data from the Regional Treatment Centre (Ontario). <i>Journal of Interpersonal Violence</i>, 15(3): 279-290.</p> <p>Marques, J. K. (1999). How to answer the question, does sex offender treatment work? <i>Journal of Interpersonal Violence</i>, 14(4): 437-451.</p> <p>Robinson, D. (1995). <i>The impact of cognitive skills training on post-release recidivism among Canadian federal offenders</i>. Research Report No. R-41. Ottawa, Ontario: Correctional Research and Development, Correctional Service Canada.</p> <p>Song, L. & Lieb, R. (1995). <i>Washington state sex offenders: Overview of recidivism studies</i>. Olympia: Washington State Institute for Public Policy.</p>
Cognitive-Behavioral Treatment in the Community for Sex Offenders	<p>Allam, J. (1999). <i>Sex offender re-conviction: Treated vs. untreated offenders</i>. West Midlands Probation Service Sex Offender Treatment Programme.</p> <p>Baird, C., Wagner, D., Decomo, B., & Aleman, T. (1994). <i>Evaluation of the effectiveness of supervision and community rehabilitation programs in Oregon</i>. Oakland, CA: National Council on Crime and Delinquency.</p> <p>Marshall, W. L., Eccles, A., & Barbaree, H.E. (1991). The treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. <i>Behaviour Research and Therapy</i>, 29(2): 129-135.</p> <p>McGrath, R. J., Hoke, S. E., & Vojtisek, J. E. (1998). Cognitive-behavioral treatment of sex offenders: A treatment comparison and long-term follow-up study. <i>Criminal Justice and Behavior</i>, 25: 203-225.</p> <p>Procter, E. (1996). A five-year outcome evaluation of a community-based treatment programme for convicted sexual offenders run by the probations service. <i>The Journal of Sexual Aggression</i>, 2(1): 3-16.</p>
Correctional Industries Programs in Prison	<p>Drake, E. (2003). <i>Class I impacts: Work during incarceration and its effects on post-prison employment patterns and recidivism</i>. Olympia: Washington State Department of Corrections.</p> <p>Saylor, R. G., & Gaes, G. G. (1996). <i>PREP: A study of "rehabilitating" inmates through industrial work participation, and vocational and apprenticeship training</i>. Washington, DC: U.S. Federal Bureau of Prisons.</p>
Domestic Violence Courts	<p>Newmark, L., Rempel, M., Diffily, K., Kane, K. M. (2001). <i>Specialized felony domestic violence courts: Lessons on implementations and impacts from the Kings County experience</i>. Washington DC: Urban Institute.</p>

Program Grouping	Study
Domestic Violence Courts, continued Drug Treatment in Jail	<p>Grover, A. R., MacDonald, J. M., Alpert, G. P., Geary, I. A., Jr. (2003). <i>The Lexington County domestic violence courts: A partnership and evaluation</i>. Columbia, SC: University of South Carolina. (National Institute of Justice Grant 2000-WT-VX-0015).</p> <p>Dugan J. R. & Everett, R. S. (1998). An experimental test of chemical dependency therapy for jail inmates. <i>International Journal of Offender Therapy and Comparative Criminology</i>, 42(4): 360-368.</p> <p>Knight, K., Simpson, D. D., & Hiller, M. L. (2003). <i>Outcome assessment of correctional treatment (OACT)</i>. Fort Worth: Texas Christian University, Institute of Behavioral Research. (National Institute of Justice Grant 99-RT-VX-KO27). Retrieved December 27, 2005, from http://www.ncjrs.org/pdffiles1/nij/grants/199368.pdf</p> <p>Peters, R. H., Kearns, W. D., Murrin, M. R., Dolente, A. S., & May, R. L. (1993). Examining the effectiveness of in-jail substance abuse treatment. <i>Journal of Offender Rehabilitation</i>, 19: 1-39.</p> <p>Taxman, F. S. & Spinner, D. L. (1997). <i>Jail addiction services (JAS) demonstration project in Montgomery County, Maryland: Jail and community based substance abuse treatment program model</i>. College Park: University of Maryland.</p> <p>Tunis, S., Austin, J., Morris, M., Hardyman, P. & Bolyard, M. (1996). <i>Evaluation of drug treatment in local corrections</i>. Washington DC: National Institute of Justice.</p>
Drug Treatment in the Community	<p>Baird, C., Wagner, D., Decomo, B., & Aleman, T. (1994). <i>Evaluation of the effectiveness of supervision and community rehabilitation programs in Oregon</i>. Oakland, CA: National Council on Crime and Delinquency.</p> <p>California Department of Corrections. (1997). <i>Los Angeles prison parole network: An evaluation report</i>. Sacramento: State of California.</p> <p>Hepburn, J. R. (2005). Recidivism among drug offenders following exposure to treatment. <i>Criminal Justice Policy Review</i>, 16(2): 237-259.</p> <p>Lattimore, P. K., Krebs, C. P., Koetse, W., Lindquist C., & Cowell A. J. 2005. Predicting the effect of substance abuse treatment on probationer recidivism. <i>Journal of Experimental Criminology</i>, 1(2): 159-189.</p>
Education/Cognitive-Behavioral Treatment for Domestic Violence Offenders	<p>Chen, H., Bersani, C., Myers, S. C., & Denton, R. (1989). Evaluating the effectiveness of a court sponsored abuser treatment program. <i>Journal of Family Violence</i>, 4(4): 309-322.</p> <p>Davis, R. C., Taylor, B. G., & Maxwell, C. D. (2000). <i>Does batterer treatment reduce violence? A randomized experiment in Brooklyn</i>. New York, NY: Victim Services Research. Retrieved December 27, 2005, from http://www.ncjrs.gov/pdffiles1/nij/grants/180772.pdf</p> <p>Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. <i>Journal of Consulting and Clinical Psychology</i>, 68(3): 468-476.</p> <p>Feder, L., & Forde, D. R. (2000). <i>A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment</i>. Memphis, TN: University of Memphis. (National Institute of Justice Grant 96-WT-NX-0008).</p> <p>Gordon, J. A. & Moriarty, L. J. (2003). The effects of domestic violence batterer treatment on domestic violence recidivism: The Chesterfield County experience. <i>Criminal Justice and Behavior</i>, 30(1): 118-134.</p> <p>Harrell, A. (1991). <i>Evaluation of court-ordered treatment for domestic violence offenders</i>. Washington, DC: Urban Institute.</p>
Electronic Monitoring	<p>Labriola, M., Rempel, M., & Davis, R.C. (2005). <i>Testing the effectiveness of batterer programs and judicial monitoring. Results from a randomized trial at the Bronx misdemeanor domestic violence court</i>. New York, NY: Center for Court Innovation. (National Institute of Justice Grant 2001-WT-BX-0506). Draft sent to M. Miller by M. Rempel.</p> <p>Baird, C., Wagner, D., Decomo, B., & Aleman, T. (1994). <i>Evaluation of the effectiveness of supervision and community rehabilitation programs in Oregon</i>. Oakland, CA: National Council on Crime and Delinquency.</p> <p>Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). Can electronic monitoring make a difference? An evaluation of three Canadian programs. <i>Crime and Delinquency</i>, 46(1): 61-75.</p> <p>Dodgson, K., Goodwin, P., Howard, P., Llewellyn-Thomas, S., Mortimer, E., Russell, N., & Weiner, M. (2001). <i>Electronic monitoring of released prisoners: An evaluation of the home detention curfew scheme</i>. Home Office Research Study 222. London: Home Office.</p> <p>Finn, M. A. & Muirhead-Steves, S. (2002). The effectiveness of electronic monitoring with violent male parolees. <i>Justice Quarterly</i>, 19(2): 293-312.</p> <p>Jolin, A. & Stipak, B. (1992). Drug treatment and electronically monitored home confinement: An evaluation of a community-based sentencing option. <i>Crime and Delinquency</i>, 38: 158-170.</p> <p>Jones, M., & Ross, D. L. (1997). Electronic house arrest and boot camp in North Carolina: Comparing recidivism. <i>Criminal Justice Policy Review</i>, 8(4): 383-404.</p> <p>Klein-Saffran, J. (1993). <i>Electronic monitoring versus halfway houses: A study of federal offenders</i>. (Doctoral dissertation. University Of California, Berkeley, 1993, UMI No. 9327445).</p> <p>Petersilia, J., Turner, S., & Deschenes, E. P. (1992). Intensive supervision programs for drug offenders. In J. M. Byrne, A. J. Lurigio, & J. Petersilia (Eds.), <i>Smart sentencing: The emergency of intermediate sanctions</i> (pp. 18-37). Newbury Park, CA: Sage.</p> <p>Petersilia, J. & Turner, S. (1990). <i>Intensive supervision for high-risk probationers: Findings from three California experiments</i>. Santa Monica, CA: RAND.</p> <p>Sugg, D., Moore, L., & Howard, P. (2001). Electronic monitoring and offending behaviour—reconviction results for the second year of trials of curfew orders. Home Office Research Findings 141. London: Home Office.</p>
Employment Training and Job Assistance Programs in the Community	<p>Anderson, D. B. & Schumacker, R. E. (1986). Assessment of job training programs. <i>Journal of Offender Counseling, Services, & Rehabilitation</i>, (10): 41-49.</p> <p>Beck, J. (1981). Employment, community treatment center placement and recidivism: A study of released federal offenders. <i>Federal Probation</i>, (45): 3-8.</p> <p>Beck, L. (1979). An evaluation of federal community treatment centers. <i>Federal Probation</i>, (43): 36-40.</p> <p>Berk, R. A., Lenihan, K. J., & Rossi, P. H. (1980). Crime and poverty: Some experimental evidence from ex-offenders. <i>American Sociological Review</i>, (45): 766-786.</p> <p>Bloom, H., Orr, L. O., Cave, G., Bell, S. H., Doolittle, F., & Lin, W. (1994). <i>The national JTPA study. Overview: Impacts, benefits and costs of Title II-A</i>. Cambridge, MA: Abt Associates Inc.</p> <p>Cave, G., Bos, H., Doolittle, F., & Toussaint, C. (1993). <i>Jobstart: Final report on a program for school dropouts</i>. New York, NY: Manpower Demonstration and Research Corporation.</p> <p>Mallar, C. D., & Thornton, C. (1978). Transitional aid for released prisoners: Evidence from the life experiment. <i>The Journal of Human Resources</i>, XIII(2): 208-236.</p> <p>Menon, R., Blakely, C., Carmichael, D., & Snow, D. (1995). Making a dent in recidivism rates: Impact of employment on minority ex-offenders. In G. E. Thomas (Ed.). <i>Race and ethnicity in America: Meeting the challenge in the 21st century</i> (pp. 279-293). Washington, DC: Taylor and Francis. See also, Finn, P. (1998). <i>Texas' Project RIO (re-integration of offenders)</i>. Washington, DC: U.S. Department of Justice.</p> <p>Milkman, R. H. (1985). <i>Employment services for ex-offenders field test—detailed research results</i>. McLean, VA: Lazar Institute.</p> <p>Rossman, S., Sridharan, S., Gouvis, C., Buck, J., & Morley, E. (1999). <i>Impact of the opportunity to succeed program for substance-abusing felons: Comprehensive final report</i>. Washington, DC: Urban Institute.</p> <p>Schochet, P. Z., Burghardt, J., & Glazerman, S. (2001). <i>National job corps study: The impacts of job corps on participants' employment and related outcomes</i>. Princeton, NJ: Mathematica Policy Research, Inc. (U.S. Department of Labor, Employment and Training Administration contract No. K-4279-3-00-80-30).</p>

Program Grouping	Study
Employment Training and Job Assistance Programs in the Community, continued Faith-Based Programs for General Offenders	<p>Uggen, C. (2000). Work as a turning point in the life course of criminals: A duration model of age, employment, and recidivism. <i>American Sociological Review</i>, 67(4): 529-546.</p> <p>Burnside, J., Adler, J., Loucks, N., & Rose, G. (2001). Kainos community in prisons: Report of an evaluation. RDS OLR 11/01. Presented to Research Development and Statistics Directorate, Home Office, HM Prison Service England and Wales and Kainos Community. Retrieved December 27, 2005, from http://www.homeoffice.gov.uk/rds/pdfs/kainos_finalrep.pdf</p> <p>Johnson, B.R. (2004). Religious programs and recidivism among former inmates in prison fellowship programs: A long-term follow-up study. <i>Justice Quarterly</i>, 21(2): 329-354.</p> <p>O'Connor, T., Su, Y., Ryan, P., Parikh, C., & Alexander, E. (1997). <i>Detroit transition of prisoners: Final evaluation report</i>. Center for Social Research, MD</p> <p>Trusty, B., & Eisenberg, M. (2003). <i>Initial process and outcome evaluation of the innerchange freedom initiative: The faith-based prison program in TDCJ</i>. Austin, TX: Criminal Justice Policy Council. Retrieved December 27, 2005, from http://www.cjpc.state.tx.us/reports/adltrehab/IFIInitiative.pdf</p> <p>Wilson, L.C., Wilson, C., Drummond, S. R., & Kelso, K. (2005). <i>Promising effects on the reduction of criminal recidivism: An Evaluation of the Detroit transition of prisoner's faith based initiative</i>. Draft report emailed to Marna Miller by Joe Williams.</p>
In-Prison Therapeutic Communities With Community Aftercare for Drug Involved Offenders	<p>Field, G. (1985). The cornerstone program: A client outcome study. <i>Federal Probation</i>, 49: 50-55.</p> <p>Inciardi, J. A., Martin S. S., & Butzin, C. A. (2004). Five-year outcomes of therapeutic community treatment of drug-involved offenders after release from prison. <i>Crime and Delinquency</i>, 50(1): 88-107.</p> <p>Knight, K., Simpson, D. D., & Hiller, M. L. (1999). Three-year reincarceration outcomes for in-prison therapeutic community treatment in Texas. <i>The Prison Journal</i>, 79(3): 337-351.</p> <p>Prendergast, M. L., Hall, E. A., Wexler, H. K., Melnick, G., & Cao, Y. (2004). Amity prison-based therapeutic community: 5-year outcomes. <i>The Prison Journal</i>, 84(1): 36-60.</p> <p>Swartz, J. A., Lurigo, A. J., & Slomka, S. A. (1996). The impact of IMPACT: An assessment of the effectiveness of a jail-based treatment program. <i>Crime and Delinquency</i>, 42(4): 553-573.</p> <p>Wexler, H. K., Falkin, G. H., Lipton, D. S., & Rosenblum, A. B. (1992). Outcome evaluation of a prison therapeutic community for substance abuse treatment. <i>Criminal Justice and Behavior</i>, 17(1): 71-92.</p>
In-Prison Therapeutic Communities Without Community Aftercare for Drug Involved Offenders	<p>Belenko, S., Foltz, C., Lang, M. A., & Sun, H. (2004). Recidivism among high-risk drug felons: A longitudinal analysis following residential treatment. <i>Journal of Offender Rehabilitation</i>, 40(1/2): 105-132.</p> <p>Gransky, L. A. & Jones, R. J. (1995). <i>Evaluation of the post-release status of substance abuse program participants: The impact incarceration program at Dixon Springs and the Gateway substance abuse program at Dwight Correctional Center</i>. Chicago: Illinois Criminal Justice Authority Report.</p> <p>Klebe, K. J., & O'Keefe, M. (2004). Outcome evaluation of the crossroads to freedom house and peer I therapeutic communities. Colorado Springs: University of Colorado. (National Institute of Justice Grant 99-RT-VX-K021).</p> <p>Mosher, C., & Phillips, D. (2002). <i>Program evaluation of the pine lodge pre-release residential therapeutic community for women offenders in Washington State, final report</i>. Pullman, WA: Washington State University. (National Institute of Justice Grant 99-RT-VX-K001).</p> <p>Oregon Department of Corrections. (1996). <i>Evaluation of the powder river and turning point alcohol and drug treatment programs</i>. Salem, OR: Oregon Department of Corrections.</p> <p>Welsh, W. N. (2003). <i>Evaluation of prison-based therapeutic community drug treatment programs in Pennsylvania</i>. Philadelphia, PA: Pennsylvania Commission on Crime and Delinquency.</p>
Intensive Supervision Of Sex Offenders in the Community	<p>Stalans, L.J., Seng, M., Yarnold, P., Lavery, T., & Swartz, J. (2001). <i>Process and initial evaluation of the Cook County adult probation department's sex offender program: Final and summary report for the period of June, 1997 to June, 2000</i>. Chicago: Illinois Criminal Justice Information Authority. Retrieved on December 28, 2005 from http://www.icjia.state.il.us/public/pdf/researchreports/An%20Implementation_Project%20in%20Cook%20County.pdf</p> <p>Stalans, L.J., Seng, M., & Yarnold, P.R. (2002). <i>Long-term impact evaluation of specialized sex offender probation programs in Lake, DuPage and Winnebago Counties</i>. Chicago: Illinois Criminal Justice Information Authority. Retrieved on December 28, 2005, from http://www.icjia.state.il.us/public/pdf/ResearchReports/Long-termDuPageWinnebago.pdf</p>
Intensive Supervision: Surveillance-Oriented Approaches	<p>Bagdon, W. & Ryan, J. E. (1993). Intensive supervision of offenders on prerelease furlough: An evaluation of the Vermont experience. <i>Forum on Corrections Research</i>, 5(2). Retrieved on December 28, 2005, from http://www.csc-scc.gc.ca/text/pblct/forum/e052/e052j_e.shtml</p> <p>Brown, K. (1999). Intensive supervision probation: The effects of supervision philosophy on intensive probationer recidivism. (Doctoral dissertation, University of Cincinnati, 1993). Retrieved on December 28, 2005, from http://www.uc.edu/criminaljustice/graduate/Dissertations/KBrown.PDF</p> <p>Byrne, J. M. & Kelly, L. (1989). <i>Restructuring probation as an intermediate sanction: An evaluation of the Massachusetts intensive probation supervision program, Executive summary</i>. Final Report to the National Institute of Justice, Research Program on the Punishment and Control of Offenders. Washington, DC: National Institute of Justice.</p> <p>Deschenes, E. P., Turner, S., & Petersilia, J. (1995). A dual experiment in intensive community supervision: Minnesota's prison diversion and enhanced supervised release programs. <i>Prison Journal</i>, 75(3): 330-357.</p> <p>Erwin, B. S., & Bennett, L. A. (1987). <i>New dimensions in probation: Georgia's experience with intensive probation supervision</i>. Research in Brief. Washington, DC: National Institute of Justice.</p> <p>Fulton, B., Stichman, A., Latessa, E., & Lawrence, T. (1998). <i>Evaluating the prototypical ISP, Iowa correctional services second judicial district. Final Report</i>. Cincinnati: Division of Criminal Justice, University of Cincinnati.</p> <p>Johnson, G. & Hunter, R. M. (1995). Evaluation of the specialized drug offender program. In R. R. Ross & R. D. Ross (Eds.), <i>Thinking straight: The reasoning and rehabilitation program for delinquency prevention and offender rehabilitation</i> (pp. 214-234). Ottawa, Canada: Air Training and Publications.</p> <p>Lichtman, C. M. & Smock, S. M. (1981). The effects of social services on probationer recidivism: A field experiment. <i>Journal of Research in Crime and Delinquency</i>, 18: 81-100.</p> <p>Pearson, F. S. (1988). Evaluation of New Jersey's intensive supervision program. <i>Crime and Delinquency</i>, 34(4): 437-448.</p> <p>Petersilia, J., Turner, S., & Deschenes, E. P. (1992). Intensive supervision programs for drug offenders. In J. M. Byrne, A. J. Lurigio, & J. Petersilia (Eds.), <i>Smart sentencing: The emergency of intermediate sanctions</i> (pp. 18-37). Newbury Park, CA: Sage.</p> <p>Petersilia, J. & Turner, S. (1990). <i>Intensive supervision for high-risk probationers: Findings from three California experiments</i>. Santa Monica, CA: RAND.</p> <p>Smith, L. G. & Akers, R. L. (1993). A comparison of recidivism of Florida's community control and prison: A five-year survival analysis. <i>Journal of Research in Crime and Delinquency</i>, 30(3): 267-292.</p> <p>Stichman, A., Fulton, B., Latessa, E., & Lawrence, T. (1998). <i>Evaluating the prototypical ISP, Hartford intensive supervision unit Connecticut office of adult probation administrative office of the courts</i>. Final Report. Cincinnati: Division of Criminal Justice, University of Cincinnati.</p> <p>Turner, S., & Petersilia, J. (1992). Focusing on high-risk parolees: Experiment to reduce commitments to the Texas Department of Corrections. <i>Journal of Research in Crime and Delinquency</i>, 29(1): 34-61.</p>

Program Grouping	Study
Intensive Supervision: Treatment-Oriented Approaches	<p>Deschenes, E. P., Turner, S., & Petersilia, J. (1995). A dual experiment in intensive community supervision: Minnesota's prison diversion and enhanced supervised release programs. <i>Prison Journal</i>, 75(3): 330-357.</p> <p>Hanley, D. (2002). Risk differentiation and intensive supervision: A meaningful union? (Doctoral dissertation, University of Cincinnati, 2002, UMI No. 3062606).</p> <p>Harrell, A., Mitchell, O., Hirst, A., Marlow, D., & Merrill, J. C. (2002). Breaking the cycle of drugs and crime: Findings from the Birmingham BTC demonstration. <i>Criminology and Public Policy</i>, 1(2): 189-216.</p> <p>Harrell, A., Mitchell, O., Merrill, J. C., & Marlowe, D. B. (2003). <i>Evaluation of breaking the cycle</i>. Washington, DC: Urban Institute. (National Institute of Justice Grant 97-IJ-CX-0013).</p> <p>Harrell, A., Roman, J., Bhati, A., & Parthasarathy, B. (2003). <i>The impact evaluation of the Maryland break the cycle initiative</i>. Washington, DC: Urban Institute.</p> <p>Paparozi, M. A. (1994). A comparison of the effectiveness of an intensive parole supervision program with traditional parole supervision. (Doctoral Dissertation. Rutgers the State University of New Jersey – New Brunswick, 1994, UMI No. 9431121).</p> <p>Petersilia, J. & Turner, S. (1990). <i>Intensive supervision for high-risk probationers: Findings from three California experiments</i>. Santa Monica, CA: RAND.</p>
Jail Diversion: Pre- and Post-Booking Programs for MICA Offenders	<p>Broner, N., Lattimore, P. K., Cowell, A. J., & Schlenger, W. E. (2004). Effects of diversion on adults with co-occurring mental illness with substance use: Outcomes from a national multi-site study. <i>Behavioral Sciences and the Law</i>, (22): 519-541</p> <p>Christy, A., Poythress, N. G., Boothroyd, R. A., Petrila, J., Mehra, S. (2005). Evaluating the efficiency and community safety goals of the Broward County Mental Health Court. <i>Behavioral Sciences and the Law</i>, 23(2):227-243.</p> <p>Cosden, M., Ellens, J., Schnell, J., & Yamini-Diouf, J. (2004). <i>Evaluation of the Santa Barbara County mental health treatment court with intensive case management</i>. Santa Barbara: University of California.</p> <p>Steadman, H. J., Cocozza, J. J., & Veysey, B. M. (1999). Comparing outcomes for diverted and nondiverted jail detainees with mental illnesses. <i>Law and Human Behavior</i>, 23(6): 615-627.</p>
Medical Treatment of Sex Offenders Psychotherapy for Sex Offenders	<p>Wille, R., & Beier, K. M. (1989). Castration in Germany. <i>Annals of Sex Research</i>, 2: 103-133.</p> <p>Hanson, R. K., Steffy, R. A. & Gauthier, R. (1993). Long-term recidivism of child molesters. <i>Journal of Consulting and Clinical Psychology</i>, 61: 646-652.</p> <p>Nutbrown, V., and Stasiak, E. (1987). <i>A retrospective analysis of O.C.I. cost effectiveness 1977-1981</i>. (Ontario Correctional Institute Research Monograph No. 2). Brampton, Ontario, Canada: Ontario Ministry of Correctional Services Ontario Correctional Institute.</p> <p>Romero, J. J. & Williams, L. M. (1983). Group psychotherapy and intensive probation supervision with sex offenders: A comparative study. <i>Federal Probation</i>, 47: 36-42.</p>
Regular Supervision Compared to No Supervision Restorative Justice Programs for Lower Risk Adult Offenders	<p>Solomon, A. L., Kachnowski, V., Bhati, A. (2005). <i>Does parole work? Analyzing the impact of postprison supervision on rearrest outcomes</i>. Washington, DC: Urban Institute.</p> <p>Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. <i>Criminal Justice and Behavior</i>, 27(3): 312-329.</p> <p>Dignan, J. (1990). <i>Repairing the damage: An evaluation of an experimental adult reparation scheme in Kettering, Northamptonshire</i>. Sheffield, UK: Centre for Criminological and Legal Research, Faculty of Law, University of Sheffield.</p> <p>Crime and Justice Research Centre Victoria University of Wellington, & Triggs, S. (2005). <i>New Zealand court-referred restorative justice pilot: Evaluation</i>. Wellington, New Zealand: Ministry of Justice. Retrieved on December 27, 2005, from http://www.justice.govt.nz/pubs/reports/2005/nz-court-referred-restorative-justice-pilot-evaluation</p> <p>Paulin, J., Kingi, V., Lash, B. (2005). <i>The Rotorua second chance community-managed restorative justice programme: An evaluation</i>. Wellington, New Zealand: Ministry of Justice. Retrieved on December 27, 2005, from http://www.justice.govt.nz/pubs/reports/2005/rotorua-second-chance-community-managed-restorative-justice/index.html</p> <p>Paulin, J., Kingi, V., Lash, B. (2005). <i>The Wanganui community-managed restorative justice programme: An evaluation</i>. Wellington, New Zealand: Ministry of Justice. Retrieved on December 27, 2005, from http://www.justice.govt.nz/pubs/reports/2005/wanganui-community-managed-restorative-justice</p> <p>Rugge, T., Bonta, J., & Wallace-Capretta, S. (2005). <i>Evaluation of the collaborative justice project: A restorative justice program for serious crime</i>. Ottawa, Ontario, Canada: Public Safety and Emergency Preparedness Canada.</p>
Therapeutic Community Programs for MICA Offenders	<p>Sacks, S., Sacks, J. Y., McKendrick, K., Banks, S., & Stommel, J. (2004). Modified TC for MICA offenders: Crime outcomes. <i>Behavioral Sciences and the Law</i>, 22(4): 477-501.</p> <p>Van Stelle, K. R., & Moberg, D. P. (2004). Outcome data for MICA clients after participation in an institutional therapeutic community. <i>Journal of Offender Rehabilitation</i>, 39(1): 37-62.</p>
Vocational Education in Prison	<p>Callan, V. & Gardner, J. (2005). <i>Vocational education and training provision and recidivism in Queensland correctional institutions</i>. Queensland, Australia: National Center for Vocational Education Research (NCVER).</p> <p>Lattimore, P. K., Witte, A. D., & Baker, J. R. (1990). Experimental assessment of the effect of vocational training on youthful property offenders. <i>Evaluation Review</i>, 14(2): 115-133.</p> <p>Saylor, W. G., & Gaes, G. G. (1996). <i>PREP: A study of "rehabilitating" inmates through industrial work participation, and vocational and apprenticeship training</i>. Washington, DC: U.S. Federal Bureau of Prisons.</p>
Work Release Programs From Prison	<p>Jeffrey, R. & Woolpert, S. (1974). Work furlough as an alternative to incarceration. <i>The Journal of Criminology</i>, 65(3): 405-415.</p> <p>LeClair, D. P. & Guarino-Ghezzi, S. (1991). Does incapacitation guarantee public safety? Lessons from the Massachusetts furlough and prerelease program. <i>Justice Quarterly</i>, (8)1: 9-36</p> <p>Turner, S. M. & Petersilia, J. (1996). Work release in Washington: Effects on recidivism and corrections costs. <i>Prison Journal</i>, 76(2): 138-164.</p> <p>Waldo, G. & Chiricos, T. (1977). Work release and recidivism: An empirical evaluation of a social policy. <i>Evaluation Quarterly</i>, 1(1): 87-108.</p>
Correctional Industries Programs in Prison	<p>Maguire, K. E., Flanagan, T. J., & Thornberry, T. P. (1988). Prison labor and recidivism. <i>Journal of Quantitative Criminology</i>, 4(1): 3-18.</p> <p>Smith, C. J., Bechtel, J., Patricky, A., & Wilson-Gentry, L. (2005). <i>Correctional industries preparing inmates for re-entry: Recidivism & post-release employment</i>. Final draft report. Email from author.</p>
Day Fines (Compared to Standard Probation)	<p>Turner, S. & Greene, J. (1999). The FARE probation experiment: implementation and outcomes of day fines for felony offenders in Maricopa County. <i>The Justice System Journal</i>, 21(1): 1-21.</p>

Document No. 06-01-1201



**Washington State
Institute for
Public Policy**

The Washington State Legislature created the Washington State Institute for Public Policy in 1983. A Board of Directors—representing the legislature, the governor, and public universities—governs the Institute and guides the development of all activities. The Institute's mission is to carry out practical research, at legislative direction, on issues of importance to Washington State.

The effects of community sanctions and incarceration on recidivism

By Paul Gendreau and Claire Goggin¹

Centre for Criminal Justice Studies, University of New Brunswick

Francis T. Cullen²

Division of Criminal Justice, University of Cincinnati
and Donald A. Andrews³

Department of Psychology, Carleton University

Within recent years "get tough" strategies have become the latest panacea for dealing with offenders. This article quantitatively summarizes a substantial body of literature that assesses the effectiveness of two types of "get tough" programmes: community sanctions and incarceration. A brief history of the development of these initiatives is provided accompanied by a meta-analytic summary of the data.

Community sanctions

At one time, some of the services provided in probation and parole settings adhered to a dynamic rehabilitative model wherein it was gratifying to discover well-conceptualized programmes of sound therapeutic integrity.⁴ Reductions in recidivism of 20%-60% were reported for some of these programmes.

What kinds of programmes were these? First, treatment staff conformed to the principles and techniques of the therapies they were employing. Secondly, staff were carefully monitored by the programme developers who themselves had excellent skills in behavioural treatment and their assessments, with ongoing training being frequently provided. Thirdly, offenders' individual differences relative to varying styles of service delivery were considered. Finally, the programmes were intense; contact between offenders and therapist was frequent and focussed on learning pro-social skills.

The following three programmes best illustrate the above. The first of these, by Walter and Mills,⁵ was a behavioural employment programme for juvenile probationers utilizing a token economy, contingency contracting, and life skills interventions. The programme was admirable in that its treatment design intimately linked the courts with community-based employers who were trained as paraprofessional behaviour modifiers. The second example came from Andrews and Kiessling's⁶ Canadian Volunteers in Corrections Programme which combined professionals with paraprofessionals in an adult probation supervision programme. The major features of the counselling and supervision practices were the use of authority, anti-criminal

modelling and reinforcement, and problem-solving techniques. The quality of interpersonal relationships was also considered when pairing offenders with probation and parole officers. The theoretical importance of this study should not be understated as the treatment guidelines employed herein were instrumental to the continuing development of the principles of effective correctional treatment literature.⁷

Thirdly, there was a series of studies by Davidson, Robinson, and Seidman that featured an amalgam of behavioural techniques, relationships skills training, child advocacy, and matching of offenders and therapists.⁸ As community psychologists, they were among the first researchers to be aware of the need to overcome system-based barriers in delivering effective interventions.

Just when it seemed, however, that progress was being made in the confirmation and promulgation of effective services for probation and parole, a counterrevolution began to evolve: the new epoch of punishment-based strategies.⁹ The reasons why this new epoch gained favour is reviewed elsewhere.¹⁰ With the exception of occasional reports of successful intervention programmes in probation and parole, distinct forms of "get tough" strategies known as intermediate sanctions began to proliferate in probation and parole settings. The term "intermediate" was derived from the notion that deterrence strategies based on excessive use of incarceration were too crude and expensive while regular probation (with or without treatment services), on the other hand, was too "soft". Interestingly, some proponents of intermediate sanctions asserted that probation could be even more punishing than prison.¹¹ The most common form of intermediate sanction was intensive supervision programming (ISP). As Billie J. Erwin so forcefully put it when referring to the Georgia ISP, considered by many to be a model for the United States: "...We are in the business of increasing the heat on probationers...satisfying the public's demand for just punishment...Criminals must be punished for their misdeeds".¹²

This new generation of ISPs quickly spread throughout the United States, and to a much lesser extent, within Canada. They turned up the heat by: greatly increasing contact between supervisors and offenders; confining offenders to their homes; enforcing curfews; submitting offenders to random drug testing; requiring offenders to pay restitution to victims; electronically monitoring offenders, and requiring offenders to pay for the privilege of being supervised. Most ISPs have employed arbitrary combinations of the above sanction types in varying degrees with the major emphasis for most being an increase in the frequency of offender-probation/parole contacts. Boot camps and quick/brief arrests or citations, often in response to spousal abuse offences, are other types of sanctions that may fall under the intermediate sanctions umbrella.

Besides serving an underlying retributive purpose and reducing prison overcrowding costs, an important expectation was that ISPs would effect pro-social conformity through the threat of punishment.¹³

How well are intermediate sanctions working? So far they appear to be "widening the net" by targeting low-risk offenders who would normally receive periods of regular probation. The data indicate that the use of intermediate sanctions can increase the number of technical violations and lead to higher rates of incarceration.¹⁴ As to recidivism, we found little evidence of the effectiveness of intermediate sanctions among this sample of studies. These results are illustrated in Table 1. Of note, a positive correlation indicates that the sanction was associated with an increase in recidivism while a negative correlation means the sanction has suppressed or decreased recidivism. Within

Category 1, ISPs, there were 47 comparisons of the recidivism rates of offenders in an ISP with those receiving regular probation. These comparisons involved 19,403 offenders with a mean treatment effect of .00, expressed as a phi coefficient (Φ), indicating no difference in percentage recidivism rates between the two groups. The recidivism rate for each of the ISP and comparison groups was 29%.

The confidence interval (CI) is a useful index of the likelihood that a given range of values will contain the "true" population parameter. In the case of ISPs, the CI about Φ is -.05 to .05, reflecting recidivism rates ranging from a 5 per cent reduction ($\Phi = -.05$) to a 5 per cent increase ($\Phi = .05$). Also of note, when a CI contains 0, one can infer a lack of significant treatment effects ($p > .05$).

The z^{\pm} value is a weighted estimate of Φ . That is, each effect size is weighted by the inverse of its variance ($\sqrt{N - 3^k}$) thereby giving more emphasis to effect sizes generated with larger sample sizes. The z^{\pm} for ISPs indicates that they were associated with a 6% increase in recidivism with an associated CI of .04 to .07.

Upon examining the mean Φ and z^{\pm} values for each of the eight types of intermediate sanctions, one can see that 13 of the 16 CI s contain 0. Only in the case of restitution and fines was there any indication of a suppression of recidivism (i.e., CI did not include 0) but these results were criterion-dependent. A summary of the data from all of the eight categories produced mean effect sizes of .00 with a CI of -.02 to .03 for Φ , and .02 for z^{\pm} with an associated CI of .01 to .03.

In fact, an examination of the effect sizes from intermediate sanctions that purported to provide a

Table 1

Mean Effect of Community Sanctions on Recidivism

Type of Sanction (k)	N	%E	%C	M Φ	CI Φ	Z [±]	CI Z [±]
1. Intensive Supervision Programs (47)	19,403	29	29	.00	-.05 to .05	.06	.04 to .07
2. Arrest (24)	7,779	38	39	.01	-.05 to .04	.00	-.02 to .02
3. Fines (18)	7,162	41	45	-.04	-.08 to .00	-.04	-.06 to -.02
4. Restitution (17)	8,715	39	40	-.02	-.15 to -.01	.03	-.01 to .05
5. Boot Camp (13)	6,831	31	30	.00	-.05 to .08	.00	-.02 to .02
6. Scared Straight (12)	1,891	46	37	.07	-.05 to .18	.04	-.01 to .09
7. Drug Testing (3)	419	13	12	.05	-.12 to .12	.00	-.10 to .10
8. Electronic Monitoring (6)	1,414	6	4	.05	-.02 to .11	.03	-.02 to .08
9. Total (140)	53,614	33	33	.00	-.02 to .03	.02	-.01 to .03

Note. k = number of effect sizes per type of sanction; N = total sample size per type of treatment; %E = percentage recidivism for the group receiving the sanction; %C = percentage recidivism for the comparison group (regular probation); M Φ = mean phi per type of sanction; CI Φ = confidence interval about mean phi; z[±] = weighted estimation of phi per type of sanction; CI z[±] = confidence interval about z[±].

modicum of "treatment" — in each case the treatment was ill-defined and, therefore, impossible to assess as to quality — an interesting result was uncovered. The addition of a treatment component produced a 10% reduction in recidivism. On this evidence, one can tentatively conclude that the effectiveness of intermediate sanctions is mediated solely through the provision of treatment.

Incarceration

The view that the experience of prison in itself acts as a deterrent has a long history in criminal justice.¹⁵ It is rooted in specific deterrence theory,¹⁶ which predicts that individuals experiencing a more severe sanction are more likely to reduce their criminal activities in the future. Research strongly indicates that both the public and many policy-makers assume incarceration has powerful deterrent effects. Amongst academics, economists have taken the lead in support of the specific deterrence model.¹⁷ They maintain that incarceration imposes direct and indirect costs on inmates (e.g., loss of income, stigmatization) such that, faced with the prospect of going to prison or after having experienced prison life, the rational individual would choose not to engage in further criminal activities.

What kind of data is used to support the hypothesis that prison time suppresses criminal behaviour? The most compelling evidence comes from some ecological studies where the results are based on rates or averages (aggregate data). An example of one of the most positive results came from a study by Fabelo¹⁸ that reported a 30% increase in incarceration rates across 50 U.S. states,

The addition of this body of evidence to the "what works" debate leads to the inescapable conclusion that, when it comes to reducing individual offender recidivism, the "only game in town" is appropriate cognitive-behavioural treatments which embody known principles of effective intervention.

corresponding with a decrease of 5% in the crime rate for a five-year period. Fabelo's data has been interpreted as convincing evidence that prisons deter crime.

To be fair to deterrence aficionados, we must acknowledge that there are a number of caveats about the potency of prison in this regard. These include the following: deterrent effects are more likely to be found among lower risk offenders, harsher prison living conditions, and aggregate data which tend to wildly inflate results in favour of deterrence.¹⁹

To return to the original question as to whether longer periods of incarceration are associated with reductions in recidivism, we examined two sets of data that addressed the above-noted caveats and provide the most exacting assessment of the issue to date. We located 222 comparisons of groups of offenders ($n = 68,248$) who spent more (an average of 30 months) versus less (an average of 17 months) time in prison. The groups were similar on approximately 1 to 5 risk factors. As seen in Table 2, offenders who did more time had slight increases in recidivism of 3% regardless of whether the effect sizes were unweighted (Φ) or weighted (Z^*).

The second sample involved 103 comparisons of 267,804 offenders who were either sent to prison for brief periods (only 18% of effect sizes had length of incarceration noted) or received a community-based sanction. Once again, the results from Table 2 indicate no deterrent effect. Using Φ as a measure of outcome, we see an increase

Table 2

Mean phi (Φ) and mean weighted phi (Z^*) for More vs. Less and Incarceration vs. Community sanctions

Type of Sanction (k)	N	M Φ (SD)	CI Φ	Z*	CI Z*
1. More vs. Less (222) ^a	68,248	.03(.11)	.02 to .05	.03	.02 to .04
2. Incarceration vs. Community (103) ^b	267,804	.07(.12)	.05 to .09	.00	.00 to .00
3. Total (325)	336,052	.04(.12)	.03 to .06	.02	.02 to .02

Note: k = number of effect sizes per type of sanction; N = total sample size per type of sanction; M Φ (SD) = mean phi and standard deviation per type of sanction;

CI Φ = confidence interval about mean phi; Z* = weighted estimation of Φ per type of sanction; CI Z* = confidence about Z*.

^a More vs. Less — mean prison time in months (k = 190): More = 30.0 mths, Less = 12.9 mths, Difference = 17.2 mths.

^b Incarceration vs. Community — mean prison time in months (k = 19): 10.5 mths.

in recidivism of 7% but no effect (0%) when effect size is weighted by sample size.

Clearly, the prison as deterrent hypothesis is not supported. The opposite conclusion, and one that is widely endorsed in some correctional circles, is that prisons do increase recidivism, in other words act as "schools for crime". This is problematic in our view. The studies in this data base are sufficient information to adequately assess this question. Moreover, the design strength of many of the comparison groups leaves much to be desired, albeit we found no correlation between quality of design and effect size (Φ). While this is the "best" available evidence with which to assess the enthusiastic claims of prison deterrence supporters, the only really

satisfactory answer to this particular question is for prison authorities to periodically assess incarcerated on a comprehensive list of dynamic risk factors and correlate time served and changes in risk while incarcerated with future recidivism. This will prove, by far, to be the most sensitive analysis. Regrettably, evaluations of this type have rarely been reported in the corrections literature.

In summary, the addition of this body of evidence to the "what works" debate leads to the inescapable conclusion that, when it comes to reducing individual offender recidivism, the "only game in town" is appropriate cognitive-behavioural treatments which embody known principles of effective intervention.²⁰ ■

1 P.O. Box 5050, Saint John, New Brunswick E2L 4L5.

2 600 Dyer Hall, Cincinnati, Ohio 45221-0389.

3 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6

4 Gendreau, P., Paparozzi, M., Little, T. and Goddard, M. (1993). "Does 'punishing smarter' work? An assessment of the new generation of alternative sanctions in probation". *Forum on Corrections Research*, 5 (3), 31-34.

5 Walter, T. L. and Mills, C. M. (1980). "A behavioral-employment intervention program for reducing juvenile delinquency". In R. R. Ross and P. Gendreau (Eds.), *Effective correctional treatment*, p. 185-206. Toronto, ON: Butterworths.

6 Andrews, D. A. and Kiessling, J. J. (1980) "Program structure and effective correctional practice: A summary of CaVic research". In R. R. Ross and P. Gendreau (Eds.), *Effective correctional treatment*, p. 439-463. Toronto, ON: Butterworths.

7 Andrews, D. A. (1995). "The psychology of criminal conduct and effective treatment". In J. McGuire (Ed.), *What works: Reducing reoffending*, p. 35-62. West Sussex, UK: John Wiley. See also Gendreau, P. and Ross, R. R. (1983-1984). "Correctional treatment: Some recommendations for successful intervention". *Juvenile and Family Court*, 34, p. 31-40.

8 C. H. Blakely, W. S. Davidson, II, C. A. Saylor and M. J. Robinson (1980). "Kentfields rehabilitation program: Ten years later". In R. R. Ross and P. Gendreau (Eds.), *Effective correctional treatment*, Toronto, ON: Butterworths, p. 439-463. See also W. S. Davidson, II and M. J. Robinson (1975). "Community psychology and behavior modification: A community based program for the prevention of delinquency", *Journal of Corrective Psychiatry and Behavior Therapy*, 21, p. 1-12.

9 Martinson, R. (1976). "California and the crossroads". In R. Martinson, T. Palmer and S. Adams (Eds.), *Rehabilitation, recidivism and research*, p. 63-74. Hackensack, NJ: National Council on Crime and Delinquency.

10 Cullen, F. T. and Gendreau, P. (in press). "Assessing correctional rehabilitation: Policy, practice, and prospects". In J. Horney (Ed.), *NIJ criminal justice 2000: Vol. 3. Changes in decision making and discretion in the criminal justice system*. Washington, DC: U.S. Department of Justice, National Institute of Justice.

11 Petersilia, J. (1990). "When probation becomes more dreaded than prison", *Federal Probation*, 54, p. 23-27.

12 Erwin, B. (1986). "Turning up the heat on probationers in Georgia". *Federal Probation*, 50, 17-24.

13 Gendreau, P., Cullen, F. T. and Bonta, J. (1994). "Intensive rehabilitation supervision: The next generation in community corrections?" *Federal Probation*, 58, p. 72-78.

14 Gendreau, P., Goggin, C. and Fulton, B. (2000). "Intensive supervision in probation and parole". In C. R. Hollin (Ed.), *Handbook of offender assessment and treatment* p. 195-204. Chichester, UK: John Wiley.

15 Cullen and Gendreau. (in press).

16 Andenaes, J. (1968). "Does punishment deter crime?" *Criminal Law Quarterly*, 11, p. 76-93.

17 von Hirsch, A., Bottoms, A. E., Burney, E. and Wikström, P.O. (1999). *Criminal deterrence and sentence severity: An analysis of recent research*. Oxford, UK: Hart Publishing.

18 Fabelo, T. (1995). *Testing the case for more incarceration in Texas: The record so far*. State of Texas: Criminal Justice Policy Council.

19 Gendreau, P., Goggin, C. and Cullen, F. (1999). *The effects of prison sentences on recidivism*. Ottawa, ON: Solicitor General Canada.

20 Andrews, D. A., Dowden, C. and Gendreau, P. (1999). *Clinically relevant and psychologically informed approaches to reduced re-offending: A meta-analytic study of human service, risk, need, responsivity, and other concerns in justice contexts*. Unpublished manuscript, Carleton University, Ottawa, ON.

**The Effectiveness of Juvenile Cognitive Behavioral and
Family Oriented Interventions: A Meta-Analysis**

Mirlinda Ndrecka, M.S, Kristin Bechtel, M.S,
Christopher T. Lowenkamp, Ph.D, and Edward J. Latessa, Ph.D
University of Cincinnati

The Effectiveness of Juvenile Cognitive Behavioral and Family Oriented Interventions: A Meta-Analysis

Introduction

On any given day, there are about 2.2 million juveniles being processed by the juvenile justice system. About 100,000 of these individuals will be placed in juvenile institutions, whereas the rest will receive some kind of community sanction (Snyder & Sickmund, 2006). Some criminal justice researchers view this situation as hopeless, concluding that no correctional treatments have any effect on recidivism (Martinson, 1974; Whitehead & Lab, 1989). Nevertheless, many meta-analyses evaluating juvenile treatments have shown that programs that include cognitive behavioral elements, and follow the “principles of effective interventions” in the course of administering treatment are successful in reducing recidivism (Izzo & Ross, 1990; Lipsey, 1995, Andrews et al., 1990; Lipsey & Landenberger, 2003). This body of research, also known as the “what works” approach to offender treatment, has been essential in reshaping the research, and methods employed by criminal justice scholars, since the 1980’s (Cann et al., 2005).

Cognitive behavioral interventions and other offender treatment programs that employ cognitive behavioral elements, such as family focused interventions have been shown to be effective in reducing recidivism with a variety of juvenile offender populations (Latessa, 2006; Van Voorhis & Lester, 2004). While cognitive behavioral programs focus on changing the anti-social aspects of individuals, family oriented interventions attempt to change these aspects in connection with the family system, and also include other environmental factors that influence deviant behavior. They see the individual as part of a variety of systems that intermingle with each other to shape behavior (Henggeler et al. 1986).

This chapter provides a succinct description of family based and cognitive behavioral interventions used with juvenile offenders. Longitudinal studies of juvenile offenders have shown that serious juvenile delinquents are at the greatest risk for committing additional offenses in the future (Weisz et al., 1991, Lewis et al., 1989). As such, their treatment and rehabilitation is imperative in preventing future criminal activity (Borduin et al., 1995). Several meta-analyses have demonstrated that treatment programs that are based on behavioral strategies such as radical behavioral, social learning and cognitive-behavioral approaches are the most successful in reducing recidivism. The present chapter will examine the effect of these types of programs on juvenile offender populations. The next section focuses on describing the philosophies that fueled these types of interventions and the research results obtained from their evaluations.

Cognitive Behavioral Interventions

Cognitive behavioral interventions attempt to change an individual's attitudes and thinking processes. Thoughts, feelings, and attitudes toward certain events influence the way individuals experience those events, and in turn, the way they react toward subsequent occurrences (Barriga et al., 2000; Van Voorhis & Lester, 2004). For example, two people could view an incident, such as being stopped by a law enforcement officer for speeding, in different ways. One could believe that the officer was "out to get them" and that he was unlucky in being caught, since everyone speeds. This individual would, then, get angry and believe that the police are misusing their resources. On the other hand, another person could accept responsibility for his behavior and believe that the officer was doing her job having a duty to protect the safety of roads by sanctioning unsafe drivers. The first instance is an illustration of anti-social thinking, while the second depicts a more pro-social attitude. By helping individuals change their

thoughts and attitudes, cognitive behavioral programs promote the development of pro-social outlooks, and attempt to increase skills to display pro-social behavior (Latessa, 2006).

Cognitive behavioral interventions are classified under two models: cognitive restructuring and cognitive skills. Cognitive restructuring interventions seek to modify the core beliefs and attitudes of an individual. They attempt to change one's thoughts. Thus, the interventions target irrational thoughts, or what Yochelson and Samenow (1976) identify as, thinking errors. Thinking errors represent anti-social thoughts, and it is by maintaining these faulty thoughts and beliefs that individuals minimize or rationalize anti-social and criminal behavior (Barriga et al., 2000; Latessa, 2006).

On the other hand, cognitive skills interventions try to increase the ability of an individual to develop pro-social thoughts, and consequently, exhibit pro-social behavior (Latessa, 2006; Van Voorhis & Lester, 2004). They seek to improve the ability of individuals to respond pro-socially to a variety of situations. The focus is on building pro-social problem-solving skills, to control and stop anti-social thinking patterns, and replace them with new pro-social ones. For example, as part of the intervention, individuals are taught techniques to control their impulsivity and coached on new skills to improve their self-control (Van Voorhis & Lester, 2004).

Both types of programs employ similar methods to achieve change in individuals. They use techniques such as modeling, during which the program facilitator demonstrates the desired behaviors for the program participants, and role-play illustration skits, which allow the participants to practice the new desired behaviors. Other techniques include reasoning exercises, group discussions, and games that encourage the reinforcement of pro-social attitudes and behaviors (Van Voorhis & Lester, 2004).

Recent programs incorporate both of these models in their curricula, attempting to influence change in their anti-social beliefs, while simultaneously increasing pro-social skills and promoting pro-social behavior. Thinking for a Change (Bush, Glick, & Taymans, 1998) is one such curriculum that targets thinking errors and anti-social beliefs while introducing new pro-social thoughts to replace the old, irrational ones. The curriculum helps participants to identify situations that present a risk for the youth to engage in anti-social and/or criminal behavior. Next, skills to avoid participating in such behavior and to respond pro-socially to those risky situations are taught and reinforced through some of the techniques previously mentioned (Latessa, 2006).

Some cognitive behavioral programs include elements of moral development and moral education in their curricula. According to Kohlberg's and Piaget's moral development theories, individuals advance through stages of moral judgment as they collect attitudes, beliefs and thoughts about their environment and their life events (Armstrong, 2003). Belonging to a certain moral development stage will dictate the behavior of an individual and the way they justify their actions. The higher one moves on the stages of moral development, the more sophisticated the thought processes become, and as a result, the more selfless and altruistic are the behaviors. Through reasoning exercises and changing the structure of beliefs, cognitive behavioral interventions attempt to move individuals upward on the moral stages continuum (Armstrong, 2003; Van Voorhis & Lester, 2004).

Following these views, programs such as EQUIP (Gibbs et al, 1995), Moral Reconciliation Therapy (MRT) (Little & Robinson, 1986) and Aggression Replacement Training (ART) (Goldstein & Glick, 1987) combine elements of cognitive behavioral interventions and moral development education (Leeman et al, 1993; Van Voorhis & Lester, 2004; Goldstein & Glick, 1987). By exposing participants to different moral dilemmas and utilizing cognitive behavioral

exercises to develop pro-social responses to these dilemmas, the curricula enable the participants to move to higher stages of moral development. Higher levels of moral reasoning correspond with the development of more pro-social skills and result in an increase in avoidance of anti-social behaviors, and pro-social behavior gains. Therefore, as individuals proceed through the treatment, they are expected to exhibit behaviors that are consistent with more advanced levels of moral judgment (Armstrong, 2003; Goldstein & Glick, 1987; Latessa, 2006; Leeman et al., 1993).

In the recent years, cognitive behavioral interventions have gained much popularity in their use with offender populations. Several characteristics of cognitive behavioral programs make them attractive for correctional agencies (Van Voorhis & Lester, 2004). These include:

- Cognitive behavioral programs are inexpensive to implement in both residential and community settings;
- Most staff can be trained to facilitate cognitive behavioral interventions and do not need to be licensed or certified clinicians;
- Cognitive behavioral interventions are applicable in both individual and group settings;
- Cognitive behavioral programs are flexible in their treatment delivery thus accommodating variations in individual's responsivity to interventions.

There are also a plethora of research studies that attest to the efficacy of cognitive behavioral interventions and the methods upon which they are based. For example, a meta-analysis of 45 treatments by Andrews et al. (1990) found that treatments that adhere to the principles of effective intervention: risk, need and responsivity, are more successful in reducing recidivism.

In addition, the programs that are designed to strictly follow structured manuals, allowing less room for deviations, and promoting treatment fidelity (Van Voorhis & Lester, 2004) are more effective when implemented. In a meta-analysis of juvenile treatments, Lipsey (1999) found that adherence to the program curriculum and design results in treatments that are more effective in reducing recidivism. Their simplicity in design and in delivery contributes to the popularity they have earned with correctional systems.

Furthermore, the very aim of cognitive behavioral treatments makes them desirable for use with offender populations. The treatment modalities attempt to change the thinking and the behavior of offenders by focusing in the present, without dwelling in past fears and/or experiences. They also teach the offenders how to respond to their current problem situations in a pro-social manner, and thus, adjust to their surroundings more efficiently. Offenders can apply the pro-social skills they have learned and rehearsed during treatment to their daily circumstances (Latessa, 2006).

Finally, many meta-analyses of offender programs have shown that cognitive behavioral interventions are very effective in reducing recidivism. A recent meta-analysis by Pearson et al. (2002) found that cognitive behavioral interventions reduced recidivism in juvenile and adult offenders placed in either institutional or community settings. Lipsey and Landenberger (2003) also found that treatment modalities that target anti-social attitudes and cognitions are more effective than treatments that do not address these issues.

Meta-analyses that have examined programs applied to juveniles specifically, have found cognitive behavioral treatments to be the most effective. Izzo and Ross' (1990) meta-analysis of juvenile programs concluded that including cognitive behavioral elements in the treatment of juveniles increases their effectiveness in reducing recidivism. After examining about 400 studies

of juvenile programs, Lipsey (1995) found that cognitive behavioral programs are amongst the most effective treatments. The efficiency of treatments also increases when cognitive behavioral interventions follow the principles of effective interventions (Andrews et al, 1990; Pealer & Latessa, 2004).

Family based interventions

The philosophy behind family based interventions maintains that the behavior of a family member is influenced by, and in turn, affects the activities of the whole family unit. Family based treatments seek to change the anti-social behavior of the delinquent individual, and also the unhealthy communication patterns that exist among family members. Improvement of the anti-social conduct and expression of wanted behavior is important in many contexts, such as school, employment and peers (Alexander & Parsons, 1982; Henggeler et al., 1986).

Family based interventions have adopted many modalities and treatment concepts throughout their development, such as the psychodynamic and the communications models (Van Voorhis & Lester, 2004). The next section focuses in examining the structure and research of but two currently popular family therapy treatment modalities: Functional Family Therapy (FFT) and Multi-Systemic Therapy (MST). Both types of family interventions include multiple treatment methods such as improving communication skills among family members using family therapy, and increasing pro-social contacts and pro-social problem-solving abilities using cognitive behavioral techniques (Alexander & Parsons, 1982; Henggeler et al., 1986). Their multi-method approach becomes important in the context of offender treatment, seeing as Izzo and Ross (1990) found improvement of effectiveness in programs that included cognitive behavioral elements during the course of treatment.

Functional Family Therapy

Functional Family Therapy (earlier version also called behavioral-systems therapy) is based on the idea that differences exist in the communication patterns among families of delinquents and non-delinquents. Usually, juvenile delinquents also belong to families that have dysfunctional communication patterns. Functional family therapy attempts to change these communication patterns to resemble healthy ones, through a series of behavioral techniques (i.e., contingency contracts) (Alexander & Parsons, 1973; Barton et al., 1985; Gordon et al., 1988).

Treatment sessions consist of family therapy during which the facilitators seek to modify verbal communications among family members by modeling healthy communication examples, role-playing, and reinforcement of wanted relationship patterns. Therapists seek to minimize and extinguish maladaptive coping skills such as blaming and punishing, by replacing them with clear and direct verbal communication of feelings. Contingency contracts among family members and reinforcement of positive communication are used to promote alternative solutions to the damaging interactions that exist (Alexander et al., 1976).

Research on the use of Functional Family Therapy has shown that the intervention is effective at improving healthy communication patterns in dysfunctional families, and reduction in delinquent behavior. Alexander and Parsons (1973) found that FFT was more effective than client-centered therapy at improving family communication. More importantly, delinquents whose families participated in the treatment showed significant reduction in delinquency rates. In addition, a long-term follow-up of the siblings of the initial delinquents that participated in the original study, found lower recidivism rates for the families that received Functional Family Therapy (Klein et al., 1977).

Replications of the 1973 Alexander and Parsons' study, conducted in different settings, also found reduced rates of recidivism for families that received Functional Family Therapy .

The rates in decline of recidivism were similar to the findings of the original study, although the replications had a few limitations (Barton et al., 1985). A more methodologically sound study conducted by Gordon and colleagues (1988) found yet a more significant reduction in recidivism for FFT participants. The authors speculated that larger reductions in recidivism resulted because of improvements in treatment delivery and longer treatment periods (Gordon et al., 1988).

Multi-Systemic Therapy

Multi-Systemic Therapy (MST) is based on the family ecological systems view that maintains individual behavior is influenced by multiple factors and multiple systems. Family relationships and communications is one of the systems that influence an individual's behavior, but there are other extra familial systems such as peers and school that also affect it. Thus, behavioral problems and changes in an individual cannot be seen only as a function of family communications. Change in behavior is achieved through change in the many systems that affect it, and how they are connected with each other (Henggeler et al., 1986).

Multi-Systemic Therapy uses treatment strategies borrowed from family therapy and behavioral therapy. As such, MST focuses on improving problem-solving skills and cognitive processes that contribute to an individual's antisocial behavior. Cognitive behavioral techniques are routinely employed in the treatment sessions to achieve behavioral change (Henggeler et al., 1992). In addition, MST attempts to improve family processes and communication, decrease associations with anti-social peers, and improve school performance. Family preservation remains one of the most important goals of Multi-Systemic Therapy; therefore, all services are developed to work with the family members and their connections to improve the problem situations (Henggeler et al., 1986).

Therapy sessions vary from family to family, although they are always conducted in the community and preferably in the family's home. Therapists are available for therapy sessions and are accessible 24 hours a day for the duration of the therapy. Family members can also initiate therapy sessions if they need the help of the therapist to resolve certain issues (Henggeler et al., 1986; Henggeler et al., 1992).

Multi-Systemic Therapy has demonstrated significant results in improving behavior and reducing delinquency rates with serious, juvenile offender populations (Henggeler et al, 1997; Henggeler et al., 1993). Significant improvements were also seen in drug use rates among drug abuse delinquent populations that underwent Multi-Systemic Therapy (Henggeler et al., 1997; Henggeler et al., 2002; Henggeler et al., 2006).

Additionally, MST has demonstrated to be a valuable remedy in the treatment of juvenile sex-offenders. After a three-year follow-up period, fewer offenders who had received Multi-Systemic Therapy were arrested for sexual crimes, and at a lower frequency (Borduin et al., 1990). Finally, the effectiveness of Multi-Systemic Therapy is robust even across cultures, as MST interventions have been replicated with juvenile populations outside the United States. In a replication of MST done with behavioral problem juveniles in Norway, the intervention was more successful than usually administered services in reducing behavioral problems, and as a result, reduced the number of out of home placements for the treatment group participants (Ogden & Halliday-Boykins, 2004; Ogden & Hagen, 2006).

Current Study

The current study provides a meta-analytic review of juvenile cognitive behavioral and family oriented treatments. As previously mentioned, many of these programs have demonstrated significant effectiveness with juvenile populations. The current study seeks to

summarize the overall magnitude of the effect that these programs have had in reducing recidivism and improving problem behavior.

Methodology

In an effort to identify potentially eligible studies, it is necessary to conduct an exhaustive review of the literature. To meet this objective, there were two primary methods employed in the literature retrieval process. First, there was an extensive search of electronic databases, including the Criminal Justice Periodical Index, PsychInfo, the Social Sciences Citation Index, Dissertation Abstracts, the National Criminal Justice Reference Service, multiple state and local government websites including the Office of Juvenile Justice and Delinquency Prevention website, and conference proceedings and unpublished papers presented at the Academy of Criminal Justice Sciences and the American Society of Criminology. Second, the ancestry method was utilized, where references cited from previously gathered studies were then tracked down for possible inclusion into the meta-analysis.

Through this process, a total of 56 studies were identified for possible inclusion into the analysis. The final study sample size was reduced to 34 studies. This is primarily due to 22 of the 56 studies being rejected for failing to meet the eligibility criteria. Specifically, these requirements were (1) the sample was to be comprised of offenders, (2) the original evaluation needed both a treatment and comparison group, (3) the outcome variable(s) had to be a measure of recidivism and (4) the findings from the original program evaluation needed to include the necessary data to compute an effect size. Measures of recidivism included technical violations,

re-arrests, re-convictions, re-incarceration, a combination of these measures of recidivism, or not reported.¹

Upon completion of coding, individual effect sizes were calculated for each study and the coding data for each eligible study were entered into a database. Next, the overall effect size was calculated. Confidence intervals were examined to determine the impact of juvenile cognitive behavioral and family oriented treatments on reducing recidivism. Finally, common language effect sizes were calculated which expresses the probability that an effect size sampled from one group will be greater than an effect size sampled from the other group (Dunlap, 1994; McGraw & Wong, 1992).

Description of sample

From these 34 studies, 27% of the sample was comprised of male-only programs and 3% were female-only. Mixing of males and females occurred in 53% of the programs. Finally, 18% of the studies did not provide these data. The majority of programs included multiple races. In particular, 47% of the 34 programs were mixed, 12% were comprised solely of White juveniles and 3% included only Black youth. In addition, 35% of the eligible studies did not present data on the racial composition of their sample. With respect to the age group in these programs, 91% were identified as juvenile only, whereas 9% of the studies suggested that the age of their sample was mixed. A majority of these programs did not include youth with mental disorders (56%), but there were 12% of the groups that did include these youth and 15% of the programs that mixed both of these groups. Further, 18% of the programs in the sample did not provide any

¹ Mixed measures of recidivism are identified as 'mixed' in the tables. While the outcome measure was labeled as recidivism in the studies, those that did not specify the exact type of recidivism examined were labeled as 'not reported' in the tables.

mental health data. Almost 40% of the sample was comprised of high-risk youth, and 12% of the 34 studies included low and moderate risk samples (9% and 3% respectively)². Twenty-four percent of the sample was comprised of mixed risk offenders and the remaining 27% of the eligible studies did not indicate the risk level (see Table A1 in Appendix A).

Concerning the offending history of the youth in these studies, 12% of the eligible studies had youth that engaged in violence, while 18% of the sample did not. Almost 30% of the programs in the overall sample mixed youth both with and without a history of violence. Finally, 41% of studies did not include a history of violence measure. Pertaining to programs with juvenile sex offenders, 3% of the programs in the overall sample included sex offenders only and 15% of the studies suggested that the programs had mixed sex offenders with non-sex offenders. Further, 27% of the programs did not have youth with a history of sex offending and 55% did not report the sex offense history of the participants in their program. Six percent of the sample contained studies that evaluated youth with a current violent offense. There were 65% of individual studies that did not specify the current offense of their sample and 15% of the overall sample included both mixed and non-violent offenders each (see Table A1 in the Appendix A).

Description of treatment

While 18% of the programs in the sample were in an institutional setting, over 75% of the programs were conducted in the community. Six percent of these studies identified that the location of the program was in a residential/group home. Concerning the primary format of treatment, 53% involved the youth and family and 24% were group sessions. In the overall

² It should be noted that criminal justice treatment providers did not report having low risk samples in their programs. Further, low risk offenders were not found in programs that lasted over 12 weeks or where the group intensity was identified as 41 hours or more.

sample, 15% of the studies included other formats of treatment and 9% that did not specify the treatment format. Regarding duration of treatment, approximately 60% of the groups lasted 12 or more weeks and 17% lasted between 1-12 weeks. Almost one-quarter of the studies did not specify the duration of treatment. Over 30% of the programs were privately managed and 65% of the treatment programs were criminal justice providers. Only 1 study did not specify the provider of the treatment program. Both cognitive behavioral and multisystemic programming were separately reported in 32% of the studies. Functional family therapy was identified in 27% of the eligible studies. One study was coded as a mixed treatment type and the remaining two included studies were identified as other programming (see Table A2 in Appendix A).

Outcome measures description

Re-arrest was identified in 62% of the studies as the primary outcome measure. Less than 10% of the evaluations examined technical violations, re-convictions, or re-incarceration as the measure for failure. Almost 20% of the studies examined a combination of recidivism measures and one study did not report the specific measure of recidivism. Length of follow-up was measured as less than two years in over 70% of the studies in comparison to nearly 30% that were over 2 years (see Table A3 in Appendix A).

To summarize, over half of the programs were mixed by sex and race and primarily contained a juvenile sample.³ A majority of the eligible studies comprised high-risk youth; however, almost 25% of the programs did not separate their groups by risk level. A majority of the eligible studies were located in the community and followed a youth and family treatment

³ It should be noted that with Functional Family Therapy and the Multi-Systemic Therapy models, that the sex of the samples would be mixed in these programs.

format. Regarding treatment models, the overall sample was almost divided into thirds between evaluations of cognitive behavioral programming Multi-Systemic Therapy and Functional Family Therapy. A majority of these programs were from independent providers and the length of programming generally lasted 12 or more weeks. For recidivism, these studies employed re-arrest as the primary measure and the follow-up length was primarily 2 years or less. Before a presentation of the findings, the interpretation of the overall effect size, confidence intervals and Q- statistic will be provided.

Interpreting the effect size, confidence intervals and Q-statistic

In order to determine if the included studies in this meta-analysis were successful in reducing recidivism, an overall effect size and confidence intervals were calculated. Interpretation of the effect size provides an indication of the overall impact that these programs have on recidivism. Regarding the confidence intervals, this range suggests if the effect size should be cautiously considered or if there should be confidence in the effect size value. Further, confidence intervals can be treated as a test of significance for the effect size. Specific examples on the interpretation of the effect size, confidence intervals and Q-statistic follows.

Interpreting the effect size:

- ❑ For example, an effect size of .30 would indicate that the treatment **decreased** recidivism by 30%.
- ❑ Similarly, an effect size of -.30 would indicate that the treatment **increased** recidivism by 30%.
- ❑ Note how the change in the sign before the effect size indicates the direction in which recidivism was affected (Lowenkamp, Smith & Bechtel, 2007).

Interpreting the confidence intervals:

- ❑ The smaller the range ($<.10$) between the upper and lower limits of the confidence interval suggests that there should be more confidence in the effect size value.
- ❑ The larger the range ($>.10$) between the upper and lower limits suggests that the effect size value should be interpreted cautiously.
- ❑ As mentioned previously, the confidence interval can be treated as a test of significance and would indicate if the effect size is significantly correlated with the outcome measure. In particular, if the range includes 0 then it would **not** be considered significant (Lowenkamp, Smith & Bechtel, 2007).

Interpreting the Q-statistic

- ❑ The Q-statistic is used as a test of homogeneity, which would suggest that the effect size from the individual studies would be equal. It is interpreted as a chi-square distribution. This value can suggest if there is any variation in the individual effect sizes which would indicate if there was a significant difference between the studies.

While the r value is provided as the mean effect size, an additional effect size indicator, the z_r has been calculated and presented in the findings. This value is calculated using the Fisher's r to Z transformation and then averaged to the mean Fisher r , which is z_r . Rosenthal (1991), however, does caution that the more conservative estimate is the mean r . It is anticipated that the comparison between the two effect size indicators will be similar. Finally, the Z statistic is also presented in the findings to examine if collectively, the studies produced a statistically significant effect on recidivism.

Findings

Before examining the effect size by the various measures such as treatment model, risk level, length of treatment, intensity of treatment, and treatment provider, a table that presents the mean effect size for the fixed effects and random effects models is presented. A fixed effects model only examines within-study variability, whereas a random effects model considers both the within-study variability and the between-study variability. Further, the assumption of the random effects model can be interpreted as though the included studies are a random sample of all possible studies examining the effectiveness of juvenile treatment. As such, the random effects model is considered a more conservative estimate, which is generally why the confidence interval is wider with this model in comparison to the fixed effects model.

Table B1 presents the mean effect size when calculated for the fixed effects and random effects models. As depicted, there are 34 studies included in the meta-analysis, with a combined total of 7,188 offenders. Both the fixed effects and random effects models produced a significant mean effect size, $r=.10$ and $r=.19$ respectively. Given that the random effects model is significant, Table 2 which examined the mean effect size by type of treatment was calculated using the random effects model.

Table B1. Mean Effect Size

Model	k	N	r	z_r	95% CI	Z (p)
Fixed Effects	34	7188	.10	.10	.08 to .12	8.43 (.0000)
Random Effects	34	7188	.19	.19	.13 to .24	6.57 (.0000)

Table B2 examines the mean effect size of cognitive behavioral, Functional Family, mixed treatment, Multi-Systemic Therapy (MST) and other juvenile programming aimed at

reducing recidivism. In this analysis, there were 11 cognitive behavioral programs identified with a total sample size of 3,817 offenders that produced a mean effect size of $r=.10$. The confidence interval for this effect size was slightly wide, but nevertheless, did reveal that the cognitive behavioral programs were significantly correlated with reductions in recidivism. Nine Functional Family Therapy program evaluations were eligible for inclusion in this study. There was a combined total of 1,157 offenders from these 9 studies and the mean effect size was $r= .27$ which was significant. Similarly, the confidence interval was slightly wide for the Functional Family Therapy programs. Only one study in the sample was a mixed treatment model. This evaluation included 149 offenders and produced a mean effect size of $r= .11$, which was not significant. Eleven Multi-Systemic Therapy evaluations were included in this analysis and there were a total of 1,072 offenders. A significant mean effect size of $r= .24$ was found and the confidence interval was slightly wide. Finally, there were two studies coded as other treatment models that included 993 offenders. These programs produced a very small effect size, $r=.05$, which was not significant. The Q-statistic (11.19, $df= 4$) indicates that there was a significant difference in treatment type as observed in the random effects model. When interpreting the significant mean effect sizes by treatment type, cognitive behavioral models decreased recidivism by 10%, Functional Family Therapy models decreased recidivism by 27% and Multi-Systemic Therapy were found to decrease recidivism by 24%. Further support for these findings is noted upon examination of the Z statistic. Multi-Systemic Therapy and Functional Family programming produced significant results as well as cognitive behavioral; however, the mixed and other models failed to reach statistical significance.

Table B2. Mean Effect Size by Type of Treatment Random Effects

Model	k	N	R	z_r	95% CI	Z (p)
CBT	11	3817	.10	.10	.01 to .19	2.06 (.039)
Family	9	1157	.27	.27	.17 to .37	5.22 (.000)
Mixed	1	149	.11	.11	-.15 to .36	0.82 (.415)
MST	11	1072	.24	.25	.16 to .34	5.49 (.000)
Other	2	993	.05	.05	-.12 to .22	0.52 (.602)

Q Between (4) = 11.19; $p = 0.02$

Table B3 presents the mean effect size by risk category. Recall, not all eligible studies reported the risk level of the offenders in their treatment program. Fourteen studies evaluated the effect of treatment on recidivism for 1,314 high-risk offenders. A significant mean effect size for the high-risk offenders was produced, $r=.23$. The confidence interval for this effect size was slightly wide. Regarding the 3 low risk juvenile studies, there was a combined total of 147 youth included in these samples. The range for this confidence interval was wide.⁴ There were 8 mixed risk groups identified with 497 total offenders. A much smaller mean effect size of $r= .10$ was calculated for the mixed risk group and a wide confidence interval was estimated. Interpretation of the Q statistic suggests that there was modest variation between the risk levels. Based on the results in Table 3, programs that targeted higher risk offenders decreased recidivism by 23%. Similarly, programs that targeted lower risk offenders were found to have significantly decreased recidivism by 22%.

⁴ Small sample sizes can often produce wide confidence intervals, which indicate that the findings should be interpreted with caution since the mean effect size may be relatively unstable.

Table B3. Mean Effect Size by Risk

Model	k	n	R	z_r	95% CI	Z (p)
Higher	14	1314	.23	.23	.16 to .31	6.00 (.000)
Low	3	147	.22	.23	.02 to .44	2.12 (.034)
Mixed	8	497	.10	.10	.01 to .19	2.15 (.031)

Q Between (2) = 5.14; $p = 0.07$

Table B4 demonstrates the mean effect sizes for the 25 studies that identified the length of treatment. Among the 6 programs where treatment lasted between 1 and 11 weeks, there were 504 offenders. A significant mean effect size of $r = .17$ was noted along with a slightly wide confidence interval. For the 19 studies with programs that operated for 12 or more weeks, there were a total of 2,491 offenders. A significant mean effect size of .15 was produced. Further, the confidence interval was slightly wide. A test of homogeneity revealed that there was little difference between the models based on length of treatment given the insignificant value of the Q statistic = .113.

Table B4. Mean Effect Size by Length of TX

Model	k	n	R	z_r	95% CI	Z (p)
1-11 weeks	6	504	.17	.18	.06 to .29	2.93 (.003)
12+ weeks	19	2941	.15	.15	.08 to .22	4.38 (.000)

Q Between (1) = .113; $p = 0.74$

Table B5 presents the mean effect sizes by intensity of treatment. There were a total of 24 studies that identified the intensity of treatment. Specifically, 7 evaluations reported that the treatment intensity lasted between 1-20 hours. There were 280 offenders within these 7 studies. A significant mean effect size of $r = .22$ was found. Regarding treatment intensity lasting 21-40 hours, there were 857 offenders in these 10 studies. A significant mean effect size of $r = .28$ was

produced. Finally, for intensity of treatment lasting 41 or more hours, there were 7 eligible studies with a total of 4,364 youth. A much smaller, but statistically significant mean effect size of $r = .12$ was revealed. The width of these three confidence intervals was similar. Evaluation of the Q statistic suggests that there was modest variation between models based on intensity of treatment. Overall, intensity of treatment between 21-40 hours was found to have significantly reduced recidivism 28% followed by treatment intensity lasting 1-20 hours significantly reducing recidivism 22%. Program intensity over 41 hours significantly reduced recidivism only 12%.

Table B5. Mean Effect Size by Intensity of TX

Model	k	N	R	z_r	95% CI	Z (p)
1-20 hours	7	280	.22	.22	.12 to .33	4.02 (.000)
21-40 hours	10	857	.28	.29	.19 to .39	5.71 (.000)
41+ hours	7	4364	.12	.12	.02 to .22	2.36 (.000)

Q Between (2) = 5.75; $p = 0.06$

Table B6 examines the mean effect size by type of treatment provider. In particular, there were 11 treatment programs that were operated by a criminal justice provider. Collectively there were 5,224 offenders who participated in these studies. A significant mean effect size of $r = .13$ was revealed with a slightly wide confidence interval. In comparison, 22 non-criminal justice treatment providers were noted for this study which included a total of 1,937 juveniles. These programs resulted in a significant mean effect size of $r = .22$. Further, the range of the confidence interval is rather narrow suggesting that this mean effect size can be confidently interpreted. To summarize, these findings indicate that the criminal justice providers significantly reduced recidivism by 13%, whereas the non-criminal justice treatment programs were able to significantly reduce recidivism 22%.

Table B6. Mean Effect Size by Provider of TX

Model	k	N	R	z_r	95% CI	Z (p)
CJ Provider	11	5224	.13	.13	.05 to .21	3.07 (.002)
Non CJ-Provider	22	1937	.22	.22	.15 to .22	6.61 (.000)

Q Between (1) = 2.94; $p = 0.09$

Table B7 presents the common language (CL) effect sizes. As previously mentioned, this statistic indicates the probability of a randomly sampled effect size from one type of treatment group would be greater than a randomly sampled effect size from another treatment group (Dunlap, 1994; McGraw & Wong, 1992).⁵ For example, when examining the first row which includes the cognitive behavioral programs, a randomly sampled CBT effect size would be larger than a MST effect size, 95% of the time. Further, a randomly sampled effect size from the CBT programs would be larger than a family group effect size 83% of the time. Regarding the other and mixed group types of treatment, a randomly selected effect size from the CBT programs would be larger than the other and mixed treatments, 72% and 77% respectively.

Table B7. Common Language Effect Size by Type of Treatment

Model	CBT- <i>CL</i>	MST- <i>CL</i>	Family- <i>CL</i>	Other- <i>CL</i>	Mixed- <i>CL</i>
CBT	--	.95	.83	.72	.77
MST	.05	--	.42	.16	.99
Family	.17	.58	--	.34	.10
Other	.28	.84	.66	--	.77
Mixed	.77	.99	.90	.23	--

⁵ Given that there are a total of 34 studies included in the meta-analysis, the CL effect size allows us to state that while the overall sample size may be low, we can say that the r value has some level of strength to it.

Discussion

Overall, the cognitive behavioral, Functional Family Therapy and Multi-Systemic Therapy approaches produced significant mean effect sizes in a conservative random effects model. These findings suggest that the three treatment models are successful in reducing recidivism rates for youth. The mean effect sizes and Z statistics produced by the Functional Family Therapy and Multi-Systemic Therapy programs were slightly larger than that of the cognitive programs. While the confidence intervals for these treatment modalities were slightly wide, this is not unusual when the number of samples in each category was rather small.

With regard to risk level, there are two important points to stress regarding how low risk cases were managed. First, low risk offenders were kept out of the programs that were operated by criminal justice providers. Second, low risk offenders were not found in programs that lasted over 12 weeks or where the group intensity was identified as 41 hours or more. When examining the mean effect sizes, programs that properly identified their high and low risk populations produced a significant effect in reducing recidivism. This is particularly relevant when examining the programming characteristics by risk level. Specifically, the mixed risk groups did worse than the high risk samples when looking at criminal justice provider programming. This may be attributed to having included lower risk youth in a criminal justice setting with higher risk offenders. Further, mixed risk groups performed poorly when compared to the high risk when treatment intensity was 41 or more hours and the programming lasted 12 or more weeks. Similarly, this may be due to exposing lower risk youth to a more intensive and longer in duration treatment program.

It would appear that while many of the programs adhered to the risk principle by identifying the appropriate group to target for intervention, programs that did not properly assess

offenders and identify their risk level based on an actuarial assessment were more likely to experience failures for their mixed risk group. Given that Functional Family Therapy, Multi-Systemic Therapy and cognitive behavioral programming is argued to follow evidence based practices, groups that combined low and high risk offender performed the worst. Perhaps this could be attributed to the inadvertent effects of mixing the low and high risk groups and exposing the lower risk offender to the antisocial behaviors of the high risk juveniles. Further, if these antisocial attitudes and behaviors are reinforced, this may perpetuate the problems not only for the low risk but also for the higher risk group (Dishion, McCord, & Poulin, 1999).

GENERAL COMMENT:

While this study provides support for Functional Family Therapy, Multi-Systemic Therapy and cognitive behavioral models, it is necessary to discuss these findings as it relates to program implementation and evaluation. First, programs which adhered to their respective models did reduce recidivism for juvenile offenders. Recall from Table B2, that the mixed and other treatment models did not significantly reduce recidivism, whereas these findings suggest that the Functional Family Therapy, Multi-Systemic and cognitive behavioral programs did. As such, implementation of the three successful models needs to be done with great fidelity. Further, evaluation of these programs should be considerate of programs which report to be following FFT, MST or cognitive behavioral programming, but are not properly implementing such models. Second, new programs should be assessing youthful offenders with a validated risk and needs assessment tool. Based on these results, the low risk group which was diverted from criminal justice programming or received a lower intensity and dosage of treatment were successful in reducing recidivism. Further, proper identification by programs of the high risk group resulted in significant reductions in recidivism. In comparison, programs which mixed the

risk level of the group were not shown to be as successful in reducing recidivism. New programs and those being evaluated need to consider the relevance for identifying the risk level of their targeted population through validated actuarial risk assessment. As gleaned from these findings, program implementation and evaluation should be addressing three main areas: (1) adherence to effective treatment models for juveniles, (2) identifying if the programs are properly identifying the risk level of the juveniles and making appropriate case management decisions, such as avoiding the mixing of the risk levels and (3) by following the risk principle in providing the most intense and highest dosage of services to the high risk youth and diverting or reducing the intensity and dosage of treatment provided to the low risk juvenile.

References

- Alexander, J. F., Barton, C., Schiavo, R. S., & Parsons, B. V. (1976). Systems-behavioral intervention with families of delinquents: Therapist characteristics, family behavior, and outcome. *Journal of Consulting and Clinical Psychology, 44*, 656-664.
- Alexander, J. F., & Parsons, B. V. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81*, 219-225.
- Alexander, J. C., & Parsons, B. V. (1982). *Functional Family Therapy*. Belmont, CA: Brooks/Cole.
- Andrews, D., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19-52.
- Armstrong, T. A. (2003). The effect of Moral Reconciliation Therapy on the recidivism of youthful offenders: A randomized experiment. *Criminal Justice and Behavior, 30*, 668-687.
- Barriga, A., Landau, J., Stinson, B., Liau, A., & Gibbs, J. (2000). Cognitive distortion and problem behaviors in adolescents. *Criminal Justice and Behavior, 27*, 36-56.
- Barton, C., Alexander, J. F., Waldron, H., Turner, C. W., Warburton, J. (1985). Generalizing treatment effects of functional family therapy: Three replications. *The American Journal of Family Therapy, 13*, 16-26.
- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 34*, 105-113.
- Bush, J., Glick, B., & Taymans, J. (1996). *Thinking for a change: An integrated cognitive behavior change program*. Longmont, CO: National Institute of Corrections.
- Cann, J., Falshaw, L., & Friendship, C. (2005). Understanding 'What Works': Accredited cognitive skills programmes for young offenders. *Youth Justice, 5*, 165-179.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm--Peer groups and problem behavior. *American Psychologist, 54*, 755-764.
- Dunlap, W. (1994). Generalizing the common language effect size indicator to bivariate normal correlations. *Psychological Bulletin, 116*, 509-511.
- Gibbs, J., Potter, G., & Goldstein, A. (1995). *The EQUIP program: Teaching youth to think and act responsibly through a peer-helping approach*. Champaign, IL: Research Press.
- Goldstein, A., & Glick, B. (1987). Aggression Replacement Training. *Journal of Counseling and Development, 65*, 356-362.

- Gordon, D. A., Arbuthnot, J., Gustafson, K. E., & McGreen, P. (1988). Home-based behavioral systems family therapy with disadvantaged juvenile delinquents. *American Journal of Family Therapy, 16*, 243-255.
- Henggeler, S. W., Brondino, M. J., Melton, G. B., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy 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., Clingempeel, W., Brondino, J. M., & 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*, 868-875.
- Henggeler, S. W., Halliday-Boykins, C. A., Cunningham, P. B., Randall, J., Shapiro, S. B., & Chapman, J. E. (2006). Juvenile drug court: Enhancing outcomes by integrating evidence-based treatments. *Journal of Consulting and Clinical Psychology, 74*, 42-54.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology, 60*, 953-961.
- Henggeler, S. W., Melton, G. B., Smith, L. A., Schoenwald, S. K., & Hanley, J. H. (1993). Family preservation using multisystemic treatment: Long-term follow-up to a clinical trial with serious juvenile offenders. *Journal of Child and Family Studies, 2*, 283-293.
- Henggeler, S. W., Rodick, J. D., Hanson, C. L., Watson, S. M., Borduin, C. M., & Urey, J. R. (1986). Multisystemic treatment of juvenile offenders: Effects on adolescent behavior and family interaction. *Developmental Psychology, 22*, 132-141.
- Izzo, R., & Ross, R. (1990). Meta-analysis of rehabilitation programs for juvenile delinquents. A brief report. *Criminal Justice and Behavior, 17*, 134-142.
- Klein, N. C., Alexander, J. F., & Parsons, B. V. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology, 45*, 469-474.
- Latessa, E. J. (2006). Effectiveness of cognitive behavioral interventions for youthful offenders – Review of the research. In B. Glick (Ed.), *Cognitive behavioral interventions for at-risk youth*. Kingston, NJ: Civic Research Institute, Inc.
- Leeman, L. W., Gibbs, J. C., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. *Aggressive Behavior, 19*, 281-292.

[Lewis, D. O., Lovely, R., Yeager, C., & Famina, D. \(1989\). Toward a theory of the genesis of violence: A follow-up study of delinquents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28, 431-436.](#)

Lipsey, M. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In McGuire (Ed.), *What Works? Reducing Reoffending*. Chichester, U.K.: Wiley.

Lipsey, M. (1999). Can intervention rehabilitate serious delinquents? *Annals of the American Academy of Political and Social Science*, 564, 142-166.

Lipsey, M., & Landenberger, N. (2003). *Cognitive-behavioral programs for juvenile and adult offenders: A meta-analysis of randomized controlled intervention studies*. Nashville, TN: Vanderbilt University Institute for Public Policy Studies, Center for Evaluation Research and Methodology.

Little, G., & Robinson, K. (1986). *Juvenile MRT: How to escape your prison*. Memphis, TN: Eagle Wing Books.

Lowenkamp, C., Smith, P., & Bechtel, K. (2007). Reducing the harm: Identifying appropriate programming for the low risk offender. *Corrections Today*, December, 50-52.

Martinson, R. (1974). What works? – Questions and answers about prison reform. *The Public Interest*, 35, 22-45.

McGraw, K.O. & Wong, S.P. (1992). A common language effect size statistic. *Psychological Bulletin*, 111, 361-365.

Ogden, T., & Hagen, K. A. (2006). Multisystemic treatment of serious behavior problems in youth: Sustainability of effectiveness two years after intake. *Child and Adolescents Mental Health*, 11, 142-149.

Ogden, T., & Halliday-Boykins, C. A. (2004). Multisystemic treatment of antisocial adolescents in Norway: Replication of clinical outcomes outside of the US. *Child and Adolescent Mental Health*, 9, 77-83.

Pealer, J. A., & Latessa, E. J. (2004). Applying the principles of effective intervention to juvenile correctional programs. *Corrections Today*, December, 26-29.

Pearson, F., Lipton, D., Cleland, C., & Yee, D. S. (2002). The effects of behavioral/cognitive behavioral programs on recidivism. *Crime and Delinquency*, 48, 476-496.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.

- Smith, P., Gendreau, P., Goggin, C. (2004), Correctional treatment: Accomplishments and realities. In P. Van Voorhis, M. Braswell, & D. Lester (Eds.), *Correctional counseling and rehabilitation* (5th ed., pp. 285-293). Cincinnati, OH: Anderson Publishing Company.
- Snyder, H. N., & Sickmund, M. (2006). *Juvenile offenders and victims: 2006 national report*. Washington, DC: U.S Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- Van Voorhis, P., & Lester, D. (2004). Cognitive therapies. In P. Van Voorhis, M. Braswell, & D. Lester (Eds.), *Correctional counseling and rehabilitation* (5th ed., pp. 183-206). Cincinnati, OH: Anderson Publishing Company.
- Weisz, J. R., Martin S. L., Walter, B. R., & Fernandez, G. A. (1991). Differential prediction of young adult arrests for property and personal crimes: Findings of a cohort follow-up study of violent boys from North Carolina's Willie M Program. *Journal of Child Psychology and Psychiatry*, 32, 783-792.
- Whitehead, J.T., & Lab, S. P. (1989). A meta-analysis of juvenile correctional treatment. *Journal of Research and Crime and Delinquency*, 26, 276-295.
- Yochelson, S., & Samenow, S. (1976). *The criminal personality, Vol. 1: A profile for change*. New York: Jason Aronson.

Studies Included in the Meta-Analysis

- Alexander, J. F., Barton, C., Schiavo, R. S., & Parsons, B. V. (1976). Systems-behavioral intervention with families of delinquents: Therapist characteristics, family behavior, and outcome. *Journal of Consulting and Clinical Psychology, 44*, 656-664.
- Alexander, J. F., & Parsons, B. V. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81*, 219-225.
- Apsche, J. A., Evile, M. M., & Murphy, C. (2004). The Thought Change system: An empirically based cognitive behavioral therapy for male juvenile sex offenders, A pilot study. *The Behavior Analyst Today, 5*, 101-107.
- Arbuthnot, J., & Gordon, D. A. (1986). Behavioral and cognitive effects of a moral reasoning development intervention for high-risk behavior-disordered adolescents. *Journal of Consulting and Clinical Psychology, 54*, 208-216.
- Armstrong, T. A. (2003). The effect of Moral Reconciliation Therapy on the recidivism of youthful offenders: A randomized experiment. *Criminal Justice and Behavior, 30*, 668-687.
- Azrin, N., Donohue, B., Besalel, V., Kogan, E., & Acierno, R. (1994). Youth drug abuse treatment: A controlled outcome study. *Journal of Child and Adolescent Substance Abuse, 3*, 1-16.
- Barnowski, R. (2002). *Preliminary findings for the juvenile rehabilitation administration's dialectic behavior therapy program*. Washington State Institute for Public Policy (Document No. 02-07-1203).
- Barnowski, R. (2002). *Washington State's Implementation of Functional Family Therapy of Juvenile Offenders: Preliminary Findings*. Washington State Institute for Public Policy.
- Barton, C., Alexander, J. F., Waldron, H., Turner, C. W., Warburton, J. (1985). Generalizing treatment effects of functional family therapy: Three replications. *The American Journal of Family Therapy, 13*, 16-26.
- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 34*, 105-113.
- 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.

- Burnette, K. D., Swan, E. S., Robinson, K. D., Robinson-Woods, M., & Little, G. L. (2003). Effects of MRT on male juvenile offenders participating in a therapeutic community program. *Cognitive-Behavioral Treatment Review*, *12*(2), 2-5.
- Burnette, K. D., Swan, E. S., Robinson, K. D., Woods-Robinson, M., & Little, G. L. (2004). Treating youthful offenders with Moral Reconciliation Therapy: A recidivism and pre-post test analysis. *Cognitive-Behavioral Treatment Review*, *3/4*, 14-15.
- Camp, B. W., Blom, G. E., Hebert, F., & van Doorninck, W. J. (1977). "Think Aloud": A program for developing self-control in young aggressive boys. *Journal of Abnormal Child Psychology*, *5*, 157-169.
- Cann, J., Falshaw, L., & Friendship, C. (2005). Understanding 'What Works': Accredited cognitive skills programmes for young offenders. *Youth Justice*, *5*, 165-179.
- Coleman, M., Pfeiffer, S., & Oakland, T. (1992). Aggression Replacement Training with behaviorally disordered adolescents. *Behavioral Disorders*, *18*, 54-66.
- Druckman, J. M. (1979). A family-oriented policy and treatment program for female juvenile status offenders. *Journal of Marriage and the Family*, *41*, 627-636.
- Friedman, A. S., Terras, A., & Glassman, K. (2002). Multimodal substance use intervention program for male delinquents. *Journal of Child and Adolescent Substance Abuse*, *11*, 43-65.
- Gibbs, J. C., Arnold, K. D., Ahlbord, H. H., & Cheesman, F. L. (1984). Facilitation of sociomoral reasoning in delinquents. *Journal of Consulting and Clinical Psychology*, *32*, 37-45.
- Glick, B., & Goldstein, A. P. (1987). Aggression Replacement Training. *Journal of Counseling and Development*, *65*, 356-362.
- Gordon, D. A., Arbuthnot, J., Gustafson, K. E., & McGreen, P. (1988). Home-based behavioral systems family therapy with disadvantaged juvenile delinquents. *American Journal of Family Therapy*, *16*, 243-255.
- Greenwood, P. W., & Turner, S. (1993). Evaluation of the Paint Creek Youth Center: A residential program for serious delinquents. *Criminology*, *31*, 263-279.
- Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: 2. Intervention. *Developmental Psychology*, *26*, 269-277.

- Henggeler, S. W., Brondino, M. J., Melton, G. B., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy 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., Clingempeel, W., Brondino, J. M., & 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*, 868-875.
- Henggeler, S. W., Halliday-Boykins, C. A., Cunningham, P. B., Randall, J., Shapiro, S. B., & Chapman, J. E. (2006). Juvenile drug court: Enhancing outcomes by integrating evidence-based treatments. *Journal of Consulting and Clinical Psychology, 74*, 42-54.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology, 60*, 953-961.
- Henggeler, S. W., Melton, G. B., Smith, L. A., Schoenwald, S. K., & Hanley, J. H. (1993). Family preservation using multisystemic treatment: Long-term follow-up to a clinical trial with serious juvenile offenders. *Journal of Child and Family Studies, 2*, 283-293.
- Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance-abusing and-dependent delinquents: Outcomes, treatment, fidelity and transportability. *Mental Health Services Research, 1*, 171-184.
- Henggeler, S. W., Rodick, J. D., Hanson, C. L., Watson, S. M., Borduin, C. M., & Urey, J. R. (1986). Multisystemic treatment of juvenile offenders: Effects on adolescent behavior and family interaction. *Developmental Psychology, 22*, 132-141.
- Ireland, J. (2004). Anger management therapy with young male offenders: An evaluation of treatment outcome. *Aggressive Behavior, 30*, 174-185.
- Jesness, C. F. (1975). Comparative effectiveness of behavior modification and transactional analysis programs for delinquents. *Journal of Consulting and Clinical Psychology, 43*, 758-779.
- Kaminer, Y., & Burlison, J. A. (1999). Psychotherapies for adolescent substance abusers: 15-Month follow-up of a pilot study. *The American Journal of Addictions, 8*, 114-119.
- Kane, L. (2003). Cognitive-behavioral anger management groups and their impact on the recidivism rates among adjudicated youth. *Dissertation Abstracts International, 63*, 3920.
- Kelley, P., Kelley, V., & Williams, B. (1989). Treatment of adolescents: A comparison of individual and family therapy. *Social Casework, 70*, 461-468.

- Kirigin, K. A., Braukmann, C. J., Atwater, J. D., & Wolf, M. M. (1982). An evaluation of teaching-family (Achievement Place) group houses for juvenile offenders. *Journal of Applied Behavior Analysis, 15*, 1-16.
- Klein, N. C., Alexander, J. F., & Parsons, B. V. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology, 45*, 469-474.
- Leeman, L. W., Gibbs, J. C., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. *Aggressive Behavior, 19*, 281-292.
- Liddle, H. A., Dakof, G. A., Parker, K., Diamond, G. S. Barrett, K., & Tejada, M. (2001). Multidimensional family therapy for adolescent drug abuse: Results of randomized clinical trial. *American Journal of Drug and Alcohol Abuse, 27*, 651-687.
- Lochman, J. E. (1992). Cognitive-behavioral intervention with aggressive boys: Three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology, 60*, 426-432.
- McPherson, S. J., McDonald, L. E., & Ryer, C. W. (1983). Intensive counseling with families of juvenile offenders. *Juvenile and Family Court Journal, 34*, 27-33.
- Myers, W. C., Burton, P. R. S., Sanders, P. D., Donat, K. M., Cheney, J., Fitzpatrick, T. M., & Monaco, L. (2000). Project Back-on-Track at one year: A delinquency treatment program for early-career juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1127-1134.
- Ogden, T., & Hagen, K. A. (2006). Multisystemic treatment of serious behavior problems in youth: Sustainability of effectiveness two years after intake. *Child and Adolescents Mental Health Volume, 11*, 142-149.
- Ogden, T., & Halliday-Boykins, C. A. (2004). Multisystemic treatment of antisocial adolescents in Norway: Replication of clinical outcomes outside of the US. *Child and Adolescent Mental Health, 9*, 77-83.
- Pullen, S. (1996). Evaluation of the Reasoning and Rehabilitation cognitive skills development program as implemented in Juvenile ISP in Colorado. Boulder, CO: Colorado Division of Criminal Justice.
- Rowland, M. D., Halliday-Boykins, C. A., Henggeler, S. W., Cunningham, P. B., Lee, T. G., Kruesi, M. J., & Shapiro, S. B. (2005). A randomized trial of multisystemic therapy with Hawaii's Felix Class youths. *Journal of Emotional and Behavioral Disorders, 13*, 13-23.

- Schaeffer, C. M., & Borduin, C. M. (2005). Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology, 73*, 445-453.
- Scherer, D. G., Brondino, M. J., Henggeler, S. W., Melson, G. B. et al. (1994). Multisystemic family preservation therapy: Preliminary findings from a study of rural and minority serious adolescent offenders. *Journal of Emotional and Behavioral Disorders, 2*, 198-206.
- Schlicter, K. J. & Horan, J. J. (1981). Effects of stress inoculation on the anger management and aggression management skills of institutionalized juvenile delinquents. *Cognitive Therapy and Research, 5*, 359-365.
- Stambaugh, L. F., Mustillo, S. A., Burns, B. J., Stephens, R. L., Baxter, B., Edwards, D., & DeKraai, M. (2007). Outcomes from wraparound and multisystemic therapy in a center for mental health services system-of-care demonstration site. *Journal of Emotional and Behavioral Disorder, 15*, 143-155.
- Timmons-Mitchell, J., Bender, M. B., Kishna, M. A., & Mitchell, C. C. (2006). An independent effectiveness trial of multisystemic therapy with juvenile justice youth. *Journal of Clinical Child and Adolescent Psychology, 35*, 227-236.
- Trupin, E. W., Stewart, D. G., Beach, B., & Boesky, L. (2002). Effectiveness of a dialectical behaviour therapy program for incarcerated female juvenile offenders. *Child and Adolescent Mental Health, 7*, 121-127.
- Wallace, A. (2001). Results of MRT utilization in the Las Cruces, New Mexico juvenile drug court. *Cognitive Behavioral Treatment Review, 10*(3/4), 1-2.

Appendix A
Descriptive Statistics

Table A-1. Demographic Data on the Samples

Variable	N	%
Race		
White	4	12
Black	1	3
Asian	1	3
Mixed	16	47
Not Reported	12	35
Gender		
Male	9	27
Female	1	3
Mixed	18	53
Not Reported	6	18
Age		
Juvenile	31	91
Mixed	3	9
Mentally Disordered		
Yes	4	12
No	19	56
Mixed	5	15
Not Reported	6	18
Risk Level		
Low	3	9
Moderate	1	3
High	13	38
Mixed	8	24
Not Reported	9	27

Table A-1. Demographic Data on the Samples – Cont.

Variable	N	%
<hr/>		
History of Violence		
Yes	4	12
No	6	18
Mixed	10	29
Not Reported	14	41
<hr/>		
History of Sex Offense		
Yes	1	3
No	9	27
Mixed	5	15
Not Reported	19	55
<hr/>		
Current Offense		
Non-violent	5	15
Violent	2	6
Mixed	5	15
Not Reported	22	65
<hr/>		

Table A-2. Data on Treatment Descriptors

Variable	N	%
Location		
Institution	6	18
Community	26	77
Residential/Group Home	2	6
Format of Treatment		
Group Session	8	24
Youth and family	18	53
Other	5	15
Not Reported	3	9
Duration of Treatment in Weeks		
1-12	6	17
12 or more	20	59
Not Reported	8	24
CJ Provider		
Yes	11	32
No	22	65
Not Reported	1	3
Type of Treatment		
CBT	11	32
MST	11	32
Family	9	27
Mixed	1	3
Other	2	6

Table A-3. Data on Outcome Measures

Variable	N	%
Type of Recidivism		
Technical Violation	2	6
Re-arrest	21	62
Re-conviction	3	9
Re-incarceration	1	3
Mixed	6	18
Not Reported	1	3
Length of Follow Up		
Less than 2 years	24	71
2 years or more	10	29

Appendix B
Analyses Results

Table B-1. Mean Effect Size

Model	k	N	r	z_r	95% CI	Z (p)
Fixed Effects	34	7188	.10	.10	.08 to .12	8.43 (.0000)
Random Effects	34	7188	.19	.19	.13 to .24	6.57 (.0000)

Table B-2. Mean Effect Size by Type of Treatment Random Effects

Model	k	N	r	z_r	95% CI	Z (p)
CBT	11	3817	.10	.10	.01 to .19	2.06 (.039)
Family	9	1157	.27	.27	.17 to .37	5.22 (.000)
Mixed	1	149	.11	.11	-.15 to .36	0.82 (.415)
MST	11	1072	.24	.25	.16 to .34	5.49 (.000)
Other	2	993	.05	.05	-.12 to .22	0.52 (.602)

Q Between (4) = 11.19; p = 0.02

Table B-3. Mean Effect Size by Risk

Model	k	n	r	z_r	95% CI	Z (p)
Higher	14	1314	.23	.23	.16 to .31	6.00 (.000)
Low	3	147	.22	.23	.02 to .44	2.12 (.034)
Mixed	8	497	.10	.10	.01 to .19	2.15 (.031)

Q Between (2) = 5.14; p = 0.07

Table B-4. Mean Effect Size by Length of TX

Model	k	n	r	z_r	95% CI	Z (p)
1-11 weeks	6	504	.17	.18	.06 to .29	2.93 (.003)
12+ weeks	19	2941	.15	.15	.08 to .22	4.38 (.000)

Q Between (1) = .113; p = 0.74

Table B-5. Mean Effect Size by Intensity of TX

Model	k	N	r	z_r	95% CI	Z (p)
1-20 hours	7	280	.22	.22	.12 to .33	4.02 (.000)
21-40 hours	10	857	.28	.29	.19 to .39	5.71 (.000)
41+ hours	7	4364	.12	.12	.02 to .22	2.36 (.000)

Q Between (2) = 5.75; p = 0.06

Table 9. Mean Effect Size by Provider of TX

Model	k	N	r	z_r	95% CI	Z (p)
CJ Provider	11	5224	.13	.13	.05 to .21	3.07 (.002)
Non CJ-Provider	22	1937	.22	.22	.15 to .22	6.61 (.000)

Q Between (1) = 2.94; p = 0.09

The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews

Mark W. Lipsey¹ and Francis T. Cullen²

¹Institute for Public Policy Studies, Vanderbilt University, Nashville, Tennessee 37212; email: mark.lipsey@vanderbilt.edu

²Division of Criminal Justice, University of Cincinnati, Cincinnati, Ohio 45226-0389; email: cullenft@email.uc.edu

Annu. Rev. Law Soc. Sci. 2007. 3:297–320

First published online as a Review in Advance on July 5, 2007

The *Annual Review of Law and Social Science* is online at <http://lawsocsci.annualreviews.org>

This article's doi:
10.1146/annurev.lawsocsci.3.081806.112833

Copyright © 2007 by Annual Reviews.
All rights reserved

1550-3585/07/1201-0297\$20.00

Key Words

treatment, deterrence, sanctions, offenders, meta-analysis

Abstract

The effects of correctional interventions on recidivism have important public safety implications when offenders are released from probation or prison. Hundreds of studies have been conducted on those effects, some investigating punitive approaches and some investigating rehabilitation treatments. Systematic reviews (meta-analyses) of those studies, while varying greatly in coverage and technique, display remarkable consistency in their overall findings. Supervision and sanctions, at best, show modest mean reductions in recidivism and, in some instances, have the opposite effect and increase reoffense rates. The mean recidivism effects found in studies of rehabilitation treatment, by comparison, are consistently positive and relatively large. There is, however, considerable variability in those effects associated with the type of treatment, how well it is implemented, and the nature of the offenders to whom it is applied. The specific sources of that variability have not been well explored, but some principles for effective treatment have emerged. The rehabilitation treatments generally found effective in research do not characterize current correctional practice, and bridging the gap between research and practice remains a significant challenge.

INTRODUCTION

On any given day in the United States, over 7 million offenders are under some form of correctional supervision (1 of every 32 adults), with approximately one-third incarcerated and the remainder on probation or parole (Glaze & Bonczar 2006). Of those incarcerated, approximately 600,000 return to the community each year (Hughes & Wilson 2002, Travis 2005). These numbers reflect an unprecedented increase during recent decades (Mauer 1999, Patillo et al. 2004). In the early 1970s, state and federal prisons housed fewer than 200,000 inmates, and the rate of incarceration had remained relatively stable for the previous half-century (Blumstein & Cohen 1973, Tonry 2004). Since then the imprisonment rates in other Western industrial nations have varied—some creeping upward, others showing slippage—but the expansion in the United States has left it with a rate that is now 5 times higher than for any other Western country and 7 to 12 times higher than most (Tonry 2004, Tonry & Farrington 2005).

There are complex reasons for the growth of the corrections population, but a contributing factor was, to use Tonry's (2004) term, a changing "sensitivity" about crime and what to do about it. During the late 1960s, crime became highly politicized with conservative officials proposing to restore law and order through punitive get-tough policies. This embrace of punishment as a solution to the crime problem represented a direct attack on the view that the correctional system should rehabilitate offenders (Cullen & Gilbert 1982), an ideal with a long history as the dominant correctional paradigm. The hegemony of this perspective is embodied in the very word corrections, which implies that the purpose of state intervention is to correct or reform offenders.

A key moment in this general critique on rehabilitation was Lipton et al.'s (1975) study that reviewed 231 evaluations of rehabilitation programs. In Martinson's (1974, p. 25) advance summary, he reported that, "[w]ith

few and isolated exceptions, the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism." Martinson (1974, p. 48) was bold enough to ask, "Does nothing work?" The implied answer was no. Soon thereafter, this study was widely interpreted as meaning that nothing works to rehabilitate offenders (Cullen & Gendreau 2000, Cullen & Gilbert 1982). Since this time, the legitimacy of correctional treatment has hinged precariously on the question of its effectiveness; after all, if rehabilitation programs do not work, the justification for their continued use evaporates.

The rehabilitation perspective embodies an assumption that the correctional system is expected to do more than exact just deserts from those who have harmed others—it is expected to reduce crime and foster public safety. To achieve this goal, correctional programs must reduce the potential for the offenders under their charge to reoffend when they are released from supervision or custody. Arguably, punitive correctional techniques might be more effective at dissuading offenders from further criminal behavior than rehabilitation treatment is for reforming them. Whether sanctions or treatment, or some combination, has the greatest effect on reoffense rates is fundamentally an empirical question. Treating it as such offers the potential for an evidence-based perspective on corrections in which offender interventions are evaluated and adopted only if they prove to inhibit criminal behavior (Cullen & Gendreau 2000, MacKenzie 2001, 2006). At present, much of what is done within corrections is not based on sound evidence but, rather, on custom, bureaucratic convenience, and political ideology with results that Latessa et al. (2002) have called "correctional quackery."

Reviewing the Research on the Effectiveness of Correctional Interventions

The purpose of this review is to assess what is known about the effects of correctional

interventions on the recidivism of the offenders exposed to them. Those effects are not easy to ascertain. Simple recidivism rates are largely a function of the input characteristics of the respective offenders, especially risk characteristics such as prior offense history, age, and gender. The only scientifically credible method for assessing intervention effects is a research design that compares recidivism rates for offenders exposed to the intervention with those for a substantially similar control group with no exposure to it. The strongest designs assign relatively large numbers of offenders randomly to intervention and control conditions, maintain high fidelity to the intervention plan, and have little attrition from the assigned conditions or the data collection on the recidivism measures. Such true experiments can be difficult to implement for correctional interventions, and much of the available research comes from quasi-experiments with nonrandomized control groups, modest sample sizes, and varying completion and attrition rates.

Hundreds of experimental and quasi-experimental studies of correctional interventions have been conducted in recent decades. Collectively they provide a large body of relevant evidence, one so large that it is not easy to summarize the patterns in their findings. A particularly comprehensive and discriminating technique for this purpose is meta-analysis, a systematic quantitative form of research synthesis that revolves around statistical effect sizes constructed to represent the differences in outcomes between intervention and control groups across multiple studies (Cooper 1998). For the recidivism outcomes of correctional interventions, the most commonly used effect size statistic is the phi coefficient—the familiar product-moment correlation coefficient applied to the relationship between a binary group variable (intervention versus control) and a binary outcome (did or did not recidivate). Meta-analysis proceeds by examining the distribution of effect sizes across studies (e.g., their

mean and variance) and their relationships with different characteristics of the studies (e.g., the type of intervention and the characteristics of the offenders).

Dozens of meta-analyses have been conducted on different and sometimes overlapping subsets of the hundreds of correctional intervention studies. It is upon these meta-analyses that we mainly rely in summarizing the available evidence about the effects of correctional interventions on recidivism (see also McGuire 2002). Given the dominant influence of punitive approaches in corrections, we first consider the effects of sanctions and supervision. We then move to a more extended assessment of the effects of rehabilitation treatment.

THE EFFECTS OF SANCTIONS AND SUPERVISION ON RECIDIVISM

The existence of a criminal justice system that threatens wrongdoers with arrest and punishment almost certainly causes many people to refrain from crime who, without any risk of detection and penalty, would break the law (Doob & Webster 2003, Levitt 2002). Our focus here, however, is not on the nature and effects of that general deterrent effect but, rather, on what is often called specific deterrence—whether the punishment offenders receive is effective in reducing their subsequent criminal behavior. Two types of research are particularly relevant to this question: evaluations of deterrence-oriented corrections programs and assessments of the effects of prison-term length. It is instructive that both literatures support the conclusion that punishment has little or no effect on recidivism (Akers & Sellers 2004, Cullen et al. 2002).

First, a number of deterrence-oriented correctional interventions aimed at increasing the punishment or control experienced by offenders have been evaluated. Perhaps the most instructive is the research on

intensive supervision programs (ISPs) in which parolees or probationers are placed in small caseloads, face regular and unannounced visits by supervising officers, and are threatened with revocation and incarceration if they misbehave. In a now-classic study, Petersilia & Turner (1993) examined ISPs across 14 sites using random-assignment experimental designs. They found no reductions in recidivism at any of the 14 sites, and in fact, the overall one-year recidivism rate for offenders in the ISPs was higher than for those in the probation-as-usual control groups (37% versus 33%).

Several meta-analyses have examined studies of the effects of increased supervision or other intermediate sanctions on recidivism in comparison to lesser or no supervision or sanctions. **Table 1** summarizes their results. To facilitate comparison, we have converted the reported mean effect sizes to phi coefficients when some other effect size statistic was used by the analyst, with negative values indicating a reduction in recidivism relative to control conditions. To make the magnitude of the recidivism effects more interpretable, we also present a standardized index that shows the corresponding percentage change in recidivism. That index assumes a 0.50 recidivism rate in the average control group and converts the phi coefficient to the equivalent proportionate reduction in that rate for the average intervention group. A phi coefficient of -0.10 , for instance, corresponds to a reduction from a 0.50 to a 0.40 recidivism rate, a 20% decrease ($0.10/0.50$).

As **Table 1** indicates, the meta-analyses that have focused on the effects of probation and parole supervision compared with no supervision, or more intensive supervision compared with regular supervision, have found modest favorable effects, ranging from 2% to 8% reductions in recidivism. Pearson et al.'s (1997) meta-analysis, which reported the largest effect, however, also included group homes in their analysis. More meta-analyses have reported on various groupings of intermediate sanctions that may include

supervision but also encompass fines, restitution contracts, prison visitation, and other such specific deterrence-oriented interventions. Their findings for mean effects on recidivism range widely, from 8% reductions to 26% increases. In general, these findings do not provide consistent support for the view that correctional supervision or intermediate sanctions are especially effective in reducing the recidivism of the offenders to whom they are applied.

Similarly, an interesting and popular deterrence-type program covered in the meta-analyses summarized in **Table 1** is one that attempts to reform offenders, especially juveniles, through a tough love approach. The main examples are prison visitation programs and boot camps. Petrosino et al.'s (2003, p. 41) meta-analysis of Scared Straight and similar programs revealed that "the intervention on average is more harmful to juveniles than doing nothing." That meta-analysis and another by Aos et al. (2001) found that the increases in recidivism produced by these programs were substantial. Similarly, MacKenzie et al.'s (2001) meta-analysis of boot-camp programs for juveniles and adults reported that these interventions had no overall effect on recidivism. For juveniles, Aos et al. (2001) found a mean increase in recidivism.

A second area of research has examined the impact of prison sentences on recidivism. As Levitt (2002, p. 443) noted, "it is critical to the deterrence hypothesis that longer prison sentences be associated with reductions in crime." However, the results are not supportive of the view that incarceration dissuades offenders from reoffending after they are released. Sampson & Laub's (1993) longitudinal study using the Gluecks' Boston-area data showed that imprisonment increased recidivism by weakening social bonds (e.g., decreased job stability). Using a matched sample of felony offenders in California, Petersilia et al. (1986) found that those sent to prison had higher recidivism rates than those placed on probation. More recently, Spohn & Holleran (2002) found a similar result for

Table 1 Meta-analyses of the effects of sanctions and supervision on recidivism

Meta-analysis report	Type of intervention	Mean effect size ^a (N)	Change in recidivism ^b
Supervision			
Pearson et al. 1997	Community supervision (adults): ordinary probation and parole, intensive supervision probation and parole, group homes	-0.04 (52)	-8%
Lipsey & Wilson 1998	Ordinary probation and parole, intensive supervision probation and parole, restitution (juveniles)	-0.04 ^c (22)	-8%
Aos et al. 2001	Intensive probation or parole supervision (juveniles)	-0.02 ^c (20)	-4%
Aos et al. 2001	Intensive probation or parole supervision (adults)	-0.01 ^c (22)	-2%
Intermediate sanctions			
Andrews et al. 1990	Criminal sanctions (juveniles and adults): regular processing versus diversion, more versus less severe dispositions, restitution contracts	+0.07 (16)	+14%
Petrosino 1997	Deterrence (juveniles and adults): intensive supervision versus traditional parole or probation, arrest versus police mediation	0.00 (23)	0%
Cleland et al. 1997	Criminal sanctions (juveniles): variations in supervision intensity, fines, criminal justice processing, or other specific deterrents	-0.04 (61)	-8%
Cleland et al. 1997	Criminal sanctions (adults): variations in supervision intensity, fines, criminal justice processing, or other specific deterrents	-0.02 (83)	-4%
Smith et al. 2002	Intermediate sanctions (juveniles and adults): intensive supervision, arrest, fines, restitution, boot camps, Scared Straight, drug testing, electronic monitoring	-0.01 (74)	-2%
Lipsey & Wilson 1998	Prison visitation, shock incarceration (juveniles)	+0.01 ^c (6)	+2%
Aos et al. 2001	Prison visitation, Scared Straight (juveniles)	+0.06 ^c (8)	+12%
Petrosino et al. 2003	Prison visitation, Scared Straight (juveniles)	+0.13 (9)	+26%
Confinement			
Pearson et al. 1997	Incarceration and punishment (adults): mostly prison or jail sentences	+0.02 (26)	+4%
Smith et al. 2002	Longer versus shorter prison sentences (juveniles and adults)	+0.03 (26)	+6%
Smith et al. 2002	Incarceration versus community supervision (juveniles and adults)	+0.07 (31)	+14%
Villettaz et al. 2006	Custodial versus noncustodial sentences (juveniles and adults)	+0.02 (5)	+4%
MacKenzie et al. 2001	Boot camps (juveniles and adults)	0.00 (44)	0%
Aos et al. 2001	Boot camps (juveniles)	+0.05 ^c (10)	+10%
Aos et al. 2001	Boot camps (adults)	0.00 ^c (22)	0%

^aPhi coefficient; unweighted mean when available. A negative sign means less recidivism for the intervention condition. Cohen's *d* effect sizes converted to phi as $\phi = d/\sqrt{4 + d^2}$; odds ratios converted to *d* as $d = \text{Log}(OR)/2$, then *d* converted to phi (this gives the phi that occurs with a 0.50 control recidivism and the given odds ratio).

^bDifference between the recidivism rate for the intervention and a control recidivism rate assumed to be 0.50 that corresponds to the given effect size.

^c Weighed or adjusted for methodological quality.

a sample from Jackson County, Missouri. Studies from Canada (Smith 2006) and the Netherlands (Nieuwbeerta et al. 2006) also show a criminogenic effect of imprisonment. As might be anticipated, none of the meta-analyses of studies of this sort (summarized in **Table 1**) found mean recidivism reductions for correctional confinement. The two meta-analyses that found essentially zero effects focused on boot camps, which feature relatively short-term custodial care. Those summarizing studies of incarceration compared with community supervision, or longer prison terms compared with shorter ones, all found that the average effect was increased recidivism.

Methodologically rigorous studies of the effects of incarceration are especially difficult to conduct—random assignment of convicted offenders to either prison or a non-prison alternative is not generally viewed as an acceptable sentencing procedure. The quasi-experimental studies that address this issue, however, use varied methods with different strengths and weaknesses. It is notable that no systematic synthesis of that research finds generally favorable effects on recidivism.

In sum, research does not show that the aversive experience of receiving correctional sanctions greatly inhibits subsequent criminal behavior. Moreover, a significant portion of the evidence points in the opposite direction—such sanctions may increase the likelihood of recidivism. The theory of specific deterrence inherent in the politically popular and intuitively appealing view that harsher treatment of offenders dissuades them from further criminal behavior is thus not consistent with the preponderance of available evidence. If, among their other purposes, correctional interventions are expected to have a net positive effect on public safety by reducing the reoffense rates of convicted offenders, reliance on punitive approaches does not appear to be sufficient for the task.

THE EFFECTS OF REHABILITATION TREATMENT ON RECIDIVISM

Rehabilitation treatment is distinguished from correctional sanctions by the centrality of interactions with the offenders aimed at motivating, guiding, and supporting constructive change in whatever characteristics or circumstances engender their criminal behavior or subvert their prosocial behavior. It is typically provided in conjunction with some form of sanction (e.g., incarceration or probation) but is not defined by that sanction and, in principle, could be delivered without any accompanying sanction. Cognitive-behavioral therapy, for instance, involves exercises and instruction designed to alter the dysfunctional thinking patterns exhibited by many offenders [e.g., a focus on dominance in interpersonal relationships, feelings of entitlement, self-justification, displacement of blame, and unrealistic expectations about the consequences of antisocial behavior (Walters 1990)].

Hundreds of studies of the effects of various rehabilitation treatments on recidivism have been conducted with both juvenile and adult offenders in community-based and residential correctional programs. The findings of those studies, in turn, have been examined in numerous meta-analyses. Some of these overlap in the studies they cover, and some researchers have contributed more than one meta-analysis. At the same time, there is considerable diversity in the meta-analytic approaches and techniques used and the potential for different meta-analyses to reach different conclusions. Our purpose here is to take a broad overview of virtually all the existing meta-analyses on rehabilitation treatments as a way to appraise the current state of evidence about their effectiveness for reducing recidivism.

The most general result available from these meta-analyses is an estimate of the overall mean effect size across diverse samples of studies of different rehabilitation treatments

Table 2 Meta-analyses of the effects of rehabilitation treatment generally on recidivism

Meta-analysis report	Age of offenders	Treatment setting	Mean effect size ^a (N)	Change in recidivism ^b
Garrett 1985	Juveniles	Residential	-0.05 ^c (19)	-10%
Whitehead & Lab 1989	Juveniles	Community and residential	-0.12 ^d (50)	-24%
Andrews et al. 1990	Juveniles and adults	Community and residential	-0.10 (88)	-20%
	Juveniles	Community and residential	-0.10 (70)	-20%
	Adults	Community and residential	-0.11 (18)	-22%
	Juveniles and adults	Community	-0.11 (68)	-22%
	Juveniles and adults	Residential	-0.07 (20)	-14%
Petrosino 1997	Juveniles and adults	Community and residential	-0.10 ^e (115)	-20%
	Juveniles	Community and residential	-0.12 ^e (55)	-24%
	Adults	Community and residential	-0.07 ^e (53)	-14%
Cleland et al. 1997	Juveniles and adults	Community and residential	-0.08 (515)	-16%
	Juveniles	Community and residential	-0.08 (288)	-16%
	Adults	Community and residential	-0.07 (227)	-14%
Lipsey & Wilson 1998	Juveniles	Community	-0.13 ^f (117)	-26%
	Juveniles	Residential	-0.07 ^f (83)	-14%
Illescas et al. 2001	Juveniles and adults	Community and residential	-0.17 ^g (22)	-34%
	Juveniles	Community and residential	-0.19 ^g (13)	-38%
	Adults	Community and residential	-0.10 ^g (15)	-20%
Latimer et al. 2003	Juveniles	Community and residential	-0.09 (156)	-18%

^aPhi coefficient; unweighted mean when available. A negative sign means less recidivism for the intervention condition. Cohen's *d* effect sizes converted to phi as $\phi = d/\sqrt{4 + d^2}$; odds ratios converted to *d* as $d = \text{Log}(OR)/2$, then *d* converted to phi (this gives the phi that occurs with a 0.50 control recidivism and the given odds ratio).

^bDifference between the recidivism rate for the intervention and a control recidivism rate assumed to be 0.50 that corresponds to the given effect size.

^cSubset with random or matched designs and recidivism outcomes.

^dComputed from **Table 1**.

^eRandomized studies only.

^fUnweighted means computed from original data.

^gEuropean studies; subset with controls.

applied to general offender samples. **Table 2** summarizes the major meta-analyses that focus on recidivism outcomes for adjudicated offenders. As shown in **Table 2**, every one of these meta-analyses found mean effect sizes¹ favorable to treatment, and none found less than a 10% average reduction in recidivism.

¹For statistical analysis, effect sizes are often weighted by a term reflecting the size of the sample on which they are based. For rehabilitation studies, however, sample size is often correlated with other study characteristics, e.g., methodological quality and how well the treatment was implemented. To avoid adjusting inappropriately for these other characteristics, we report the unweighted effect size means in **Table 1** whenever available.

Most of their mean effect sizes represent recidivism reductions in the 20% range, varying upward to nearly 40%. It is especially notable that there is no overlap in the range of mean effect sizes found in meta-analysis of rehabilitation treatment and that found for meta-analyses of the effects of sanctions and supervision (**Table 1**). The smallest mean recidivism effect size found in any meta-analysis of a general collection of rehabilitation studies is bigger than the largest one found in any meta-analysis of the effects of sanctions.

More meta-analysis has been completed on treatment for juveniles than for adults,

making it difficult to assess whether the effects are comparable. Three of the four broad meta-analyses that cover both (Cleland et al. 1997, Illescas et al. 2001, Petrosino 1997) found larger effects for juveniles, although the differences are slight in two cases. None of the three, however, explored differences in the nature and quality of the treatments provided to juveniles in comparison to adult offenders or, conversely, the differential responsiveness of juveniles and adults to similar treatments. On balance, therefore, we have less synthesis of rehabilitation research for adult offenders and, correspondingly, less ability to examine the robustness of the findings across multiple analyses.

Table 2 also differentiates, where possible, the average findings for community-based treatment for offenders under probation or parole supervision and treatment in residential settings for incarcerated offenders. This distinction also has not been explored fully. Favorable effects from treatment are found in both settings, but in the two meta-analyses that break out separate results (Andrews et al. 1990, Lipsey & Wilson 1998), the mean effect sizes for community-based treatment are larger than those for residential treatment. Neither addresses the question of whether this difference is associated with differences in the nature or quality of the treatment, the characteristics of the offenders treated in these different settings, or the influence of these quite different contexts on the treatment effects.

Skeptics might question whether the broadly positive average effects of rehabilitation treatments found in the studies included in these meta-analyses actually reflect the benefits of treatment or some equally broad and pervasive upward bias in the effect estimates generated by those studies. One such possible bias that is well known to meta-analysts is the tendency for published studies to show larger effects than unpublished ones, presumably because of the selection processes associated with the development, submission, and review of manuscripts for journal publica-

tion (Rothstein et al. 2005). Mean effect sizes from meta-analyses that include only published studies, or which greatly overrepresent them, may thus be inflated. This is an unlikely explanation for the findings in **Table 2**, however. Most of the meta-analyses in **Table 2** include unpublished studies, which, even if underrepresented, should diminish the influence of publication bias on their results. In addition, direct comparisons between the mean effect sizes for published and unpublished studies appear in some meta-analyses of specific treatments (we discuss these more fully below). A few of these do find larger effects reported in published studies (Gallagher et al. 1999, Landenberger & Lipsey 2005, Reitzel & Carbonell 2006), but others find the reverse (Illescas et al. 2001, Wilson et al. 2006) or differences going both ways (Mitchell et al. 2006). In all cases, however, the unpublished studies also show mean positive effects, and the differences between published and unpublished studies are not large enough to account for the generally positive overall effects. This is perhaps not surprising in a research area in which, historically, finding and reporting no difference have not been viewed as uninteresting, and indeed, at one time were almost normative.

Another possible source of broad bias in the findings of rehabilitation studies relates to the quality of the research designs, in particular whether the treatment and control groups are created through random assignment in a true experimental design. Random assignment is not always feasible in criminal justice settings, and many studies use weaker quasi-experimental designs in which control groups are selected from convenient groups of untreated offenders, usually with some attempt to match relevant background characteristics. Quasi-experimental studies are not automatically biased, but they are vulnerable to bias stemming from initial uncontrolled differences between the comparison groups that then carry forward to produce differences on the outcome measures that mimic treatment effects. If such bias occurs

and, furthermore, tends to be in the same direction across different quasi-experimental studies, the mean effects found in those studies would accordingly underestimate or overestimate the actual treatment effects. Weisburd et al. (2001), for instance, examined the studies in the Maryland report on what works in crime prevention (Sherman et al. 1997) and found that studies with weaker designs were more likely to report favorable intervention effects than those with stronger designs. If the large number of quasi-experimental studies of rehabilitation effects is similarly biased, this might account for the generally positive effects summarized in **Table 2** (see also Farabee 2006).

Many meta-analyses of rehabilitation have investigated this source of possible bias by looking at the relationship between effect sizes and the methodological quality of the studies that generated them, especially with regard to the use of randomized or non-randomized designs. **Table 3** summarizes their findings, indicating whether larger effect estimates were found for the studies with stronger designs, weaker designs, or neither (phi coefficients equal to within ± 0.01). In some cases there were fairly large differences, in others relatively small, but overall there was little indication of a consistent bias. In particular, the mean effect sizes from non-randomized studies, or those judged to have

Table 3 Meta-analyses that compare effects from weaker and stronger research designs

Meta-analysis report (intervention)	Mean effect size for weaker designs ^a (N)	Mean effect size for stronger designs ^a (N)	Favors
Andrews et al. 1990 (mixed)	-0.11 (42 ^b)	-0.10 (38 ^b)	Neither
Dowden et al. 2003 (relapse prevention)	-0.13 (18 ^b)	-0.21 (6 ^b)	Stronger
Feder & Wilson 2005 (batterers)	+0.07 (4)	-0.13 (7)	Stronger
Gallagher et al. 1999 (sex offenders)	-0.22 (23)	-0.44 (3)	Stronger
Hanson et al. 2002 (sex offenders)	-0.12 (17)	+0.01 (3)	Weaker
Latimer 2001 (family)	-0.18 (19 ^b)	-0.10 (16 ^b)	Weaker
Lösel & Schmucker 2005 (sex offenders)	-0.15 (23)	-0.10 (6)	Weaker
MacKenzie et al. 2001 (boot camps)	-0.01 (39)	+0.07 (5)	Weaker
Mitchell et al. 2006 (counseling)	-0.11 (23)	-0.02 (2)	Weaker
Mitchell et al. 2006 (drug relapse)	-0.04 (18)	-0.22 (2)	Stronger
Mitchell et al. 2006 (therapeutic community)	-0.07 (28)	-0.16 (2)	Stronger
Pearson et al. 2002 (behavioral)	-0.10 (61)	-0.21 (7)	Stronger
Illescas et al. 2001 (mixed)	-0.17 (19)	-0.02 (3)	Weaker
Wilson & Lipsey 2000 (challenge programs)	-0.10 ^c (13)	-0.09 ^c (9)	Neither
Wilson et al. 2000 (vocational)	-0.11 (50)	-0.10 (3)	Neither
Wilson et al. 2005b (boot camps)	-0.01 (39)	+0.02 (4)	Weaker
Wilson et al. 2006 (drug courts)	-0.12 (49)	-0.13 (5)	Neither
Correlations and regression coefficients			
Cleland et al. 1997 (mixed)	Regression coefficient for random = -0.025		Weaker
Dowden & Andrews 1999 (female offenders)	Partial correlation for random = -0.10		Weaker
Landenberger & Lipsey 2005 (cognitive-behavioral)	Correlation for random = 0.04		Stronger

^aPhi coefficient; a negative sign means less recidivism for the intervention condition. Cohen's *d* effect sizes converted to phi as $\phi = d/\sqrt{4+d^2}$; odds ratios converted to *d* as $d = \text{Log}(OR)/2$, then *d* converted to phi.

^bEstimated.

^cComputed from table 7 in Wilson & Lipsey (2000).

weaker designs, were sometimes found to be larger, sometimes smaller, and sometimes substantially the same as those from randomized studies. The high proportion of quasi-experimental studies among those investigating rehabilitation effects, therefore, surely adds variability to the effect estimates but, overall, does not appear to bias them in one direction or the other.

In this regard, the systematic difference in the mean outcomes of the studies of sanctions and treatments mentioned above (**Tables 1 and 2**) is informative. Studies of sanctions, especially regular and intensive supervision and intermediate sanctions such as restitution and boot camps, use a mix of experimental and quasi-experimental designs similar to those used to study treatment. Any pervasive bias associated with the inclusion of studies with weaker research designs in meta-analyses of treatment studies should similarly bias the meta-analysis results for sanctions. The dramatically larger effect sizes found for treatment under these circumstances are thus difficult to attribute entirely to methodological bias. Similarly, many of the meta-analyses of sanctions and treatment include comparable mixes of published and unpublished studies. Indeed, in some cases the results for these different interventions are breakouts from the same overarching meta-analyses and thus involve similar literature search strategies and inclusion criteria. Whatever publication bias is present, therefore, should affect both sets of studies and cannot by itself explain the substantially larger effects found for treatment.

The global question of whether rehabilitation treatment works is thus answered affirmatively by the favorable mean effects on recidivism found by every meta-analyst who has conducted a systematic synthesis of a broad sample of the available experimental and quasi-experimental research. No general bias in the findings of that research or the meta-analyses that summarize them has yet been demonstrated which is sufficient to negate the overall positive findings. It is the

case, however, that the available research is unevenly distributed and synthesized. Treatment effects for juvenile offenders have been more thoroughly analyzed and documented than for adult offenders, and possible differential effects of treatment in community and residential settings have not been well explored. These are matters of practical importance to the juvenile and criminal justice system policy makers responsible for rehabilitation programs, and they warrant more attention from researchers.

The Importance of the Large Variability in Rehabilitation Effects

Knowing the average effects of rehabilitation treatment has little specific practical or theoretical value unless all treatments produce essentially that average effect. That is most decidedly not what has been found for rehabilitation treatments. One of the most general and striking findings of research on this topic is the great variability of the recidivism effects across different treatments and different studies. Within any broad sample of studies, one finds many near-zero and even negative effect sizes at one end of the effect size distribution, whereas the other end extends to impressively large effects representing reductions in recidivism of 50% and higher. A certain amount of that variability, of course, reflects only statistical noise and unsystematic differences in study methods and procedures. However, much of it is related to substantive characteristics of the treatments and the offender samples to which they are applied (Wilson & Lipsey 2001). The most important challenge for contemporary rehabilitation research is to identify the factors that most influence the likelihood of positive treatment effects. Such knowledge is needed to support the design of optimally effective treatment in practice settings and to guide theory toward a better understanding of the change mechanisms through which offender behavior can be altered. Research to date has been dominated by issues of whether anything works, with

relatively little attention to questions of what works best, for whom, under what circumstances, and why. In the sections that follow, we review the limited insight that current research provides into such matters.

Type of treatment and the magic bullet hypothesis.

In some human service areas, different treatment approaches within the broad mainstream of practice seem to produce relatively similar effects. For instance, relatively modest differences have been found between the effects of different intervention modalities for school-violence prevention programs (Wilson et al. 2003a). Rehabilitation programs for offenders, in contrast, show marked differentiation in the effects of different types of treatment. **Table 4** summarizes the meta-analysis research on the effects of relatively specific treatments and treatment approaches. It is sequenced to keep results for the same or similar interventions together while ordering them roughly from the largest mean effect sizes to the smallest. These mean effect sizes range from a near-zero recidivism reduction (with one showing an increase in recidivism) to values representing more than 50% reductions. Focusing on specific treatments means that many of these findings are based on a small number of studies and thus are not stable, but even allowing for that, the range of mean effect sizes is quite remarkable.

One thing to note about the findings in **Table 4**, incidentally, is the further support they give to the positive effects of rehabilitation treatment. Despite many differences between studies, when the results of those investigating any given therapeutically oriented treatment are averaged together, the results are positive in the vast majority of cases. The only instances of treatment showing zero effects on recidivism, or recidivism increases, are for small sets of studies classified by Garrett (1985) as involving behavioral or psychodynamic treatment. Two later meta-analyses that used a behavioral classification both found positive effects on

recidivism (Gottshalk et al. 1987, Pearson et al. 2002) in contrast to the negative effect Garrett reported. There is thus little indication of treatment ineffectiveness in these results and a near universal indication that most of the rehabilitation treatments with sufficient research to be included in a meta-analysis are effective. Moreover, those treatments that show the largest average effects tend to be those based on better developed theory and research about their approach to bringing about change—for example, multidimensional treatment foster care, multisystemic therapy, family therapy, treatment for sex offenses, and cognitive-behavioral therapy.

Although the type of treatment clearly matters in relation to recidivism effects, it is not clear what accounts for those effects in the different treatment types. Rehabilitation treatments of a given type do not generally follow a common well-defined treatment protocol. Research-based manualized programs, such as multisystemic therapy and functional family therapy, have that character, but they are not typical of the programs represented in the research literature. Much of the available research involves more generic treatment types, such as family counseling or vocational training, which vary from provider to provider. Moreover, treatment elements are often mixed and combined in varied ways (e.g., drug-education classes combined with individual counseling and vocational training).

Most revealing, perhaps, is that even for a relatively well-defined program type, different studies of different program implementations show variable effects. Virtually all the meta-analyses summarized in **Table 4** that examined the variation in effect sizes across studies found significant heterogeneity. Programs that, on average, show relatively large effects nonetheless produce small effects in some instances, and generally weaker programs sometimes show large effects. Factors other than the type of treatment, therefore, must be influencing the effects actually

Table 4 Meta-analyses of the effects of specific treatment types on recidivism

Meta-analysis report	Treatment	Mean effect size ^a (N)	Change in recidivism ^b
Andrews et al. 1990	Appropriate: behavioral and social-learning treatment addressing risk and needs (juveniles and adults)	-0.30 (39)	-60%
Lipsey & Wilson 1998	Interpersonal skills (juveniles)	-0.18 ^c (6)	-36%
Aos et al. 2001	Multidimensional treatment foster care (juveniles)	-0.18 ^c (2)	-36%
Curtis et al. 2004	Multisystemic therapy (juveniles)	-0.24 (7)	-46%
Aos et al. 2001	Multisystemic therapy (juveniles)	-0.15 ^c (3)	-30%
Littell et al. 2005	Multisystemic therapy (juveniles)	-0.08 (5)	-16%
Lipsey & Wilson 1998	Multimodal (juveniles)	-0.14 ^c (23)	-28%
Lipsey & Wilson 1998	Teaching family home (juveniles)	-0.16 ^c (6)	-32%
Woolfenden et al. 2002	Family and parenting (juveniles)	-0.27 (5)	-52%
Latimer 2001	Family intervention (juveniles)	-0.15 (35)	-30%
Lipsey & Wilson 1998	Family counseling (juveniles)	-0.13 ^c (8)	-26%
Aos et al. 2001	Family therapy (juveniles)	-0.10 ^c (13)	-20%
Lipsey & Wilson 1998	Individual counseling (juveniles)	-0.16 ^c (16)	-32%
Garrett 1985	Life skills (juveniles)	-0.15 (3)	-30%
Dowden & Andrews 1999	Programs for females (juveniles and adults)	-0.14 (26)	-28%
Reitzel & Carbonell 2006	Programs for sex offenders (juveniles)	-0.24 (9)	-46%
Gallagher et al. 1999	Programs for sex offenders (juveniles and adults)	-0.18 (26)	-36%
Hanson et al. 2002	Programs for sex offenders (juveniles and adults)	-0.14 (31)	-28%
Lösel & Schmucker 2005	Programs for sex offenders (juveniles and adults)	-0.13 (49)	-26%
Hall 1995	Programs for sex offenders (juveniles and adults)	-0.12 (12)	-24%
Aos et al. 2001	Programs for sex offenders (juveniles)	-0.06 ^c (5)	-12%
Aos et al. 2001	Cognitive-behavioral therapy for sex offender (adults)	-0.05 ^c (7)	-10%
Wilson et al. 2005a	Cognitive-behavioral therapy (juveniles and adults)	-0.16 (11)	-32%
Dowden et al. 2003	Relapse prevention (juveniles and adults)	-0.15 (24)	-30%
Pearson et al. 2002	Cognitive-behavioral therapy (juveniles and adults)	-0.14 (44)	-28%
Landenberger & Lipsey 2005	Cognitive-behavioral therapy (juveniles and adults)	-0.11 (58)	-22%
Aos et al. 2001	Aggression replacement training (juveniles)	-0.09 ^c (4)	-18%
Aos et al. 2001	Cognitive-behavioral therapy (adults)	-0.04 ^c (14)	-8%
Tong & Farrington 2006	Reasoning and rehabilitation cognitive-behavioral therapy (juveniles and adults)	-0.04 (15)	-8%
Lösel 1995	Social-therapeutic prisons (adults)	-0.12 (11)	-24%
Pearson et al. 1997	Milieu therapy (adults)	-0.12 (16)	-24%

(Continued)

Table 4 (Continued)

Meta-analysis report	Treatment	Mean effect size ^a (N)	Change in recidivism ^b
Lipsey & Wilson 1998	Milieu therapy (juveniles)	-0.06 ^c (3)	-12%
Pearson et al. 1998	Challenge programs (juveniles and adults)	-0.15 (12)	-30%
Wilson & Lipsey 2000	Challenge programs (juveniles)	-0.09 (22)	-18%
Lipsey & Wilson 1998	Challenge programs (juveniles)	-0.05 ^c (9)	-10%
Nugent et al. 2003	Victim-offender mediation (juveniles)	-0.09 (15)	-18%
Wilson et al. 2006	Drug courts (juveniles and adults)	-0.12 (50)	-24%
Lowenkamp et al. 2005	Drug courts (juveniles and adults)	-0.07 (22)	-14%
Aos et al. 2001	Drug courts (adults)	-0.04 ^c (27)	-8%
Pearson et al. 1997	Drug and alcohol treatment (adults)	-0.10 (41)	-20%
Mitchell et al. 2006	Drug treatment (juveniles and adults)	-0.08 (52)	-16%
Pearson & Lipton 1999b	Programs for drug abusers (juveniles and adults)	-0.07 (20)	-14%
Lipsey & Wilson 1998	Drug treatment (juveniles)	-0.05 ^c (5)	-10%
Aos et al. 2001	Therapeutic community (adults)	-0.03 ^c (16)	-6%
Aos et al. 2001	Drug treatment (adults)	-0.02 ^c (27)	-4%
Lipsey & Wilson 1998	Behavioral programs (juveniles)	-0.20 ^c (9)	-40%
Pearson et al. 2002	Behavioral and incentive programs (juveniles and adults)	-0.07 (23)	-14%
Gottshalk et al. 1987	Behavioral programs (juveniles)	-0.06 (14)	-12%
Garrett 1985	Behavioral programs (juveniles)	0.04 (6)	+8%
Lipsey & Wilson 1998	Social casework (juveniles)	-0.07 ^c (6)	-14%
Feder & Wilson 2005	Programs for batterers (adults)	-0.06 (9)	-12%
Pearson et al. 1997	Group counseling (adults)	-0.06 (17)	-12%
Lipsey & Wilson 1998	Group counseling, guided group (juveniles)	-0.04 ^c (25)	-8%
Wilson et al. 2000	Educational, vocational, and work programs (adults)	-0.10 (53)	-20%
Pearson & Lipton 1999a	Educational and vocational programs (juveniles and adults)	-0.05 (72)	-10%
Aos et al. 2001	Education, vocational, and employment programs (adults)	-0.03 ^c (16)	-6%
Visher et al. 2005	Employment programs (adults)	-0.01 (10)	-2%
Lipsey & Wilson 1998	Vocational and employment programs (juveniles)	0.00 ^c (10)	0%
Garrett 1985	Psychodynamic treatment (juveniles)	0.00 (10)	0%

^aPhi coefficient; unweighted mean when available. A negative sign means less recidivism for the intervention condition. Cohen's *d* effect sizes converted to phi as $\phi = d/\sqrt{4 + d^2}$; odds ratios converted to *d* as $d = \text{Log}(OR)/2$, then *d* converted to phi (this gives the phi that occurs with a 0.50 control recidivism and the given odds ratio).

^bDifference between the recidivism rate for the intervention and a control recidivism rate assumed to be 0.50 that corresponds to the given effect size.

^cWeighted or adjusted for methodological quality.

achieved; no programs or program types have been identified that consistently produce positive effects. The main implication of this situation is that effective programs cannot be defined adequately in terms of the type of

treatment they represent. It follows that the widespread model-program lists and rankings of named programs and program types that identify them this way can, at best, provide only general guidance for effective programs.

It is simply not consistent with the research evidence to view rehabilitation programs as well-defined magic bullets, the right one of which, if found, will have a big impact on recidivism.

An alternative perspective more consistent with the nature of the variability found in treatment effects is to characterize more and less effective programs in terms of treatment principles. From this perspective, we ask not what program packages are most effective, but what characteristics are common to effective programs. Unfortunately, neither current research nor meta-analysis of that research is sufficiently differentiated to provide good empirical guides to effective program principles. Moreover, whatever theory underlies the different treatment approaches is also not generally well developed enough to support much conceptual analysis of similarities and differences.

Andrews and his colleagues have gone the furthest in attempting to delineate the principles that characterize effective rehabilitation treatments (Andrews 1995, Andrews et al. 1990, Gendreau 1996). With regard to the nature of the treatment provided, they describe a need principle and a responsivity principle that are associated with the likelihood of positive effects (they also advanced a risk principle, which we address below). According to the need principle, treatment has larger effects if it addresses the criminogenic needs of the offender—those dynamic risk factors predictive of subsequent criminal conduct. Criminogenic needs include antisocial attitudes and peer associations, lack of self-control and self-management skills, drug dependencies, and other such malleable characteristics associated with criminal offense rates. The responsivity principle, in turn, identifies effective treatment as that which is generally capable of actually bringing about change in the targeted criminogenic needs and which is specifically matched to the learning styles and characteristics of the offenders treated. This principle skates on the edge of circularity—effective treatment is that which is capable

of affecting risk factors for recidivism; treatment that changes those risk factors is effective. Andrews et al. escape this circularity by drawing on a large body of research and theory about behavioral change to define responsive treatments as those that use the cognitive-behavioral and social-learning approaches shown to be generally effective in influencing a variety of behaviors. Fundamentally, then, the responsivity principle claims that there are larger effects from treatments that provide learning and skill-building experiences aimed at changing specific problem behaviors through such techniques as practice, role playing, modeling, feedback, verbal guidance, and reinforcement.

In a series of meta-analyses, Andrews et al. have shown that studies of interventions they judge as conforming to their need and responsivity principles do indeed show considerably larger effects on recidivism than those that do not (Andrews et al. 1990, Andrews & Bonta 2006, Gendreau et al. 2006). In one recent meta-analysis, for instance, they showed that programs that departed from the need, responsivity, and risk principles had a mean effect size in the vicinity of zero, whereas those that embodied those principles achieved an effect size of $\phi = 0.26$, equivalent to a recidivism reduction of approximately 50% (Andrews & Bonta 2006, p. 335). Few other primary researchers or meta-analysts have explored these treatment principles or proposed any alternatives [Cleland et al. (1997) are an exception among the meta-analysts]. The general notion that rehabilitation treatment is effective to the extent that it targets malleable risk factors for recidivism and uses techniques that, in fact, induce positive change in those risk factors is plausible and consistent with the evidence on effective treatments. It has many testable implications for the variables that should mediate recidivism reductions (criminogenic risk factors), the matching of treatment with offender characteristics according to their particular risk factors, and the differential effectiveness of different approaches.

The critical importance of treatment implementation and integrity. A major source of variability in the treatment effects on the offenders' recidivism relates to how well the treatment program is implemented. Implementation has several facets. One is simply whether the treatment delivered is the treatment intended, a matter of treatment integrity or fidelity. Therapists may, for instance, substitute their own preferred treatment techniques for those prescribed by the rehabilitation program. Or, a poor quality version of the treatment may be delivered by poorly trained or unmotivated providers. Another facet of implementation is dosage—the amount of treatment offered and received. The strength of an otherwise effective treatment may be diminished if too little is offered (e.g., five sessions are provided when ten are required to have good effects). It may also be undermined by the lack of participation by the offenders being served if they are present but not engaged, have poor attendance at treatment events, or drop out before completing treatment.

It is a truism that a treatment that is not delivered cannot have effects. What is not so obvious is how frequently treatments are poorly implemented, even in research studies, and how readily that compromises program effects. Unfortunately, the degree and quality of implementation are not well documented in most treatment-effectiveness studies. For a research synthesis to examine these factors, one must use very approximate variables to represent the nature of the implementation in each study (e.g., crude treatment completion rates, indications of any monitoring of service delivery, the number of sessions or duration, miscellaneous reports of implementation problems). Despite their coarseness, the meta-analyses that include such indicators universally find that they are strongly related to the size of the effects on recidivism (e.g., Andrews & Dowden 2005, Landenberger & Lipsey 2005, Latimer 2001, Lipsey & Wilson 1998, Lösel & Schmucker 2005).

Andrews & Dowden (2005) conducted the most extensive analysis of the relationship between indicators of the integrity of treatment implementation and recidivism effect sizes. They found correlations ranging from 0.06 to 0.39, with especially revealing relationships appearing for such indicators as having a treatment manual ($r = 0.24$), staff trained in the treatment ($r = 0.26$), and clinical supervision of treatment delivery ($r = 0.20$). These correlations, however, are mainly based on simple indications in the research reports about whether these elements were present. It is not clear how often such information goes unreported or how much difference the extent or quality of such implementation characteristics makes to treatment effectiveness. As Gendreau et al. (1999) observed, program implementation seems to be a forgotten issue in rehabilitation research.

Characteristics of the offenders. Another aspect of rehabilitation treatment that research has not explored well is the potential for differential effects for different offenders. The available research has touched on several relevant types of offender characteristics. One category relates to familiar demographic distinctions: age, gender, and ethnicity. Another has to do with the level of risk for subsequent offending—characteristics of offenders such as prior offense histories and associations with criminal peers that are predictive of the probability of recidivism. The third concerns the treatment needs of different offenders—the particular problems and circumstances that most strongly propel their criminal behavior, for example, drug addiction, poor impulse control, and unemployment.

As noted earlier, few meta-analyses have directly compared the effects of treatment on juveniles versus adults, and none has done so while attempting to hold other factors constant. Those that have made age comparisons most often find at least slightly larger mean effects for juveniles (e.g., Dowden & Andrews 2000, Illescas et al. 2001, Landenberger & Lipsey 2005, Lösel & Schmucker 2005,

Petrosino 1997), although the reverse has been found in some meta-analyses of drug treatments (Mitchell et al. 2006, Wilson et al. 2006). Gender differences have been examined even less often, in part because most treatment studies use all male or nearly all male samples. Dowden & Andrews (1999) examined a small set of studies with female samples and found a mean recidivism effect size comparable to those found elsewhere for male samples. They also showed that Andrews et al.'s (1990) principles of effective treatment (risk, need, and responsivity) were associated with better outcomes for female offenders, as had been found with predominately male samples. Racial and ethnic differences have hardly been examined at all, although for juvenile offenders Wilson et al. (2003b) showed that mainstream treatments without cultural tailoring were as effective for minority youth as for white youth.

The most fully documented relationship between an offender characteristic and treatment effects is for the characteristic of reoffense risk. Among their principles of effective treatment, Andrews et al. (1990) argued that larger effects should be found for higher-risk offenders (their risk principle). Higher-risk offenders have a greater need for treatment and also have more room for improvement from effective treatment. Andrews and his colleagues have shown that there are indeed larger treatment effects for higher-risk cases for violent offenders (Dowden & Andrews 2000) and female offenders (Dowden & Andrews 1999). Similar differences have been shown in meta-analyses of community-based treatment for juveniles (Lipsey & Wilson 1998), treatment for sex offenders (Hall 1995, Reitzel & Carbonell 2006), and specific treatment types, such as cognitive-behavioral therapy (Landenberger & Lipsey 2005, Tong & Farrington 2006) and drug treatment (Lowenkamp et al. 2005). In one especially revealing research synthesis, Lowenkamp et al. (2006) analyzed recidivism effects for 97 correctional programs in Ohio that involved matched comparison

groups and detailed risk assessments for the participating offenders. They found larger effects for treatment groups with greater overall proportions of high-risk participants and for programs that provided more units of service or longer service to those among the participants who were at higher risk.

Specific treatment needs of offenders, such as substance-abuse problems, can also be conceptualized as risk factors. These represent dynamic risk factors—that is, malleable risk factors that, in principle, can be changed by effective treatment but that are predictive of subsequent offending (Andrews et al. 1990). As such, they contrast with the static risk factors that dominate most risk-assessment instruments and are not susceptible to change (e.g., characteristics of the offender's prior offense history). They also contrast with other treatment needs offenders may have that are not related to the likelihood of subsequent offending (not criminogenic), such as self-esteem. The need principle of Andrews et al. (1990) posits that treatment that addresses these criminogenic needs or dynamic risk factors has larger effects on recidivism. In several meta-analyses, Andrews and his colleagues have categorized treatments as targeting such needs or not and have shown that this distinction is related to recidivism effects (Dowden & Andrews 1999, 2000). As noted earlier, they have also shown that “appropriate” treatment that reflects all three of their effective-treatment principles (need, risk, and responsivity) produces much larger effects than treatments judged not appropriate by these principles (Andrews & Bonta 2006, Andrews et al. 1990, Cleland et al. 1997). Somewhat analogous analyses have shown the effectiveness of targeted treatment for offenders with specific problems, such as substance abusers and sex offenders (see **Table 4**).

Most research on rehabilitation treatments, however, is not specific about the needs the treatment is intended to address and rarely involves any explicit matching of treatment to needs. The judgments Andrews and colleagues must make to identify treatments that

meet their need principle are thus necessarily rather broad. Moreover, for most of the relevant needs of offenders, limited research is available to indicate which treatments are most effective in addressing any specific need or frequent combinations of such needs. For diagnostic purposes, needs-assessment instruments are generally less well developed than risk-assessment instruments and less widely used. The research currently available suggests generally that identifying criminogenic needs and addressing them with treatments especially effective for those needs produce relatively favorable recidivism outcomes. However, the research directly testing this proposition in all its ramifications is not sufficient for differentiating the needs most important to target or the treatments with the greatest impact on them.

The next generation of research. As the discussion above indicates, there are many questions about the sources of variability in the effects of rehabilitation treatments that have not been adequately addressed by the research available to date. What research is available now firmly establishes the general point that rehabilitation works. The most important task for the next generation of research is to address the questions of when, why, and for whom it works best. The outlines for the corresponding research agenda are relatively clear. We need research that provides more detailed descriptions of the nature of the treatment and treatment components provided, and the characteristics of the offenders who receive that treatment. The critical tasks for such research are determining what aspects of treatment most facilitate positive effects and what needs, risk factors, demographic profiles, and the like relate to differential responsiveness to treatment. There is particular utility for systematic information about the differential responsiveness of females and racial minorities and the relative benefits of treatments specially tailored for them (e.g., culturally sensitive). Also important is the identification of the pathways

through which treatment has its effects—for example, the mediating changes in needs, risk factors, cognitions, and motivation that bridge between treatment and recidivism effects. With such research should come better developed theories of change that can help explain the effects of current interventions and guide efforts to create better ones.

In addition, we need to know more about the dimensions of effective treatment implementation. There are strong indications in the extant research that the quality with which treatments are implemented is nearly as important as what treatments are implemented. Better information is needed, however, about the relative contributions of factors such as provider training, clinical supervision, and the monitoring of client participation to effective implementation. Research is also needed to clarify how best to conceptualize and measure treatment dosage (amount) and fidelity and their relationship to outcomes. Stated in research-design terms, the greatest need is not for more research on the main effects of treatment but, rather, for research on moderator and mediator relationships aimed at explaining differential effects.

CONCLUSION: CORRECTIONAL INTERVENTION AND PUBLIC SAFETY

This review of the research evidence about the effects of correctional interventions on offender reoffense rates, as with virtually every other such review in the past 30 years, falls under the long shadow of Lipton et al.'s (1975) review and Martinson's (1974) disparaging interpretation of the research evidence available at the time. During the intervening decades, hundreds of additional studies have been conducted, and techniques for systematically summarizing and analyzing the findings of intervention studies have advanced greatly. In particular, meta-analysis has developed as a way to conduct research reviews that makes the criteria for including and excluding studies explicit, represents study characteristics

systematically, captures the direction and magnitude of the empirical findings in a differentiated quantitative form, and allows for analysis of the distribution of those findings across studies with regard to the overall average and the factors related to their differences.

In this review we attempt to catalog every meta-analysis conducted on studies of correctional interventions and summarize the most general and robust of their collective findings. Some of these meta-analyses have broad scope, whereas others are narrow. Some are elaborate and some are relatively simple. Some are well done and a few are rather inept. Across this diversity, however, there is striking consistency on two key points. First, every meta-analysis of studies that compare recidivism outcomes for offenders receiving greater versus lesser or no sanctions has found, at best, modest mean recidivism reductions for the greater sanctions and, at worst, increased recidivism for that condition. Second, every meta-analysis of large samples of studies comparing offenders who receive rehabilitation treatment with those who do not has found lower mean recidivism for those in the treatment conditions. Moreover, the least of those mean reductions is greater than the largest mean reductions reported by any meta-analysis of sanctions. In addition, nearly all the meta-analyses of studies of specific rehabilitation treatments or approaches show mean recidivism reductions, and the great majority of those are greater than the largest reductions found in any meta-analysis of sanctions.

There are deficiencies in the underlying studies and the meta-analyses of those studies that could upwardly bias the statistical effect sizes that are at the heart of these findings. The main candidates are inflated effect estimates from poorly controlled quasi-experiments and overrepresentation of published studies, which often report larger effects than unpublished ones. Neither of these, however, is sufficient to account for the generally positive effects observed for rehabilitation treatment. Subsets of better

controlled studies also show positive effects, and the average differences between the findings of methodologically stronger and weaker studies do not consistently favor the weaker studies. Regarding publication bias, many meta-analyses include a large proportion of unpublished studies and, when separated out, the mean effect sizes for published studies are not consistently larger than for unpublished studies. In addition, any general bias of this sort would be expected to apply to studies of sanctions as well as to rehabilitation treatment and thus cannot easily account for the dramatic difference in their findings.

The preponderance of research evidence, therefore, supports the general conclusion that rehabilitation treatment is capable of reducing the reoffense rates of convicted offenders and that it has greater capability for doing so than correctional sanctions. The volume of research and the consistency of the findings of the systematic reviews make this a sufficiently sound general conclusion, bordering on beyond a reasonable doubt, to provide a basis for correctional practice and policy. The gap between this body of research and current practice and policy, however, is large and not easily bridged.

Research and Practice

The research reviewed here demonstrates that there are rehabilitation treatments with the potential to substantially reduce the recidivism of offenders in the correctional system and, in that way, reduce crime and enhance public safety. That does not mean, however, that the rehabilitation programs currently being used in correctional practice actually have those salutary effects. The increased punitive emphasis of recent decades has led to less rehabilitation programming, resulting in many offenders not being exposed to any significant treatment at all (Tewksbury et al. 2000). Moreover, the types of programs used in correctional practice are not the same mix represented in the research literature. Educational and vocational programs, for instance,

are common in correctional settings, with the latter often being no more than work assignments in a custodial setting. The treatments on which we have research, however, are more likely to have been developed from theory and prior research and to focus more directly on criminal behavior.

The rehabilitation treatments on which the available research is based also differ from correctional practice in another important way. Many of the research studies involve treatments that were developed by the researcher or delivered by the researcher, for example, with the researcher or developer selecting and training the personnel and monitoring the quality of service. Treatments provided in the context of such research and demonstration projects are not necessarily representative of typical correctional practice. Nor are their results representative—the recidivism effects for treatments in which the researcher is involved are larger than those for similar treatments without such involvement (Petrosino & Soydan 2005). In one meta-analytic comparison (Lipsey 1999), the mean effect size for research and demonstration programs was twice as large as that found in evaluations of routine practice programs in which the researcher had no role in design or implementation.

In short, the research on rehabilitation treatment reviewed here provides an encouraging indication of the relatively large effects that might be attainable in actual practice, but cannot be interpreted as evidence that current practice has such effects or, indeed, that it has any positive effects at all. We have too little systematic research on the nature of the rehabilitation programs that are actually in use in correctional practice to fully appraise the gap between research and practice, but there is every reason to believe it is enormous. The greatest obstacle to using rehabilitation treatment effectively to reduce criminal behavior is not a nothing-works research literature with nothing to offer but, rather, a correctional system that does not use the research available and has no history of doing so.

There are many aspects of rehabilitation treatment that are poorly understood and in need of additional research, as we note throughout this review, but the greatest challenge is the problem of technology transfer (Cullen & Gendreau 2000). This challenge is not unique to corrections. Even in the field of medicine, in which there are strong professional norms to base treatment on research evidence, the difficulty of influencing medical practice with the latest scientific knowledge has proven formidable. Nonetheless, the credibility of calls for effective correctional intervention depends on making concerted efforts to use evidence-based treatments (Cullen & Gendreau 2000, MacKenzie 2006).

There is much to be done on the research side of that exchange. We need a better understanding of how to package findings about effective treatment in ways that facilitate their dissemination and application in correctional settings. We especially need a better understanding of the constraints inherent in the organizational context of correctional programs and how to tailor evidence-based treatment to those contexts in ways that make them easy to adopt and, most especially, to implement well and sustain. On the other side of the exchange, it is essential that correctional systems attend to research evidence when making decisions about how much emphasis to place on rehabilitation treatment, which programs to implement, and how to implement them in ways that ensure they are effective. That will not happen spontaneously; it will require political and legislative action, such as the recent spate of state laws mandating the use of evidence-based practice (e.g., Washington, Oregon, North Carolina).

With regard to the potential for supportive political action, it is important to note that the American public is not antagonistic to offender rehabilitation. There is a widespread myth that the public harbors exclusively punitive sentiments in the domain of crime control. This view draws legitimacy from opinion polls showing that Americans endorse capital punishment, the use of “harsher courts,” and

prison sentences for many offenses (Cullen et al. 2000). But this is only half the story.

Survey evidence across three decades reveals that Americans also embrace efforts to intervene constructively with adult offenders and, in particular, with at-risk children and juvenile delinquents (Cullen 2006; Cullen et al. 2000, 2007). This research shows that upward of 8 in 10 Americans believe that rehabilitation is an important goal of adult corrections (Cullen et al. 2000). Support for treating youngsters is nearly universal; in one study, 97% stated that rehabilitation was an important goal of juvenile prisons (Cullen et al. 2007). Furthermore, in several stud-

ies in which respondents were asked whether the crime problem should be addressed by spending tax dollars on “early intervention programs” or on “building more prisons,” over three-fourths preferred expanded prevention efforts over the option of increasing imprisonment (Cullen et al. 2000, 2007). Taken together, these findings suggest that the American public favors a balanced approach to corrections that not only punishes but also tries to save the wayward. Indeed, clearly there is ample ideological room to implement rehabilitation programs that can be shown to improve the lives of offenders and, in so doing, enhance public safety.

DISCLOSURE STATEMENT

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

- Akers RL, Sellers CS. 2004. *Criminological Theories: Introduction, Evaluation, and Application*. Los Angeles: Roxbury. 4th ed.
- Andrews DA. 1995. The psychology of criminal conduct and effective treatment. In *What Works: Reducing Reoffending—Guidelines from Research and Practice*, ed. J McGuire, pp. 35–62. New York: Wiley
- Andrews DA, Bonta J. 2006. *The Psychology of Criminal Conduct*. Newark, NJ: Lexis/Nexis. 4th ed.
- Andrews DA, Dowden C. 2005. Managing correctional treatment for reduced recidivism: a meta-analytic review of programme integrity. *Br. Psychol. Soc.* 10:173–87
- Andrews DA, Zinger I, Hoge RD, Bonta J, Gendreau P, Cullen FT. 1990. Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology* 28:369–404
- Aos S, Phipps P, Barnoski R, Lieb R. 2001. *The Comparative Costs and Benefits of Programs to Reduce Crime*. Olympia: Wash. State Inst. Public Policy
- Blumstein A, Cohen J. 1973. A theory of the stability of punishment. *J. Crim. Law Criminol.* 64:198–206
- Cleland CM, Pearson FS, Lipton DS, Yee D. 1997. *Does age make a difference? A meta-analytic approach to reductions in criminal offending for juveniles and adults*. Presented at Annu. Meet. Am. Soc. Criminol., San Diego
- Cooper HM. 1998. *Synthesizing Research: A Guide for Literature Reviews*. Thousand Oaks, CA: Sage. 3rd ed.
- Cullen FT. 2006. It’s time to reaffirm rehabilitation: reaction essay. *Criminol. Public Policy* 5:665–72
- Cullen FT, Fisher BS, Applegate BK. 2000. Public opinion about punishment and corrections. In *Crime and Justice: A Review of Research*, ed. M. Tonry, pp. 1–79. Chicago: Univ. Chicago Press

- Cullen FT, Gendreau P. 2000. Assessing correctional rehabilitation: policy, practice, and prospects. In *Criminal Justice 2000: Policies, Processes, and Decisions of the Criminal Justice System*, ed. J Horney, 3:109–75. Washington, DC: U.S. Dept. Justice, Natl. Inst. Justice
- Cullen FT, Gilbert KE. 1982. *Reaffirming Rehabilitation*. Cincinnati: Anderson
- Cullen FT, Pratt TC, Micelli SL, Moon MM. 2002. Dangerous liaison? Rational choice theory as the basis for correctional intervention. In *Rational Choice and Criminal Behavior: Recent Research and Future Challenges*, ed. AR Piquero, SG Tibbetts, pp. 279–96. New York: Routledge
- Cullen FT, Vose BA, Jonson CL, Unnever JD. 2007. Public support for early intervention: Is child saving a “habit of the heart?” *Victims Offenders* 2:1–16
- Curtis NM, Ronan KR, Borduin CM. 2004. Multisystemic treatment: a meta-analysis of outcome studies. *J. Fam. Psychol.* 18:411–19
- Doob AN, Webster CM. 2003. Sentence severity and crime: accepting the null hypothesis. In *Crime and Justice: A Review of Research*, ed. M Tonry, 30:143–95. Chicago: Univ. Chicago Press
- Dowden C, Andrews DA. 1999. What works for female offenders: a meta-analytic review. *Crime Delinq.* 45:438–52
- Dowden C, Andrews DA. 2000. Effective correctional treatment and violent reoffending: a meta-analysis. *Can. J. Criminol.* 42:449–67
- Dowden C, Antonowicz D, Andrews DA. 2003. The effectiveness of relapse prevention with offenders: a meta-analysis. *Int. J. Offender Ther. Comp. Criminol.* 47:516–28
- Farabee D. 2006. Why the expansion of offender programs is premature. *J. Community Correct.* 16:9–10
- Feder L, Wilson DB. 2005. A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers’ behavior? *J. Exp. Criminol.* 1:239–62
- Gallagher CA, Wilson DB, Hirschfield P, Coggshall MB, MacKenzie DL. 1999. A quantitative review of the effects of sex offender treatment on sexual reoffending. *Correct. Manag. Q.* 3:19–29
- Garrett CJ. 1985. Effects of residential treatment on adjudicated delinquents: a meta-analysis. *J. Res. Crime Delinq.* 22:287–308
- Gendreau P. 1996. The principles of effective intervention with offenders. In *Choosing Correctional Interventions That Work: Defining the Demand and Evaluating the Supply*, ed. AT Harland, pp. 117–30. Newbury Park, CA: Sage
- Gendreau P, Goggin C, Smith P. 1999. The forgotten issue in effective correctional treatment: program implementation. *Int. J. Offender Ther. Comp. Criminol.* 43:180–87
- Gendreau P, Smith P, French SA. 2006. The theory of effective correctional intervention: empirical status and future directions. In *Taking Stock: The Status of Criminological Theory—Advances in Criminological Theory*, ed. FT Cullen, JP Wright, KR Blevins, 15:419–46. New Brunswick, NJ: Transaction
- Glaze LE, Bonczar TP. 2006. *Probation and Parole in the United States, 2005*. Washington, DC: Bureau Justice Stat., U.S. Dept. Justice
- Gottshalk R, Davidson WS II, Mayer J, Gensheimer GK. 1987. Behavioral approaches with juvenile offenders: a meta-analysis of long-term treatment efficacy. In *Behavioral Approaches to Crime and Delinquency: A Handbook of Application, Research, and Concepts*, ed. EK Morris, CJ Braukmann, 15:389–422. New York: Plenum
- Hall GCN. 1995. Sexual offender recidivism revisited: a meta-analysis of recent treatment studies. *J. Clin. Consult. Psychol.* 63:802–9

- Hanson RK, Gordon A, Harris AJR, Marques JK, Murphy W, et al. 2002. First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse* 14:169–94
- Hughes T, Wilson DJ. 2002. *Reentry Trends in the United States*. Washington, DC: Bureau Justice Stat., U.S. Dept. Justice
- Illescas SR, Sanchez-Meca JS, Genovés VG. 2001. Treatment of offenders and recidivism: assessment of the effectiveness of programmes applied in Europe. *Psychol. Spain* 5:47–62
- Landenberger NA, Lipsey MW. 2005. The positive effects of cognitive-behavioral programs for offenders: a meta-analysis of factors associated with effective treatment. *J. Exp. Criminol.* 1:451–76
- Latessa EJ, Cullen FT, Gendreau P. 2002. Beyond correctional quackery: professionalism and the possibility of effective treatment. *Fed. Probat.* 66(2):43–49
- Latimer J. 2001. A meta-analytic examination of youth delinquency, family treatment, and recidivism. *Can. J. Criminol.* 43:237–53
- Latimer J, Dowden C, Morton-Bourgon KE. 2003. Treating youth in conflict with the law: a new meta-analysis. *Rep. RR03Y7-3e*. Ottawa, Can.: Dept. Justice
- Levitt SD. 2002. Deterrence. In *Crime: Public Policies for Crime Control*, ed. JQ Wilson, J Petersilia, pp. 435–50. Oakland, CA: ICS Press
- Lipsey MW. 1999. Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Va. J. Soc. Policy Law* 6:611–41
- Lipsey MW, Wilson DB. 1998. Effective intervention for serious juvenile offenders. In *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions*, ed. R Loeber, DP Farrington, 13:313–45. Thousand Oaks, CA: Sage
- Lipton D, Martinson R, Wilks J. 1975. *The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies*. New York: Praeger
- Littell JH, Popa M, Forsythe B. 2005. *Multisystemic Therapy for Social, Emotional, and Behavioral Problems in Youth Aged 10–17*. Campbell Collab. Libr. <http://www.campbellcollaboration.org/frontend.asp>
- Lösel F. 1995. The efficacy of correctional treatment: a review and synthesis of meta-evaluations. In *What Works: Reducing Reoffending, Guidelines from Research and Practice*, ed. J McGuire, 4:79–111. New York: Wiley
- Lösel F, Schmucker M. 2005. The effectiveness of treatment for sexual offenders: a comprehensive meta-analysis. *J. Exp. Criminol.* 1:117–46
- Lowenkamp CT, Holsinger AM, Latessa EJ. 2005. Are drug courts effective: a meta-analytic review. *J. Community Correct.* 28:5–10
- Lowenkamp CT, Latessa EJ, Holsinger AM. 2006. The risk principle in action: What have we learned from 13676 offenders and 97 correctional programs? *Crime Delinq.* 52:77–93
- MacKenzie DL. 2001. Evidence-based corrections: identifying what works. *Crime Delinq.* 4:457–71
- MacKenzie DL. 2006. *What Works in Corrections? Reducing the Criminal Activities of Offenders and Delinquents*. New York: Cambridge Univ. Press
- MacKenzie DL, Wilson DB, Kider SB. 2001. Effects of correctional boot camps on offending. *Ann. Am. Acad. Polit. Soc. Sci.* 578:126–43
- Martinson R. 1974. What works? Questions and answers about prison reform. *Public Interest* 35:22–54
- Mauer M. 1999. *Race to Incarcerate*. New York: New Press
- McGuire J. 2002. Integrating findings from research reviews. In *Offender Rehabilitation and Treatment: Effective Programmes and Policies to Reduce Re-Offending*, ed. J. McGuire, pp. 3–38. West Sussex, UK: Wiley

- Mitchell O, Wilson DB, MacKenzie DL. 2006. *The Effectiveness of Incarceration-Based Drug Treatment on Criminal Behavior*. Campbell Collab. Libr. <http://www.campbellcollaboration.org/frontend.asp>
- Nieuwbeerta P, Nagin DS, Blokland AAJ. 2006. *The relationship between first imprisonment and criminal career development: a matched samples comparison*. Presented at 2nd Annu. Workshop Criminol. Econ., Queenstown, Maryland
- Nugent WR, Williams M, Umbreit MS. 2003. Participation in victim-offender mediation and the prevalence of subsequent delinquent behavior: a meta-analysis. *Utah Law Rev.* 137:137-66
- Patillo M, Weiman D, Western B, eds. 2004. *Imprisoning America: The Social Effects of Mass Incarceration*. New York: Russell Sage Found.
- Pearson FS, Lipton DS. 1999a. *The effectiveness of educational and vocational programs: CDATE meta-analyses*. Presented at Annu. Meet. Am. Soc. Criminol., Toronto
- Pearson FS, Lipton DS. 1999b. A meta-analytic review of the effectiveness of corrections-based treatments for drug abuse. *Prison J.* 79:384-410
- Pearson FS, Lipton DS, Cleland CM. 1997. *Rehabilitative programs in adult corrections: CDATE meta-analysis*. Presented at Annu. Meet. Am. Soc. Criminol., San Diego
- Pearson FS, Lipton DS, Cleland CM, Yee DS. 1998. *A review and meta-analysis of the effects of experiential challenge programs on recidivism*. Pres. Annu. Meet. Am. Soc. Criminol., Washington DC
- Pearson FS, Lipton DS, Cleland CM, Yee DS. 2002. The effects of behavioral/cognitive-behavioral programs on recidivism. *Crime Delinq.* 48:476-96
- Petersilia J, Turner S. 1993. Intensive probation and parole. In *Crime and Justice: A Review of Research*, ed. M Tonry, 17:281-335. Chicago: Univ. Chicago Press
- Petersilia J, Turner S, Peterson J. 1986. *Prison versus Probation in California: Implications for Crime and Offender Recidivism*. Santa Monica, CA: RAND
- Petrosino A. 1997. *'What works?' revisited again: a meta-analysis of randomized field experiments in rehabilitation, deterrence, and prevention*. PhD diss. Rutgers, State Univ. New Jersey, Newark
- Petrosino A, Soydan H. 2005. The impact of program developers as evaluators on criminal recidivism: results from meta-analyses of experimental and quasi-experimental research. *J. Exp. Criminol.* 1:435-50
- Petrosino A, Turpin-Petrosino C, Buehler J. 2003. Scared straight and other juvenile awareness programs for preventing juvenile delinquency: a systematic review of the randomized experimental evidence. *Ann. Am. Acad. Polit. Soc. Sci.* 589:41-62
- Reitzel LR, Carbonell JL. 2006. The effectiveness of sexual offender treatment for juveniles as measured by recidivism: a meta-analysis. *Sex Abuse* 18:401-21
- Rothstein H, Sutton AJ, Borenstein M, eds. 2005. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. New York: Wiley
- Sampson RJ, Laub JH. 1993. *Crime in the Making: Pathways and Turning Points Through Life*. Cambridge, MA: Harvard Univ. Press
- Sherman LW, Gottfredson DC, MacKenzie DL, Eck JE, Reuter P, Bushway SD. 1997. *Preventing Crime: What Works, What Doesn't, What's Promising*. Washington, DC: U.S. Dept. Justice, Natl. Inst. Justice
- Smith P. 2006. *The effects of incarceration on recidivism: a longitudinal examination of program participation and institutional adjustment in federally sentenced adult male offenders*. PhD thesis. Univ. New Brunswick. 121 pp.
- Smith P, Goggin C, Gendreau P. 2002. The effects of prison sentences and intermediate sanctions on recidivism: general effects and individual differences. *Rep. JS42-103/2002*. Ottawa, Can.: Solicitor Gen.

- Spohn C, Holleran D. 2002. The effect of imprisonment on recidivism rates of felony offenders: a focus on drug offenders. *Criminology* 40:329–57
- Tewksbury R, Erickson DJ, Taylor JM. 2000. Opportunities lost: the consequences of eliminating Pell grant eligibility for correctional education students. *J. Offender Rehab.* 31:43–56
- Tong LSJ, Farrington DP. 2006. How effective is the “Reasoning and Rehabilitation” programme in reducing reoffending? A meta-analysis of evaluations in four countries. *Psychol. Crime Law* 12:3–24
- Tonry M. 2004. *Thinking About Crime: Sense and Sensibility in American Penal Culture*. New York: Oxford Univ. Press
- Tonry M, Farrington DP, eds. 2005. *Crime and Punishment in Western Countries, 1980–1999—Crime and Justice: A Review of Research*, Vol. 33. Chicago: Univ. Chicago Press
- Travis J. 2005. *But They All Come Back: Facing the Challenges of Prisoner Reentry*. Washington, DC: Urban Inst. Press
- Villettaz P, Killias M, Zoder I. 2006. The effects of custodial vs noncustodial sentences on reoffending. A systematic review of the state of knowledge. Campbell Collab. Libr. <http://www.campbellcollaboration.org/frontend.asp>
- Visher CA, Winterfield L, Coggeshall MB. 2005. Ex-offender employment programs and recidivism: a meta-analysis. *J. Exp. Criminol.* 1:295–315
- Walters GD. 1990. *The Criminal Lifestyle: Patterns of Serious Criminal Conduct*. Thousand Oaks, CA: Sage
- Weisburd D, Lum CM, Petrosino A. 2001. Does research design affect study outcomes in criminal justice? *Ann. Am. Acad. Polit. Soc. Sci.* 578:50–70
- Whitehead JT, Lab SP. 1989. A meta-analysis of juvenile correctional treatment. *J. Res. Crime Delinq.* 26:276–95
- Wilson DB, Bouffard LA, MacKenzie DL. 2005a. A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Crim. Justice Behav.* 32:172–204
- Wilson DB, Gallagher CS, MacKenzie DL. 2000. A meta-analysis of corrections-based education, vocation, and work programs for adult offenders. *J. Res. Crime Delinq.* 37:347–68
- Wilson DB, Lipsey MW. 2001. The role of method in treatment effectiveness research: evidence from meta-analysis. *Psychol. Methods* 6:413–29
- Wilson DB, MacKenzie DL, Mitchell FG. 2005b. Effects of correctional boot camps on offending. Campbell Collab. Libr. <http://www.campbellcollaboration.org/frontend.asp>
- Wilson DB, Mitchell O, MacKenzie DL. 2006. A systematic review of drug court effects on recidivism. *J. Exp. Criminol.* 2:459–87
- Wilson SJ, Lipsey MW. 2000. Wilderness challenge programs for delinquent youth: a meta-analysis of outcome evaluations. *Eval. Prog. Plan.* 23:1–12
- Wilson SJ, Lipsey MW, Derzon JH. 2003a. The effects of school-based intervention programs on aggressive and disruptive behavior: a meta-analysis. *J. Consult. Clin. Psychol.* 71:136–49
- Wilson SJ, Lipsey MW, Soydan H. 2003b. Are mainstream programs for juvenile delinquency less effective with minority youth than majority youth? A meta-analysis of outcomes research. *Res. Soc. Work Pract.* 13:3–26
- Woolfenden SR, Williams K, Peat J. 2002. *Family and Parenting Interventions in Children and Adolescents with Conduct Disorder and Delinquency Aged 10–17*. Cochrane Libr. <http://www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME>



Contents

Frontispiece	
<i>Kitty Calavita</i>	x
Immigration Law, Race, and Identity	
<i>Kitty Calavita</i>	1
Accountability, Quantification, and Law	
<i>Wendy Nelson Espeland and Berit Irene Vannebo</i>	21
How Autonomous Is Law?	
<i>Christopher Tomlins</i>	45
Half Empty, Half Full, or Neither: Law, Inequality, and Social Change in Capitalist Democracies	
<i>Robin Stryker</i>	69
The Rule of Law	
<i>John K.M. Obnesorge</i>	99
Islamic Constitutionalism	
<i>Saïd Amir Arjomand</i>	115
The Emergence, Content, and Institutionalization of Hate Crime Law: How a Diverse Policy Community Produced a Modern Legal Fact	
<i>Valerie Jenness</i>	141
Restorative Justice: What Is It and Does It Work?	
<i>Carrie Menkel-Meadow</i>	161
Law and Collective Memory	
<i>Joachim J. Savelsberg and Ryan D. King</i>	189
Law and Lawyers Preparing the Holocaust	
<i>Michael Stolleis</i>	213
The Death of Socialist Law?	
<i>Inga Markovits</i>	233
Legal Innovation and the Control of Gang Behavior	
<i>Eva Rosen and Sudhir Venkatesh</i>	255

Punishment Beyond the Legal Offender <i>Megan Comfort</i>	271
The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews <i>Mark W. Lipsey and Francis T. Cullen</i>	297
The Socio-Legal Implications of the New Biotechnologies <i>Alain Pottage</i>	321
The Frontiers of Intellectual Property: Expanded Protection versus New Models of Open Science <i>Diana Rhoten and Walter W. Powell</i>	345
Personal Information, Borders, and the New Surveillance Studies <i>Gary T. Marx and Glenn W. Muschert</i>	375
Institutional Perspectives on Law, Work, and Family <i>Catherine Albiston</i>	397
Implicit Social Cognition and Law <i>Kristin A. Lane, Jerry Kang, and Mabzarin R. Banaji</i>	427

Indexes

Cumulative Index of Contributing Authors, Volumes 1–3	453
Cumulative Index of Chapter Titles, Volumes 1–3	455

Errata

An online log of corrections to *Annual Review of Law and Social Science* articles may be found at <http://lawsocsci.annualreviews.org>

January 2004

OUTCOME EVALUATION OF WASHINGTON STATE'S RESEARCH-BASED PROGRAMS FOR JUVENILE OFFENDERS

SECTION I: INTRODUCTION

In 1997, the Washington State Legislature passed the Community Juvenile Accountability Act (CJAA).¹ The primary goal of the CJAA is to reduce juvenile crime, cost effectively, by establishing "research-based" programs in the state's juvenile courts.² The basic idea is straightforward: taxpayers are better off if their dollars fund programs that have been proven to be effective in achieving key policy outcomes, in this case reduced re-offending.

Washington's effort is part of a nationwide trend to use research evidence to inform policy and program choices. The University of Colorado's Center for the Study and Prevention of Violence refers to research-based programs as "Blueprint Programs" when they meet strict scientific standards and have sufficient documentation to permit replication.³

The CJAA represents the nation's first statewide experiment of research-based programs for juvenile justice. Because the selected treatment programs had already been researched and found to be successful elsewhere in the United States, usually as small scale pilot projects, the question here was whether they work statewide in a "real world" setting. This report indicates that the answer to this question is yes—when the programs are competently delivered.

The specific research-based programs implemented in Washington were selected after the Washington State Institute for Public Policy (Institute) reviewed the national research literature.⁴

¹ RCW 13.40.500 - 540

² RCW 13.40.510

³ <www.colorado.edu/cspv>

⁴ S. Aos, P. Phipps, R. Barnoski, and R. Lieb, *The Comparative Costs and Benefits of Programs to Reduce Crime, Version 4.0* (Olympia: Washington State Institute for Public Policy, May 2001).

SUMMARY

In 1997, the Washington State Legislature passed the Community Juvenile Accountability Act (CJAA). The primary goal of the CJAA is to reduce juvenile crime, cost effectively, by establishing "research-based" programs in the state's juvenile courts. The basic idea is straightforward: taxpayers are better off if their dollars fund programs that have been proven to be effective in achieving key policy outcomes, in this case reduced re-offending.

The CJAA funded the nation's first statewide experiment concerning research-based programs for juvenile justice. Because selected treatment programs had already been researched elsewhere in the United States, usually as small scale pilot projects, the question here was whether they work when applied statewide in a "real world" setting. This report indicates that the answer to this question is yes—when the programs are competently delivered.

The basic findings are these:

1. When Functional Family Therapy (FFT) is delivered competently, the program reduces felony recidivism by 38 percent. The cost-benefit analyses find that FFT generates \$2.77 in savings (avoided crime costs) for each taxpayer dollar spent on the program, regardless of therapist competence. For competent FFT therapists, the savings are greater—\$10.69 in benefits for each taxpayer dollar spent.
2. When competently delivered, Aggression Replacement Training (ART) has positive outcomes with estimated reductions in 18-month felony recidivism of 24 percent and a benefit to cost ratio of \$11.66.
3. The Coordination of Services program achieved a decrease in 12-month felony recidivism, and the estimated benefit to cost ratio is \$7.89.
4. Because of problems implementing the Institute's evaluation design, no findings are associated with Multi-Systemic Therapy (MST). If the courts and the state wish to continue funding MST, the Institute recommends re-evaluating the program.

These findings affirm the merit of the legislature's investment in research-based programs for juvenile offenders. The next step is to implement the CJAA quality assurance standards so taxpayers can fully benefit from these programs.

Reports published by the Institute are available at www.wsipp.wa.gov. For further information, contact Robert Barnoski, (360) 586-2744, barney@wsipp.wa.gov; or Steve Aos (360) 586-2740, saos@wsipp.wa.gov

The following four CJAA programs were selected by Washington's 33 juvenile courts:

- *Functional Family Therapy (FFT)* was selected by 14 juvenile courts: Benton/Franklin, Grant, Grays Harbor, King, Kitsap, Klickitat, Lincoln, Pierce, Skagit, Snohomish, Spokane, Thurston, Whatcom, and Yakima;
- *Aggression Replacement Training (ART)* was selected by 26 courts: Adams, Asotin, Benton/Franklin, Chelan, Clallam, Clark, Cowlitz, Grant, Island, Jefferson, King, Kitsap, Kittitas, Lewis, Mason, Okanogan, Pacific/Wahkiakum, Pierce, Skamania, Snohomish, Spokane, Stevens, Thurston, Walla Walla, Whitman, and Yakima;
- *Coordination of Services (COS)* was selected by Snohomish Juvenile Court; and
- *Multi-Systemic Therapy (MST)* was selected by King, Kitsap, and Pierce Juvenile Courts.⁵

The Legislature directed the state's Juvenile Rehabilitation Administration (JRA) to oversee funding and quality adherence for the CJAA. In 1997, the Legislature also directed the Institute to determine whether the funded programs reduced recidivism.⁶ The juvenile courts and JRA formed the CJAA Committee for purposes of coordination and decision-making.

The evaluation relied on the following schedule:⁷

July 1998 State funding begins.
January 1999 CJAA program implementation.
July 1999 Program evaluation begins.
September 2000 Study samples include sufficient numbers of youth.
September 2002 Preliminary 12-month recidivism measurement period ends.
March 2003 Final 18-month recidivism measurement period ends.⁸
December 2003 Final report.

The CJAA specified that local juvenile courts target both diverted and adjudicated juvenile offenders for the programs and use a risk assessment to identify

⁵ These counties use the Juvenile Accountability Incentive Block Grant (JAIBG) funds, not CJAA funds, for MST.

⁶ RCW 13.40.500 – 540, Community Juvenile Accountability Act.

⁷ R. Barnoski, *The Community Juvenile Accountability Act: Program Evaluation Design* (Olympia: Washington State Institute for Public Policy, 1998).

⁸ The recidivism measurement period includes an 18-month follow-up period for re-offending and then a one-year period to allow for offenses to be adjudicated.

appropriate youth. The Institute worked with the Washington State Association of Juvenile Court Administrators to develop the Washington State Juvenile Court Assessment (WSJCA).⁹ This comprehensive assessment measures risk and protective factors identified by research as associated with juvenile criminality. The WSJCA classifies youth as low-, moderate-, or high-risk for re-offense. The WSJCA also produces a profile of risk measures for these domains: school, free-time, peers, family, mental health, aggression, anti-social attitudes, or social skills. The CJAA Committee determined that only moderate- to high-risk youth with a specific risk profile are considered for ART, FFT, and MST, while COS is for low-risk youth. Using the assessment to screen for program eligibility created a pool of youth across the courts with similar risk and protective factors who could potentially benefit from the program.

In 2002, two preliminary Institute reports¹⁰ found that FFT and ART appeared to reduce recidivism during a 12-month follow-up period. This final report contains 18-month follow-up data and supports the preliminary findings. The appendix to this report includes technical results and computations.¹¹

Overview of Findings

Exhibit 1 summarizes results for the four research-based programs.

- When FFT is delivered competently, the program reduces felony recidivism by 38 percent. The cost-benefit analyses find that FFT generates \$2.77 in savings (avoided crime costs) for each taxpayer dollar spent on the program, regardless of therapist competence. For competent FFT therapists, the savings are greater—\$10.69 in benefits for each taxpayer dollar spent.
- When competently delivered, ART has positive outcomes with estimated reductions in 18-month felony recidivism of 24 percent and a positive benefit to cost ratio of \$11.66.

⁹ R. Barnoski, *Washington State Juvenile Court Assessment Manual, Version 2.0* (Olympia: Washington State Institute for Public Policy, 1999).

¹⁰ R. Barnoski, *Washington State's Implementation of Functional Family Therapy for Juvenile Offenders: Preliminary Findings* (Olympia: Washington State Institute for Public Policy, 2002); R. Barnoski, *Washington State's Implementation of Aggression Replacement Training for Juvenile Offenders: Preliminary Findings* (Olympia: Washington State Institute for Public Policy, 2002).

¹¹ R. Barnoski, *Outcome Evaluation of Washington State's Research-Based Programs for Juvenile Offenders: Appendix* (Olympia: Washington State Institute for Public Policy, 2004)

- The COS program achieved a decrease in 12-month felony recidivism and a favorable estimated benefit to cost ratio of \$7.89.
- Because of problems implementing the Institute's evaluation design, no findings are associated with Multi-Systemic Therapy (MST). If the courts and the state wish to continue funding MST, the Institute recommends re-evaluating the program.

For these programs to achieve success, this evaluation found that the programs must be consistently delivered in a competent manner that follows the programs' specifications. In fact, the findings indicate that incompetent delivery may increase recidivism of participants. Without quality assurance efforts, the program may not only fail to reduce recidivism, it may actually increase recidivism.

The 2003 Washington State Legislature acted on the Institute's preliminary CJAA evaluation results¹² by directing the Institute to develop adherence and outcome standards for juvenile justice research-based programs.¹³ The subsequent Institute report¹⁴

includes guidelines for overseeing the delivery of programs and developing quality assurance measures. The CJAA statute requires JRA to submit annual reports to the legislature about the CJAA programs. The Institute's report recommends that JRA present measures of adherence to the standards in their annual reports. The Institute's recommended adherence standards include measures of competent program delivery, estimated recidivism reductions, and estimated returns from the state's investment in research-based programs. The legislation also states that courts shall not continue to use programs that do not comply with these standards.

The legislature took a calculated risk when it launched a policy to identify and fund research-based programs. Additionally, policymakers invested resources in a rigorous outcome evaluation to learn whether the programs are a cost-effective state investment. The gamble paid off; this evaluation found that using research-based programs can produce benefits to taxpayers in excess of their costs.

Exhibit 1
Summary of Outcome Evaluation Findings

PROGRAM	NUMBER OF YOUTH		ADJUSTED 18-MONTH FELONY RECIDIVISM ^A		REDUCTION IN RECIDIVISM	BENEFIT TO COST ^B (2002 DOLLARS)
	CONTROL	PROGRAM	CONTROL	PROGRAM		
Functional Family Therapy: Competent	313	181	27.0%	16.7%	-38.1%**	+\$10.69
Functional Family Therapy: Not Competent	313	206	27.0%	31.5%	+16.7%	-\$4.18
Functional Family Therapy: Total	313	387	27.0%	24.2%	-10.4%	+\$2.77
Aggression Replacement Training: Competent	417	501	24.8%	18.8%	-24.2%**	+\$11.66
Aggression Replacement Training: Not Competent	108	203	24.8%	26.5%	+6.9%	-\$3.10
Aggression Replacement Training: Total	525	704	24.8%	20.8%	-16.1%	+\$6.71
Coordination of Services ^C	171	171	3.3%	1.4%	-57.6%*	+\$7.89

^A Recidivism is defined as reconvictions in the Washington State court system. The rates shown are adjusted to account for systematic differences between the program and control groups using means in the equations from the logistic regressions.

^B To be conservative, the benefit-cost ratios are based on reduced estimates of program effects to account for the less-than-random-assignment research designs. The FFT effect size was reduced 25 percent, ART 50 percent, and COS 50 percent. The estimated cost per youth is \$2,100 for FFT, \$745 for ART, and \$400 for COS.

^C Adjusted 12-month felony recidivism rate.

* Statistically significant reduction in recidivism at the .15 level.

** Statistically significant reduction in recidivism at the .05 level.

¹² Barnoski, *Washington State's Implementation of Functional Family Therapy for Juvenile Offenders*; Barnoski, *Washington State's Implementation of Aggression Replacement Training for Juvenile Offenders*.

¹³ RCW 13.40.530

¹⁴ R. Barnoski, S. Aos, R. Lieb, *Recommended Quality Control Standards: Washington State Research-Based Juvenile Offender Programs* (Olympia: Washington State Institute for Public Policy, December 2003).

SECTION II: EVALUATION DESIGN

The 1997 Washington State Legislature directed the Institute to determine whether the programs funded by the CJAA reduce recidivism. The best way to answer this question is to compare the recidivism rates of eligible youth randomly assigned to either the control or the program group.¹⁵ Any outcome differences between the two groups can then be attributed to the program. Since this approach was not seen as feasible by all juvenile courts, a pseudo-random assignment process was used. For the CJAA evaluation, control groups of juvenile offenders who did not receive a CJAA program were selected using the “waiting line” approach. This method takes advantage of the fact that CJAA resources were not sufficient to allow every eligible youth to enter a CJAA program.

In the waiting line approach, all juvenile offenders are assessed by court staff using the Washington State Juvenile Court Assessment (WSJCA).¹⁶ The WSJCA was specifically developed by the Institute and the juvenile courts for the CJAA because the enabling legislation required youth be screened for program eligibility and an assessment be used to determine the programs most likely to change behaviors of juvenile offenders.

The WSJCA involves a two-stage process. First, all adjudicated youth are assessed with a pre-screen instrument that determines the youth’s level of risk. The level of risk is determined by the pre-screen criminal history and social history risk scores. Second, only the moderate- to high-risk youth are assessed with the full instrument to determine their risk profile.

The full assessment is organized into nine domains: school, free-time, employment, relationships, family (current and prior), drug/alcohol, mental health, anti-social attitudes, and skills. For each domain, a risk or protective factor score is computed. Another score was developed to measure aggression.

The validity of both the pre-screen and full WSJCA is supported by an Institute study.¹⁷ The eligibility criteria developed by the CJAA Committee for the four treatment programs are displayed in Exhibit 2.

¹⁵ R. Barnoski, *Standards for Improving Research Effectiveness in Adult and Juvenile Justice* (Olympia: Washington State Institute for Public Policy, December 1997).

¹⁶ Barnoski, *Washington State Juvenile Court Assessment Manual*.

¹⁷ R. Barnoski, *Assessing Risk for Re-Offense: Validating the Washington State Juvenile Court Assessment* (Olympia: Washington State Institute for Public Policy, forthcoming).

These criteria match the youth’s risk profile to the program that addresses those risk factors.

Exhibit 2
CJAA Program Eligibility Criteria

CJAA PROGRAM	ELIGIBILITY CRITERIA
Coordination of Services	Low-risk.
Aggression Replacement Training	Moderate- or high-risk, and: a score of at least one for a weapon, violent misdemeanor, or felony conviction; or a dynamic risk factor score of at least 2 out of 13 on aggression; or a dynamic risk factor score of at least 7 out of 28 on attitudes/behavior; or a dynamic risk factor score of at least 9 out of 36 on skills.
Functional Family Therapy	Moderate- or high-risk and a dynamic risk factor score of at least 6 out of 24 on current family.
Multi-Systemic Therapy	High-risk and a dynamic risk factor score of at least 6 out of 24 on current family.

Youth who met the selection criteria and had a sufficient period of time on supervision to complete the program were assigned by court staff to the appropriate CJAA program.¹⁸ When the program reached capacity (all therapists had full caseloads or sessions were full), the remaining eligible youth were assigned by court staff to the control group and never participated in the program; instead, they received the usual juvenile court services. The assignment process started in July 1999, and sufficient sample sizes were attained by September 2000.

The procedures for this assignment process varied from court to court. In some courts, the assignment of youth was random (using the last digit of their juvenile number), in some courts it occurred on a first-come, first-served basis, while in others, the courts exercised some discretion in group assignments.

Discussions with court staff in some counties indicated that youth viewed as most in need of services may have received preferential assignment to the program groups. Because of this potential bias in the assignment process, the evaluation’s analyses use multivariate statistical techniques to control for systemic differences between the program and control groups on key characteristics

¹⁸ Some exceptions were created for youth with mental health and acute drug/alcohol problems that would prevent participation in the program.

from the WSJCA (gender, age, and domain risk and protective factor scores). From these analyses, mean-adjusted recidivism rates are calculated. These adjusted rates provide estimates of the impact of the program which are not confounded by systematic differences between the groups.¹⁹

The evaluation design incorporated a time period for service providers to learn the treatment program before youth were included in the outcome evaluation. For the FFT and MST interventions, only youth whose service provider had at least 90 days of supervised experience were included in the study. Because the Institute did not have access to the identities of ART instructors for each class, it was not possible to follow this procedure. As a remedy, ART participants during the first year of implementation are excluded from the study.

To measure recidivism, the Institute follows the definition for recidivism established by the 1997 Legislature.²⁰ Recidivism is measured using conviction rates for subsequent juvenile or adult offenses. In Washington, all convictions in juvenile and adult criminal courts are recorded in statewide databases maintained by the state's Administrative Office of the Courts and the Department of Corrections. Three reconviction rates are reported:

- Total misdemeanor and felony convictions;
- Felony convictions; and
- Violent felony convictions.

The follow-up "at-risk" period for each youth is 18 months.²¹ In calculating rates, the Institute allows a 12-month period for an offense to be adjudicated by the courts.

This research design provides a strong means to test whether the CJAA programs lowered recidivism rates. As previously mentioned, this is not a perfect random assignment research design, because the treatment and control groups may differ for reasons other than CJAA program participation. Fortunately, the WSJCA data allow for rigorous statistical modeling to control for potential pre-existing differences.

SECTION III: FUNCTIONAL FAMILY THERAPY

What Is Functional Family Therapy? Functional Family Therapy (FFT) is a structured family-based intervention that works to enhance protective factors and reduce risk factors in the family. FFT is a three-phase program. The first phase is designed to motivate the family toward change. The second phase teaches the family how to change a specific critical problem identified in the first phase. The final phase helps the family generalize their problem-solving skills.²² FFT has been identified by the University of Colorado's Center for the Study and Prevention of Violence as a Blueprint Program.²³

Trained FFT therapists have caseloads of 10 to 12 families, and the intervention involves about 12 visits during a 90-day period. Between January 1999 and September 2001, 14 of Washington's 34 juvenile courts implemented FFT, and approximately 400 families and 40 therapists participated in the program. Court staff use the WSJCA to assess whether youth are eligible for FFT: a youth must have at least a moderate-risk level with family problems indicated by a family dynamic risk factor score above the eligibility cut-off value (6 out of 24 points).

The average cost of FFT reported by JRA is \$2,100 per family. Some juvenile courts trained their own staff as therapists, some courts hired therapists, while other courts contracted with private therapists. FFT, Inc., now based in Seattle, trains and supervises the clinical practices of FFT therapists.

The question for this study is whether FFT works in a setting where FFT, Inc. is *not* directly involved with the families. That is, can FFT be implemented by 14 independent juvenile courts with sufficient consistency and program fidelity to reduce recidivism and make the \$2,100 cost per program participant a wise use of taxpayer dollars?

¹⁹ These calculations use the means of the WSJCA factors of the total sample for both the program and control groups in determining the adjusted rate. Barnoski, *Outcome Evaluation Appendix*.

²⁰ Barnoski, *Standards for Improving Research Effectiveness*.

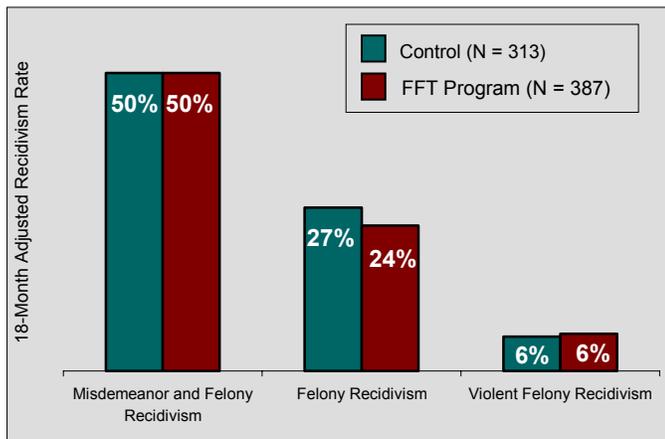
²¹ One CJAA program, Coordination of Services, was not implemented until 2001, and, therefore, only a 12-month follow-up period could be measured.

²² For information about Functional Family Therapy, see <www.fftinc.com>.

²³ Panels of experts have determined that Blueprint Programs meet a standard of scientific evidence which provides a high degree of confidence that the programs can achieve their objectives. See <www.colorado.edu/cspv/blueprints>.

FFT Results: Exhibit 3 shows the three adjusted 18-month recidivism rates for youth in the control group versus all youth receiving FFT, regardless of therapist competence.²⁴ For example, the adjusted 18-month felony recidivism rate for the control group is 27 percent compared with 24 percent for the FFT group. There are no statistically significant differences for the three types of recidivism. Does this mean that, contrary to the national FFT findings, FFT in Washington State does not reduce recidivism? The next section takes a look “under the hood” to better understand these results.

Exhibit 3
Adjusted 18-Month Recidivism Rates
FFT vs. Control Group



No statistically significant differences.

Therapist Adherence to FFT: Although the courts hire or contract with the therapists, JRA and FFT, Inc. manage the quality assurance process for the FFT therapists in Washington State. State funding was used to assign a qualified JRA staff person with a master’s degree in counseling (Dana Phelps) to receive FFT, Inc. training and help manage FFT delivery. Ms. Phelps assisted FFT with training, therapists’ consultations, and corrective actions throughout the state. As a result, she became very familiar with all the state’s FFT therapists.

Because Washington’s experience was the first statewide implementation of FFT in the nation, the process of program management on a large scale was developed as the program was implemented. That is, the therapists were learning FFT, and the state and FFT, Inc. were learning how to train and manage a large number of therapists. FFT, Inc.’s

²⁴ The multivariate statistical analyses use data from the WSJCA (gender, age, criminal history, social history, and other risk and protective factors) to control for systemic differences between the program and control groups. The calculations for the adjusted recidivism rates from the multivariate logistic regression are given in *Outcome Evaluation Appendix*, Exhibit A-1.

computer information system for recording data about therapist competence was completed after the evaluation was underway. Therefore, the therapist ratings used for this evaluation were based on Ms. Phelps’ recollections, combined with those of the FFT, Inc. consultants, rather than “real time” measurement. The ratings were obtained before Ms. Phelps knew any of the study outcomes.

Despite the imprecise rating of therapists during the study period, the preliminary FFT findings²⁵ demonstrated that the group of FFT therapists rated as competent had reduced the 12-month felony recidivism rates of youth (p=.08). In addition, the preliminary results showed that the group of therapists who were not competent may have *increased* the felony recidivism rates of youth. Since the ratings created valid distinctions among therapists, the ratings continued to be used.

Exhibit 4 displays, for each therapist group, the number of therapists during the study period with a minimum 90 days of supervised experience delivering FFT. The exhibit also includes the number of families seen by these therapists. Therapists judged as highly competent and competent are combined into a total competent group, and those rated as either not competent or borderline competent are combined into a total not competent group. Together, 48.4 percent (16) of the 33 therapists are rated by FFT, Inc. and JRA as competent or highly competent; these therapists treated 46.8 percent of the families in the study.

Exhibit 4
FFT Therapist Competence Ratings

FFT THERAPIST GROUPS	THERAPISTS		FAMILIES	
	Number	Percent-age	Number	Percent-age
Not Competent	11	33.3	118	30.5
Borderline	6	18.2	88	22.7
Total Not Competent	17	51.5	206	53.2
Competent	8	24.2	103	26.6
Highly Competent	8	24.2	78	20.2
Total Competent	16	48.4	181	46.8
Total	33	100.0	387	100.0

Note: Four therapists are excluded because their competence was not known by the raters.

Exhibit 5 compares key characteristics of youth in the three study groups. These characteristics, based on the WSJCA, include age and gender, the

²⁵ Barnoski, *Washington State’s Implementation of Functional Family Therapy for Juvenile Offenders*.

two pre-screen risk scores, as well as the full assessment domain scores.

Exhibit 5
Comparison of WSJCA Characteristics for Control and FFT Groups

VARIABLE	CONTROL GROUP	YOUTH SEEN BY COMPETENT THERAPISTS	YOUTH SEEN BY NOT COMPETENT THERAPISTS
Number of Youth	313	181	206
Male Gender ^A	80%	81%	75%
Age ^{AB}			
13	10%	18%	11%
14	16%	19%	18%
15	21%	25%	23%
16	24%	20%	26%
17	29%	18%	21%
Average Age ^{AB}	15.5	15.0	15.3
Pre-Screen Average Risk Scores			
Criminal History ^A	8.0	7.7	7.1
Social History	9.0	9.3	9.1
Full Assessment Average Domain Risk Scores			
Aggression	2.2	2.4	2.3
Attitude ^{AB}	8.5	9.5	8.5
Drug/Alcohol	5.4	5.4	5.3
Employment (Protective) ^B	1.4	1.0	1.1
Family	14.1	14.3	13.6
Free-Time	2.0	2.0	1.9
Mental Health ^A	2.1	2.3	1.9
Prior Family ^A	15.3	15.9	15.0
Relationship ^A	10.9	10.3	12.6
School	11.5	12.5	12.7
Skill	18.8	19.5	18.7

^A Statistically significant difference between youth seen by therapists rated competent versus those seen by therapists not competent.

^B Statistically significant difference between youth seen by competent therapists versus those in the control group.

Statistically significant differences were found between the study groups on several characteristics.

Competent Therapists Versus Control Group:

For youth seen by competent therapists versus those in the control group, statistically significant differences exist for these variables: age, attitude, and employment. These differences indicate that the youth seen by competent therapists are slightly higher risk than youth in the control group (p<.05).

Competent Versus Not Competent Therapists:

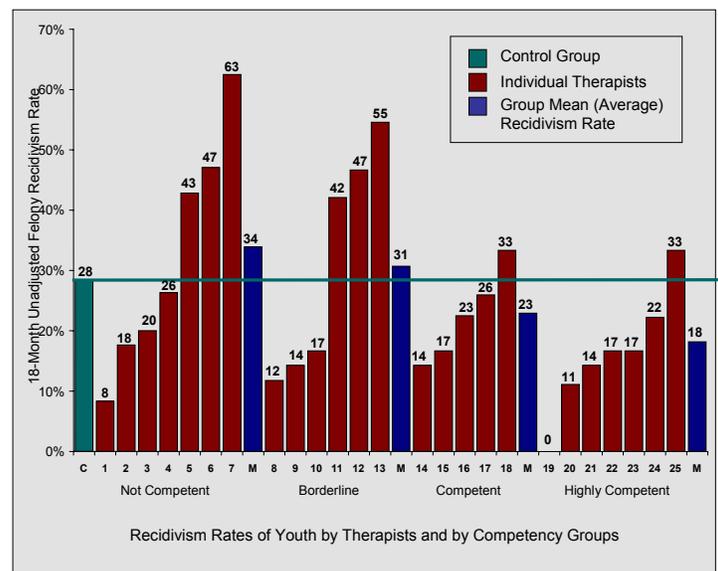
Comparing youth seen by competent FFT therapists with those seen by therapists who are not competent, the following characteristics are significantly different: gender, age, criminal history, attitude, employment

(protective), mental health, prior family, and relationships. With the exception of relationships, the competent therapists saw youth whose characteristics indicate a higher risk to re-offend.

These findings may indicate two flaws in the assignment process: youth viewed as most in need of services may have received preferential assignment to FFT rather than the control group, and the higher-risk youth may have received preferential assignment to the better therapists. Multivariate statistical analyses were used to compensate for these differences; the findings are as follows.

Therapists' FFT Competence and Recidivism Outcomes: Exhibit 6 shows the felony recidivism rates for youth grouped by their individual therapist's competence rating. The mean (average) recidivism rates for each therapist group and the control group are also included. The results are for the 25 therapists who saw at least six youth.

Exhibit 6
18-Month Felony Recidivism Rates for Youth Assigned to Individual FFT Therapists



The exhibit shows that the youth in the competent and highly competent therapist groups have lower average felony recidivism rates than the youth in either the control group or the not competent or borderline competent therapist groups. These results occurred even though the competent and highly competent therapists were assigned, on average, slightly higher-risk youth. Exhibit 6 also shows that within each group of therapists, the recidivism rates vary considerably. In particular, the youth treated by five therapists judged as not competent or borderline competent have low recidivism rates (therapists 1, 2, 8, 9, and 10).

Conversely, the youth seen by two therapists judged as competent or highly competent have high recidivism rates (therapists 18 and 25). One possible explanation for these results is that some therapists may be misclassified.

To determine the relationship between therapist competence and recidivism, competence is included as an additional variable in the multivariate analysis. Exhibit 7 compares the resulting 18-month adjusted recidivism rates for three study groups. Exhibit 8 presents the same data by the more detailed rating of therapist competence.²⁶

Youth seen by the competent therapists have an 18 percent felony recidivism rate compared with 27 percent for the control group, a statistically significant reduction of 38 percent. For violent felony recidivism, the competent therapist group has a 3 percent rate compared with 6 percent for the control group, a 50 percent reduction that is statistically significant at the $p=.115$ probability level.

Exhibit 7
Adjusted 18-Month Recidivism Rates: Control vs. Not Competent and Competent FFT Therapist Groups

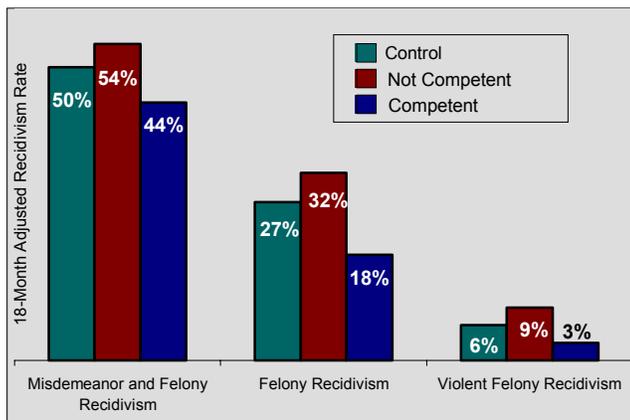


Exhibit 8
Adjusted 18-Month Recidivism Rates Control vs. FFT Therapist Groups

STUDY GROUP	MISDEMEANOR AND FELONY	FELONY	VIOLENT FELONY
Control	49.6%	27.0%	5.5%
Not Competent	51.2%	32.8%	10.7%
Borderline	58.3%	29.9%	7.8%
Total Not Competent	54.3%	31.5%	9.5%*
Competent	49.1%	17.6%*	3.1%
Highly Competent	37.3%	15.3%*	2.4%
Total Competent	44.1%	16.7%*	2.8%
All FFT Youth	49.6%	24.2%	6.2%

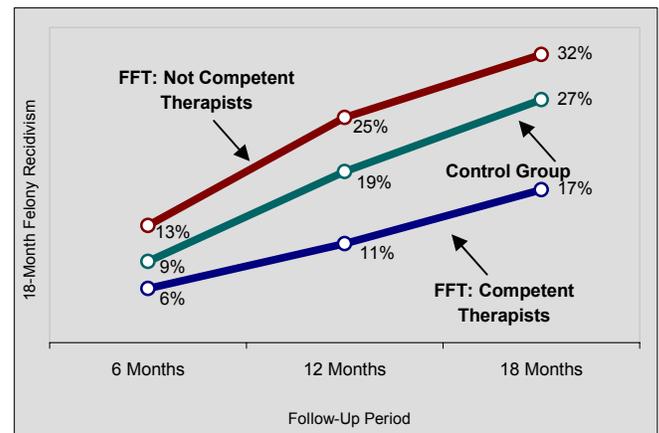
*Statistically significant at the .05 probability level.

²⁶ The calculations for the adjusted recidivism rates from the multivariate logistic regression are given in *Outcome Evaluation Appendix*, Exhibit A-2.

The results shown in Exhibits 6 through 8 illustrate the critical role of FFT therapist competence. This finding is especially significant, because recidivism may be exacerbated by therapists who do not competently follow the model.

The next step in examining FFT effectiveness is to see how well the reductions in recidivism by competent therapists hold up over time. For this sample, we examine 6-month, 12-month, and 18-month adjusted felony recidivism rates.²⁷ Exhibit 9 compares these adjusted rates for the three study groups over time. The reduction in felony recidivism between the control and competent therapist groups at 12 months is 40 percent compared with 38 percent at 18 months, indicating that FFT's suppression effect on felony recidivism is relatively constant.

Exhibit 9
Adjusted Felony Recidivism Rates at 6-, 12-, and 18-Month Follow-up Periods



FFT Cost-Benefit Analysis: The cost-benefit analysis, described in Section VII, determines whether Washington citizens receive a positive return on their dollars spent on FFT. When FFT is delivered by competent therapists, it generates \$10.69 in benefits (avoided crime costs) for each dollar spent on the program. When not competently delivered, FFT costs the taxpayer \$4.18. Averaging these results for all youth receiving FFT, regardless of therapist competence, results in a net savings of \$2.77 per dollar of costs.

FFT Conclusions: When the FFT model is delivered competently, the program reduces felony and violent felony recidivism cost effectively.

²⁷ The Institute will continue tracking the recidivism of these groups to determine if the FFT effect is sustained over longer follow-up periods. The calculations for the adjusted recidivism rates from the multivariate logistic regression are given in *Outcome Evaluation Appendix*, Exhibit A-3.

SECTION IV: AGGRESSION REPLACEMENT TRAINING

What Is Aggression Replacement Training?

Aggression Replacement Training (ART) is a 10-week, 30-hour intervention administered to groups of 8 to 12 juvenile offenders three times per week. The program relies on repetitive learning techniques to teach participants to control impulsiveness and anger and use more appropriate behaviors. In addition, guided group discussion is used to correct anti-social thinking. Although ART does not meet the strict scientific standards required to be a Blueprint Program by the Center for the Study and Prevention of Violence, three research studies support the effectiveness of ART in reducing recidivism.²⁸

The CJAA Committee decided that CJAA funds could be used for ART when court probation staff or private contractors received Washington State ART training. The cost for ART in Washington State is approximately \$745 per youth.

The CJAA Committee established the eligibility criteria for ART. Eligible youth must have at least a moderate risk level. In addition, the youth must have a problem with aggression, pro-social attitudes, or pro-social skills as indicated by relevant scores on the WSJCA scales.²⁹

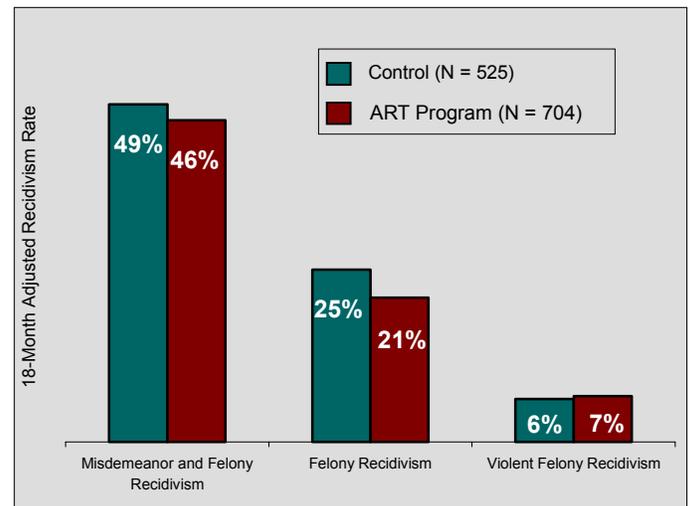
ART was the most widely implemented CJAA program, with 26 juvenile courts participating and more than 100 instructors. During the first year, courts were sending new instructors to training, replacing existing instructors, and changing instructional teams. Information identifying individual ART instructors was not recorded by the courts, so it was not possible to know the level of instructor expertise for individual youth.

Because of this flux in instructors during the first year, questions emerged about the quality of the program's delivery during 1999, the first year of implementation. A multivariate analysis of 18-month felony recidivism³⁰ revealed that, compared with control group youth, youth receiving ART during 2000 had significantly better results than

youth receiving ART during 1999. To allow for the courts to gain sufficient experience and stability in the delivery of ART, this study excludes youth assigned to ART and the control groups during 1999 and only includes youth assigned during 2000.

ART Results: Exhibit 10 shows the three adjusted recidivism rates of youth in the control group versus the ART group for 2000.³¹ The 18-month adjusted felony recidivism rate for the control group is 25 percent compared with 21 percent for ART (a 16 percent reduction in felony recidivism rates). The finding for felony recidivism is statistically significant at the $p=.125$ probability level. There are no statistically significant differences in misdemeanor and felony recidivism and violent felony recidivism rates. As with FFT, we now examine how competent delivery affects these results.

Exhibit 10
Adjusted 18-Month Recidivism Rates
Control vs. ART Groups During 2000



Instructional Team Adherence to ART: Unlike Functional Family Therapy and Multi-Systemic Therapy, no national organization provides training and consultation for ART. Although Barry Glick, an expert from New York State, provided the initial training in Washington State, the juvenile courts and JRA had to develop the quality assurance capacity for this program. Fortunately, the state already had a well respected expert in ART, Chris Hayes from Snohomish County Juvenile Court. Mr. Hayes worked with JRA on a half-time basis to train CJAA-funded ART instructors, establish a quality assurance process and a training curriculum, as well as a procedures manual.

²⁸ Aos, et al., *The Comparative Costs and Benefits of Programs to Reduce Crime*.

²⁹ A score of at least one for a weapon, violent misdemeanor, or felony conviction or a dynamic risk factor score of at least 2 out of 13 on aggression; dynamic risk factor score of at least 7 out of 28 on attitudes/behavior or a dynamic risk factor score of at least 9 out of 36 on skills.

³⁰ Logistic regression was used with an interaction term accounting for the study year and study group (ART vs. control). The interaction term was statistically significant ($p<.07$) and indicated better outcomes in the year 2000.

³¹ The calculations for the adjusted recidivism rates from the logistic models are given in *Outcome Evaluation Appendix*, Exhibit B-1.

When analyzing data for the Institute's preliminary report, we found the effectiveness of ART in reducing recidivism varied from court to court.³² In response, the Institute asked Mr. Hayes to rate various attributes of ART delivery in each court. Because he was not able to observe every instructional team, Mr. Hayes could only provide information for each court as a whole. The ratings would have been more accurate if they were applied to each instructional team. Despite this shortcoming, the preliminary report found that the courts judged to be competently delivering ART had significantly reduced 12-month felony recidivism ($p=.05$). Mr. Hayes' ratings are used in this report.

In addition, Mr. Hayes identified two courts that consistently delivered ART with the highest degree of fidelity to the model: Okanogan and Pierce. The ratings of competent and highly competent ART courts are comparable to the ratings of competent and highly competent FFT therapists.

Exhibit 11 presents the number of courts and youth involved in the ART evaluation during 2000. Five courts were rated as not delivering ART competently; 108 youth were in the control group and 203 in ART. Twenty-one courts were judged as delivering ART competently to 501 youth. The two highly competent courts provided ART to 99 youth.

Exhibit 11
ART Evaluation Study Groups in 2000

ART GROUP	NUMBER OF COURTS	NUMBER OF YOUTH		
		Control	ART	Total
Not Competent	5	108	203	311
Competent	19	299	402	701
Highly Competent	2	118	99	217
Total Competent	21	417	501	918
Total	26	525	704	1,229

The characteristics of the control and ART groups in the year 2000 are compared in Exhibit 12.

Exhibit 12
Comparison of Characteristics Between Control Group and ART Groups in 2000 for Competent and Not Competent Delivery of ART Courts

VARIABLE	ALL ART COURTS		COMPETENT ART DELIVERY		NOT COMPETENT ART DELIVERY	
	CONTROL	ART	CONTROL	ART	CONTROL	ART
	Number of Youth	525	704	417	501	108
Male	81%	80%	81%	81%	81%	79%
Age at Adjudication	15.5	15.2**	15.4	15.1**	15.6	15.4
Criminal History	8.1	8.3	7.7	7.9	9.5	9.4
Social History	8.6	8.1**	8.6	8.2*	8.4	7.9
Aggression	2.1	2.3	2.1	2.2	2.1	2.3
Drug/Alcohol	5.2	4.5**	5.1	4.5**	5.5	4.5*
Employment (Protective)	1.4	1.0**	1.4	1.0**	1.5	0.9**
Family	9.0	9.1	9.3	9.7	8.0	7.4
Free-Time	1.7	1.6	1.7	1.6	1.9	1.7
Mental Health	2.2	2.1	2.3	2.2	1.9	1.8
Prior Family	12.8	12.1	13.4	12.8	10.7	10.3
Relationship	9.8	9.1	10.0	9.6	8.8	7.8
School	11.5	10.6**	11.7	10.8*	10.9	10.1
Skill	17.8	17.7	18.4	18.8	15.2	14.9
Attitude	7.5	7.5	7.7	7.8	6.6	6.7

* Statistically significant difference at the .05 probability level.
**Statistically significant difference at the .01 probability level.

All ART Courts: For ART and control group youth in all courts, five variables have statistically significant differences between the groups: age, social history risk, drug/alcohol risk, employment (protective), and school risk. For example, the average age of ART youth is 15.2, while the average age of control group youth is 15.5. Lower age indicates increased risk.

Competent ART Delivery: For the courts judged competent, significant differences exist between the control and ART group youth on five variables. The competent ART group has lower risk scores than the control group on social history, drug/alcohol, and school risk, but a lower protective factor score for employment. The average age of ART youth is 15.1, while the average age of control group youth is 15.4.

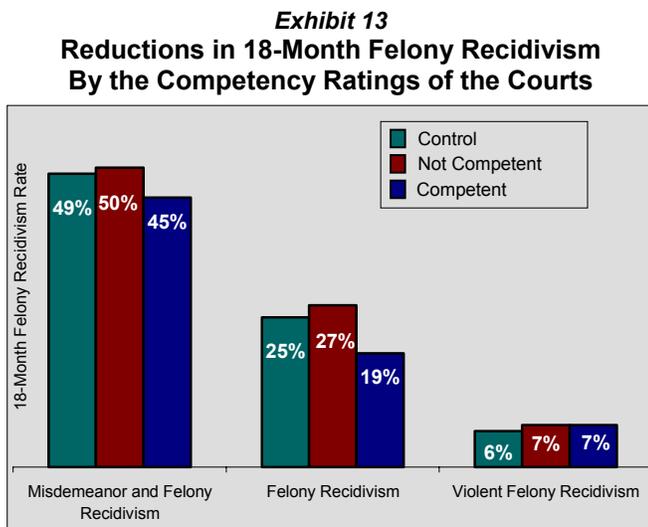
Not Competent ART Delivery: For the courts judged not competent, statistically significant differences also exist between the ART and control groups; in this case for two variables: drug/alcohol risk and employment. The ART group has a lower drug/alcohol risk but a lower protective factor score for employment.

³² Barnoski, *Washington State's Implementation of Aggression Replacement Training for Juvenile Offenders*.

In conclusion, there are some differences between the youth in ART and those in the control group. Multivariate statistical analyses are, therefore, used next to adjust for these systematic differences.

ART Court Competency Ratings and Felony Recidivism: Exhibit 13 displays the adjusted felony recidivism rates by court competency ratings. Exhibit 14 presents the same data by the more detailed rating of competence.³³ The exhibits illustrated these findings:

- For the five courts rated as not competent, the adjusted 18-month felony recidivism rate is 27 percent compared with 25 percent for the control group. This difference is not statistically significant.
- For the 21 courts rated as either competent or highly competent, the 18-month felony recidivism rate is 19 percent. This is a 24 percent reduction in felony recidivism compared with the control group, which is statistically significant.
- The two highly competent courts have statistically significant reductions in both misdemeanor and felony recidivism and felony recidivism, but not violent felony recidivism.



These findings are similar to those in the preliminary report which were based on 12-month recidivism rates and included youth in the study during 1999. The competency ratings continue to influence the results for ART on felony recidivism during its second year. The next step is to see how well these results hold up over time.

³³ The calculations for the adjusted recidivism rates from the logistic models are given in *Outcome Evaluation Appendix*, Exhibit B-2.

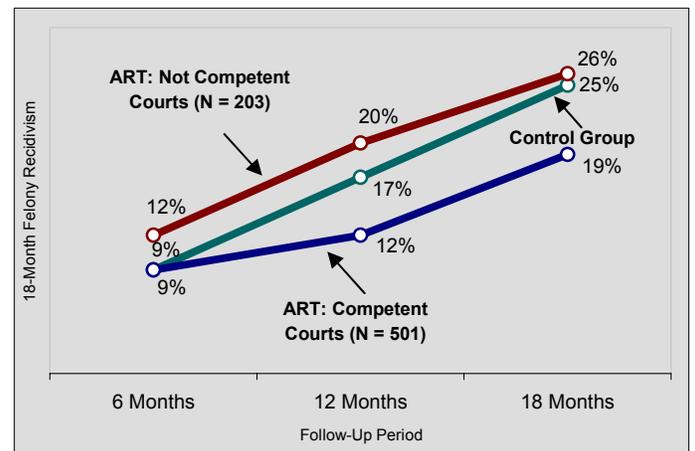
Exhibit 14
Adjusted 18-Month Recidivism Rates Control vs. ART Groups in 2000

STUDY GROUP	YOUTH	MISDEMEANOR AND FELONY	FELONY	VIOLENT FELONY
Control	525	48.6%	24.8%	6.2%
Not Competent	203	50.4%	26.5%	6.8%
Competent	402	47.0%	20.3%	6.6%
Highly Competent	99	36.4%*	12.9%*	6.4%
Total Competent	501	44.9%	18.8%*	6.6%
All ART Youth	704	46.3%	20.8%	6.6%

* Statistically significant at the .05 probability level.

In Exhibit 15, the 6-month, 12-month, and 18-month adjusted felony recidivism rates are displayed for the control group and the competent and not competent ART court groups during 2000.³⁴ The exhibit illustrates that the differences between the control and competent ART court groups first appear at the 12-month follow-up period and continue to the 18-month period. Conversely, the difference that existed at 6-months between the control and not competent ART court groups disappeared by the 18-month period.

Exhibit 15
Adjusted Felony Recidivism Rates 6-, 12-, and 18-Month Follow-up Periods



ART Cost-Benefit Analysis: The cost-benefit analysis, described in Section VII, determines whether Washington citizens receive a positive return on their dollars spent on ART. These analyses find that ART generates \$6.71 in benefits (avoided crime costs) for each taxpayer dollar spent on the program. For courts where ART was competently delivered, the savings are greater—\$11.66 in benefits for each dollar spent on the program.

³⁴ The calculations for adjusted recidivism rates from the logistic models are given in *Outcome Evaluation Appendix*, Exhibit B-3.

ART Conclusions: When ART is delivered competently, the program reduces felony recidivism and is cost effective. For courts rated as competent in delivering ART during 2000, there was a 24 percent reduction in 18-month felony recidivism compared with the control group, which is statistically significant. There is clear evidence that outcomes for ART have improved between its first and second year of operation in Washington, presumably because the courts and program instructors are getting better at delivering ART.

SECTION V: COORDINATION OF SERVICES

What Is Coordination of Services? Coordination of Services (COS), developed by Patrick Tolan, Ph.D.,³⁵ provides an educational program to low-risk juvenile offenders and their parents. The goals of COS are to describe the consequences of continued delinquent behavior, stimulate goal setting, review the strengths of the youth and family, and explain what resources are available for helping to achieve a positive pro-social future for the youth. COS is not a Blueprint Program, having one outcome study supporting this program's effectiveness in reducing recidivism.³⁶

COS was implemented in the Snohomish County Juvenile Court and called the "WayOut" program; Dr. Tolan consulted in training the program providers. WayOut consists of two all-day classes scheduled on consecutive Saturdays. In addition to the juvenile court, several community groups participate in the program: YMCA, WSU Cooperative Extension, Compass Health, 4-H, Snohomish Police, CORE Teen Seminars, and Snohomish County Health Communities Task Force. WayOut costs approximately \$400 per family.

The following are key features of WayOut:

- Low-risk juvenile offenders are court-mandated to attend, thus assuring a captive audience of youth who are at a crossroads when early intervention can make a difference.

- Parents/guardians are also required to attend, thus providing an opportunity to teach parent and child the same skills simultaneously. Additionally, the participants are given a vehicle to open lines of communication and make shifts in thinking.
- Community groups present participants with information concerning the services they provide.

Graduating from WayOut allows the juvenile participants to complete their court-mandated community service hours. The WayOut program coordinator reported that during 2000, ten two-day educational seminars were conducted. Over 90 percent of the youth assigned to the program attended with a parent or guardian.

Adherence to the COS Model: The Institute did not obtain ratings of how well WayOut followed Dr. Tolan's COS model. Conversations with the WayOut service providers indicated they adjusted the original design somewhat.

Evaluation Design: The evaluation design for COS is different from FFT and ART. To simplify procedures for juvenile court staff, the Institute created the control group from the full population of low-risk youth in Snohomish County. Pre-screen data from the WSJCA were used for matching, because a full assessment is not completed for low-risk youth.

Individual control group youth were matched to each WayOut youth on risk level, age, gender, criminal history score, and social history score. Each control group youth had the same risk level, age, and gender values as the WayOut youth. In addition, the WayOut and control youth were matched to within three points, out of a 31 possible points, on criminal history scores, and to within three points on social history scores (18 possible points).

The follow-up period had to be altered for the evaluation of WayOut. The Institute's data on WayOut youth starts in 2000, so only a 12-month follow-up period could be used for the 342 youth in the study sample.

³⁵ Director, Institute for Juvenile Research, University of Illinois at Chicago.

³⁶ Patrick Tolan, M. Shelley Perry, Theodore Jones, "Delinquency Prevention: An Example of Consultation in Rural Community Mental Health," *Journal of Community Psychology* 15 (1987): 43-50.

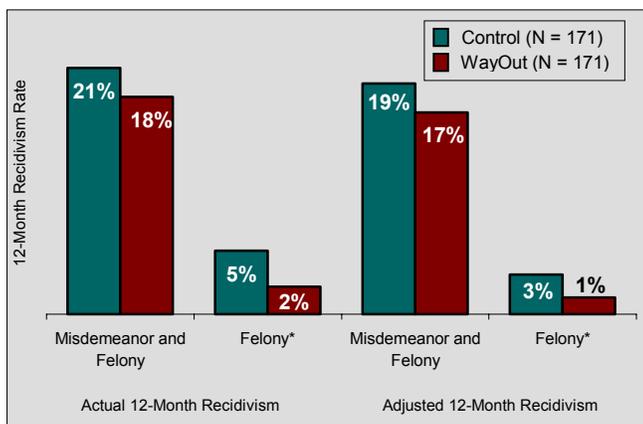
Exhibit 16 displays key characteristics of WayOut and control group youth. No differences were found between the groups.

Exhibit 16
Comparison of Characteristics Between Control and WayOut Groups

VARIABLE	CONTROL	WAYOUT
Number of Youth	171	171
Male	74.9%	74.9%
Age	15.4	15.4
Criminal History	4.3	4.3
Social History	3.6	3.6
Risk Level: Low	87.1%	87.1%
Moderate	8.8%	8.8%
High	4.1%	4.1%

WayOut Results: Exhibit 17 shows both the adjusted and actual 12-month felony recidivism rates for WayOut and the control groups.³⁷ Because these are mostly low-risk youth, the number of those re-offending was expected to be relatively small. Of the 342 youth in the sample, 63 re-offended with a misdemeanor, and 13 re-offended with a felony. These low recidivism rates make it less likely to observe statistically significant differences between the groups. Only three youth re-offended with a violent felony, so the violent felony recidivism rates are too small to analyze.

Exhibit 17
Adjusted and Actual 12-Month Recidivism Rates Control vs. WayOut Groups



* Statistically significant at the .15 probability level.

The 12-month felony recidivism rate for the control group is 5 percent compared with 2 percent for the WayOut group, a 55 percent reduction. The adjusted rates are similar and produce a 59 percent reduction in 12-month felony recidivism. Both these differences are statistically significant at the $p=.15$

probability level. The percent reduction for misdemeanor and felony recidivism is about 12 percent; this difference is not statistically significant.

COS Cost-Benefit Analysis: The cost-benefit analysis, in Section VII of this report, determines whether Washington citizens receive a positive return on their dollars spent on COS. These analyses find that COS generates \$7.89 in savings (avoided crime costs) for each taxpayer dollar spent on the program.

COS Conclusions: The program achieved a cost effective decrease in 12-month felony recidivism, which is close to statistical significance at $p=.15$.

SECTION VI: MULTI-SYSTEMIC THERAPY

What Is Multi-Systemic Therapy? Multi-Systemic Therapy (MST) is an intervention for youth that focuses on improving the family's capacity to overcome the known causes of delinquency.³⁸ Its goals are to promote parents' ability to monitor and discipline their children and replace deviant peer relationships with pro-social friendships. Like FFT, MST is a Blueprint Program.

Trained MST therapists, working in teams consisting of one Ph.D. clinician and three or four clinicians with masters' degrees, have a caseload of four to six families. The intervention typically lasts between three and six months. MST, Inc., in Charleston, South Carolina, trains and clinically supervises all MST therapists. MST, Inc. indicates that costs are approximately \$5,000 per family.

Although MST is on the list of CJAA research-based programs, no juvenile court chose to implement MST using this source of funds. Rather, three counties chose to use federal funding—the Juvenile Accountability Incentive Block Grant (JAIBG). The courts contracted with two organizations to provide MST: Seattle Children's Home in King County and Bold Solutions in Pierce and Kitsap Counties. To be eligible for MST, the CJAA Committee decided that a youth must have a high risk level and family problems as indicated by a family dynamic risk factor score above the eligibility cut-off value (6 out of 24 points). Between January 1999 and September 2001, MST was delivered to 97 families.

³⁷ The calculations for the adjusted recidivism rates are given in *Outcome Evaluation Appendix*, Exhibit C-1.

³⁸ <www.mstservices.com>

Research literature has demonstrated that MST reduces recidivism of juvenile offenders when delivered by MST, Inc. therapists, or therapists under the direct supervision of MST, Inc. The question for this study is whether MST is effective in recidivism reduction in a setting where MST, Inc. is less directly involved with the families. That is, can MST be implemented successfully by three independent juvenile courts with sufficient consistency and program fidelity to reduce recidivism and make the \$5,000 cost per program participant a wise use of taxpayer dollars?

MST Implementation Problems: The research design for MST follows the designs used for FFT and ART. However, the implementation of MST differed in the following ways:

- MST was implemented in only three courts.
- JRA staff did not work closely with the MST courts and providers because MST is not funded under CJAA.
- The number of youth in the MST treatment and control groups is small, which makes finding statistically significant differences less likely.
- The number of youth assigned to individual MST therapists is small, making it difficult to calculate valid recidivism rates for the youth treated by individual therapists.
- Significant differences exist between the MST and the control groups on the WSJCA scores, which raises doubts about the comparability of these groups on key variables.
- The recidivism rates for the control groups for the two organizations are very different. This indicates a strong selection bias in assigning youth to the control or MST groups.

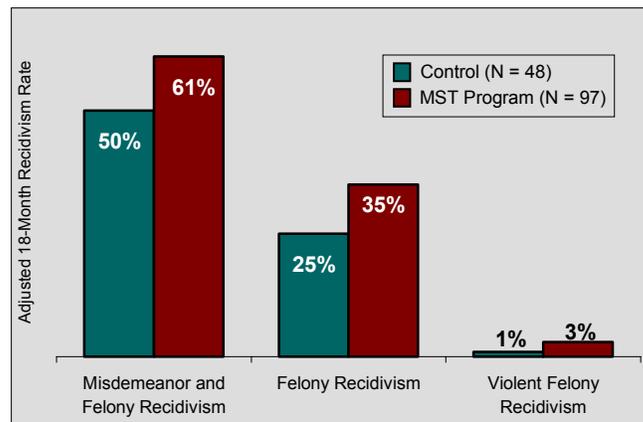
These differences threaten the evaluation’s ability to conclusively indicate whether MST is able to reduce recidivism as implemented in Washington State.

MST Results: Exhibit 18 shows the three adjusted recidivism rates of youth in the MST study groups.³⁹ The 18-month adjusted felony recidivism rate for the control group is 25 percent compared with 35 percent for MST. Although it appears that MST participants had higher recidivism rates, none of the differences in recidivism rates between the two groups is statistically significant. Before

³⁹ The calculations for the adjusted recidivism rates from the logistic models are given in *Outcome Evaluation Appendix, Exhibit D-1*.

reaching any conclusions, the data needs further examination.

Exhibit 18
Adjusted 18-Month Recidivism Rates
Control vs. MST Groups



No statistically significant differences.

Therapist Adherence to MST: MST, Inc. manages the quality assurance process for therapists in both agencies.

MST therapists ask each family to complete the Therapist Adherence Measure (TAM) questionnaire about their treatment. The results are used by the MST clinical supervisor to assess how well each therapist is delivering MST. The TAM measures views of the family receiving treatment and does not represent an independent assessment of how well the therapist adheres to the MST model. Therefore, the TAM was not used by the Institute for this outcome evaluation.

Another MST, Inc. instrument, the Nine Principles Review Form, is used by MST consultants to assess how well therapists follow the nine MST principles. However, no MST expert knew the therapists in both organizations well enough to assess competent delivery. Therefore, the Institute asked the clinical supervisor in the two agencies to rate their therapists retrospectively.

The rating distributions for the clinical supervisors were very different; the Children’s Home ratings were much higher than the Bold Solutions ratings. This result may reflect real differences in therapist behavior, or the use of different “anchor points,”⁴⁰ by

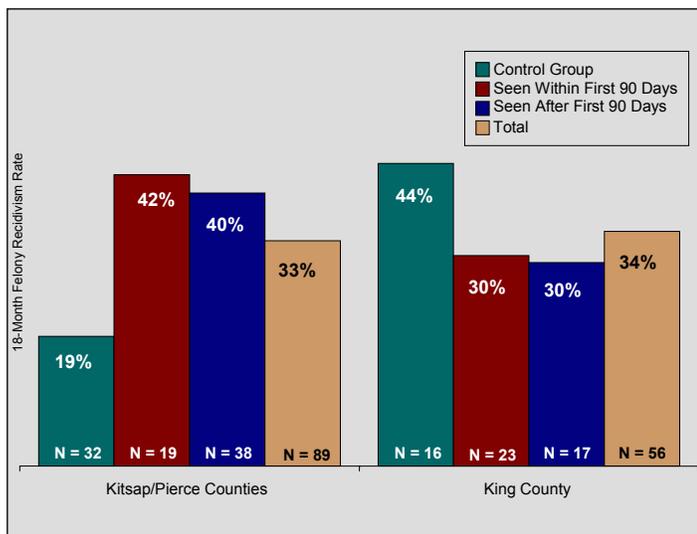
⁴⁰ An anchor point refers to the tendency to pick responses on a subjective scale within a specific range. For example, on a five-point scale, from very bad to very good, some people will anchor their responses around the scale value of 2, being uncomfortable giving high ratings, while others may anchor their responses around 4, being uncomfortable giving low ratings. This problem can be overcome by reducing the subjectivity of the scale.

the two supervisors. Even after standardizing the ratings for each supervisor,⁴¹ only a few items from the Nine Principles Review Form were correlated with recidivism. Therefore, these ratings could not be used to assess therapist competence.

However, these results led to an examination of the outcomes for each organization. The recidivism rates of youth seen in the two courts are examined separately in Exhibit 19. The recidivism rates for youth are separated into two groups: those seen within the therapists' first 90 days of MST practice and those seen subsequently.

First, the felony recidivism rates for all youth within the two courts are similar: 33 percent for Kitsap/Pierce and 34 percent for King. In the WSJCA validation study, the statewide 18-month felony recidivism rate for youth assessed as high risk is estimated as 33 percent. The recidivism rates of all youth in each court are nearly identical to the expected rate. This finding indicates that the youth selected for inclusion in the study for each court are comparable.

Exhibit 19
18-Month Felony Recidivism Rates for Youth in the MST Evaluation



However, the recidivism rates for the control groups for the two organizations are remarkably different: 19 percent for Kitsap/Pierce versus 44 percent for King. Correspondingly, the recidivism rates for the MST groups are also very different with Kitsap/Pierce having much higher recidivism rates than King. This result raises a concern that the assignment of cases to the MST and control groups may not have been

random and may have occurred differently in the two courts. In this event, the findings could be due to the assignment process, not the program.

Exhibit 20 reveals systematic differences between the groups on key characteristics from the WSJCA. For example, in King County, 63 percent of the MST group is male compared with 100 percent of the control group. The King County MST group has significantly higher risk scores in four domains: prior family, attitude, mental health, and relationship. For Kitsap/Pierce, the MST group has higher risk scores for four domains: social history, free-time, mental health, and skill.

Exhibit 20
Comparison of MST and Control Groups On Key Characteristics

VARIABLE	KING COUNTY		KITSAP/PIERCE COUNTIES	
	Control Group	MST Group	Control Group	MST Group
Number of Youth	16	40	32	57
Male Gender	100%	63%***	78%	81%
Average Age at Adjudication	15.5	15.0	15.2	15.3
WSJCA Pre-Screen Average Risk Scores				
Criminal History	9.8	9.0	8.5	8.6
Social History	8.9	10.1	9.6	10.9**
WSJCA Full Assessment Average Risk Scores				
Family	10.6	13.0	13.8	14.5
Prior Family	11.6	16.4*	15.4	16.8
Attitude	7.8	11.9**	8.7	9.2
Drug/Alcohol	6.2	6.7	5.2	6.3
Employment (Protective)	1.0	0.9	1.6	1.1
Free Time	2.6	2.5	1.6	2.2**
Mental Health	1.2	2.7**	2.1	2.7*
Relationship	9.2	12.7*	14.0	15.2
School	11.2	13.9	12.9	12.7
Skill	21.2	21.7	17.4	21.0**
Aggression	2.2	2.5	3.1	2.5

*Statistically significant at the .10 probability level.

**Statistically significant at the .05 probability level.

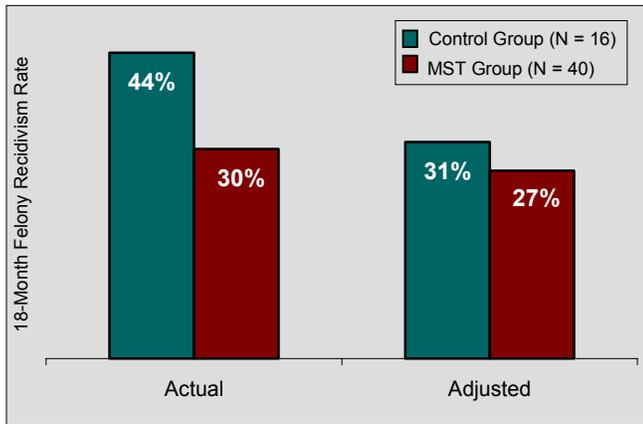
***Statistically significant at the .01 probability level.

Because of the differences between the study groups shown in Exhibits 19 and 20, separate multivariate analyzes for each location are necessary in an attempt to adjust for these differences.

⁴¹ The mean rating for each supervisor was subtracted from each therapist's rating, and the resulting difference was divided by standard deviation of the supervisor's ratings.

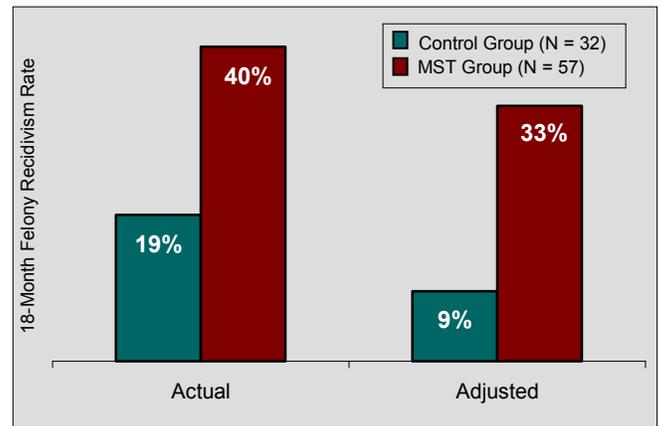
King County Analysis: Exhibit 21 shows both the adjusted and actual 18-month felony recidivism rates for King County.⁴² The model includes the same independent variables used in the modeling of outcomes for FFT. The inclusion of these independent variables reduced the recidivism rate for the control group from 44 percent to 31 percent and the MST group from 30 percent to 27 percent. That is, much of the difference in the felony recidivism rates between the control and MST groups arises from differences in the risk level between the two groups. The estimate of the effect of MST on recidivism was in the right direction, decreasing recidivism by 11.8 percent, but the difference is not statistically significant. With this small sample, much larger effect sizes are needed to achieve statistical significance.

Exhibit 21
King County Adjusted and Actual
18-Month Felony Recidivism Rates
Control Group vs. MST Group



Kitsap and Pierce County Analysis: Exhibit 22 shows both the adjusted and actual 18-month felony recidivism rates for Pierce and Kitsap Counties.⁴³

Exhibit 22
Kitsap/Pierce Counties Adjusted and Actual
18-Month Felony Recidivism Rates
Control Group vs. MST Group



The Kitsap/Pierce model includes the same independent variables used in the modeling of outcomes for MST in King County. The inclusion of these independent variables exacerbated the difference in recidivism rates between the control group and the MST group. The recidivism rate for the control group decreased from 19 percent to 9 percent, while the MST group's recidivism rate decreased from 40 percent to 33 percent. The estimate of the negative effect of MST on recidivism is statistically significant. These results suggest that MST youth had higher rates of recidivism in Kitsap/Pierce than the control group, or that the statistical modeling did not successfully control for systematic differences between treatment and control groups in Kitsap/Pierce.

MST Conclusions: The implementation of MST in Washington State threatened the validity of the evaluation's results. Therefore, this evaluation cannot conclusively indicate whether or not MST, as implemented in Washington State, had any effect on recidivism.

⁴² The calculations for the adjusted recidivism rates are given in *Outcome Evaluation Appendix*, Exhibit C-1.

⁴³ The calculations for the adjusted recidivism rates are given in *Outcome Evaluation Appendix*, Exhibit D-2.

SECTION VII: COST-BENEFIT ANALYSIS

The results of the outcome evaluation of the CJAA programs were described in the preceding sections. The results included the findings of the cost-benefit analyses presented in this section. FFT, ART, and COS cost taxpayers, respectively, \$2,100, \$745, and \$400 per program participant. The cost-benefit question is whether the reduction in recidivism, if any, leads to more benefits than costs. Simply put, are taxpayers better off as a result of the CJAA programs?

To answer this question, the Institute relied on a cost-benefit model developed in recent years.⁴⁴ The model estimates how reductions in crime translate into taxpayer benefits and crime victim benefits. For this evaluation, the model quantifies the dollar value of costs that are avoided when recidivism is reduced by FFT, ART, and COS.

To be conservative, the cost-benefit model uses reduced estimates of program effects to account for the CJAA's less-than-random-assignment

research designs. The FFT effect size is reduced 25 percent, ART 50 percent, and COS 50 percent.

When crimes are avoided, taxpayers do not have to spend as much money on the criminal justice system. Fewer crimes also mean that there are fewer crime victims. This cost-benefit analysis of Washington's CJAA programs estimates the present value of avoided crimes to both taxpayers and crime victims. From the present-value sum of these benefits, we then subtract the costs of the CJAA programs to determine the economic "bottom line."

In this evaluation, we only estimated the effect that the CJAA programs have on crime outcomes. We did not attempt to determine whether the programs improve other outcomes, such as decreases in substance abuse or increases in education levels. As a result, our cost-benefit analysis does not include these other potential, but unmeasured, benefits of the CJAA programs.

Exhibit 23
Summary of Cost-Benefit Results^A

	FFT		ART		COS
	Competent	Not Competent	Competent	Not Competent	
Change in Number of Felony Convictions as a Result of the Program, Per Program Participant	-0.44	+0.17	-0.17	+0.05	-0.08
Program Costs Per Participant	\$2,100	\$2,100	\$745	\$745	\$400
Program Benefits					
• Taxpayer Benefits (avoided criminal justice costs)	\$9,003	-\$3,521	\$3,483	-\$927	\$1,462
• Crime Victim Monetary Costs Avoided	\$4,478	-\$1,751	\$1,732	-\$461	\$570
• Crime Victim Quality of Life Costs Avoided	\$8,967	-\$3,507	\$3,469	-\$923	\$1,124
• Total Taxpayer and Crime Victim Costs Avoided ^B	\$22,448	-\$8,779	\$8,684	-\$2,312	\$3,155
Benefit-to-Cost Ratios From Three Perspectives					
• Taxpayer	\$4.29	-\$1.68	\$4.68	-\$1.24	\$3.65
• Taxpayer and Crime Victim (Monetary Only)	\$6.42	-\$2.51	\$7.00	-\$1.86	\$5.08
• Total Taxpayer and All Crime Victim	\$10.69	-\$4.18	\$11.66	-\$3.10	\$7.89

^A Detailed cost-benefit results for each program are in Appendix E.

^B Totals may not add due to rounding.

⁴⁴ For a complete description of the cost-benefit methods we used in this analysis, see: Aos et al., *The Comparative Costs and Benefits of Programs to Reduce Crime*.

Exhibit 23 summarizes the results of the cost-benefit analysis of FFT, ART, and COS. The table shows the estimated number of felony convictions avoided by the programs from the time youth are 15 years old until they are 30 years old. Exhibit 23 also shows the per-participant cost for each program in 2002 dollars. These program costs were obtained from JRA and reflect the actual spending by the juvenile courts on the programs divided by the total number of youth who entered the program.

The program benefits section of Exhibit 23 displays the present value of the estimated benefits that are generated from the reduced crime from the three programs. The total dollar value of these benefits are shown in their three component parts: those benefits that accrue to taxpayers because of the reduced number of criminal justice system costs, those that accrue to crime victims for monetary (out-of-pocket) costs that are avoided, and those that accrue to crime victims for quality of life cost savings.

The final section of Exhibit 23 displays benefit-cost ratios (benefits divided by program costs) from three perspectives. The taxpayer perspective considers only taxpayer benefits divided by taxpayer costs. The results indicate, for example, that FFT generates \$4.29 in taxpayer savings (avoided costs) for each dollar spent on the program when competently delivered. This means that from the perspective of the taxpayer, FFT is a good investment: each dollar spent will return over ten dollars (present value terms) in taxpayer savings over the next 15 years.

The additional two perspectives for the benefit-cost ratios shown in Exhibit 23 include crime victim costs avoided in addition to those that accrue just to taxpayers. The second perspective includes only so-called crime victim “monetary” costs avoided by the reduction in crime. These victim costs include only those out-of-pocket expenses (e.g. medical costs, lost wages) that victims suffer when crimes occur. The FFT program, for example, generates \$6.42 in benefits for each dollar of costs when victim monetary benefits are added to the taxpayer benefits. The final perspective on program benefits includes a broader, and sometimes more controversial, definition of crime victim costs of crime: quality of life losses that victims suffer when crime occurs. After including these quality-of-life benefits, the FFT benefit-to-cost ratio increases to \$10.69 of benefits per dollar of cost.⁴⁵

To obtain the overall benefit to cost ratio for FFT and ART, regardless of therapist competence, the benefit to cost ratios for competent and not competence service are averaged.

SECTION VIII: SUMMARY OF FINDINGS

✓ Functional Family Therapy

Youth seen by competent therapists have an 18 percent felony recidivism rate compared with 27 percent for the control group, a statistically significant reduction of 38 percent ($p=.01$). For violent felony recidivism, the competent therapist group has a 3 percent rate compared with 6 percent for the control group, a 50 percent reduction that is statistically significant at the $p=.115$ probability level.

The cost-benefit analysis, as shown in Exhibit 1, determines whether Washington citizens receive a positive return on their dollars spent on FFT. When FFT is delivered by competent therapists, it generates \$10.69 in benefits (avoided crime costs) for each dollar spent on the program. When not competently delivered, FFT costs the taxpayer \$4.18. Averaging these results for all youth receiving FFT, regardless of therapist competence, results in a net savings of \$2.77 per dollar of costs.

✓ Aggression Replacement Training

For the 21 courts rated as either competent or highly competent in delivering ART, the 18-month felony recidivism rate is 19 percent. This is a 24 percent reduction in felony recidivism compared with the control group, which is statistically significant ($p=.03$).

The cost-benefit analysis, as shown in Exhibit 1, determines whether Washington citizens receive a positive return on their dollars spent on ART. When ART is delivered by competent courts, it generates \$11.66 in benefits (avoided crime costs) for each dollar spent on the program. When not competently delivered, ART costs the taxpayer \$3.10. Averaging these results for all youth receiving ART, regardless of court competence, results in a net savings of \$6.71 per dollar of costs.

⁴⁵ A more detailed discussion of the crime victim cost definitions is contained in Aos et al., *The Comparative Costs and Benefits of Programs to Reduce Crime*.

✓ **Coordination of Services**

The 12-month felony recidivism rate for the control group is 5 percent compared with 2 percent for the WayOut group, a 55 percent reduction. The adjusted rates are similar and produce a 59 percent reduction in 12-month felony recidivism. Both of these differences are statistically significant at the $p=.15$ probability level.

The cost-benefit analyses find that COS generates \$7.89 in savings (avoided crime costs) for each taxpayer dollar spent on the program.

✓ **Multi-Systemic Therapy**

The evaluation of MST, as implemented in Washington State, cannot conclusively indicate whether MST was able to reduce recidivism.

✓ **Overall**

These findings indicate that research-based programs can reduce recidivism. However, without quality assurance, programs may not only fail to reduce recidivism, they may actually increase recidivism. The 2003 Washington State Legislature acted on the preliminary CJAA evaluation results by directing the Institute to develop adherence and outcome standards for juvenile justice research-based programs (RCW 13.40.530), which were published in December 2003.

This report affirms the merit of Legislature's investment in research-based programs for juvenile offenders. The next step is to implement the quality assurance standards so that taxpayer benefits can confidently be obtained for each dollar spent on the CJAA programs.

ACKNOWLEDGEMENTS

The Institute wishes to thank all of the juvenile court administrators for their cooperation and support of this project evaluation. In particular, we wish to thank the juvenile court members of the CJAA Committee: Dick Carlson, Barb Carr, Dan Erker, Audrey Fethers, Harris Haertel, Bruce Knutson, Ken Trull, Larry Wardle, and Bernie Warner, as well as Kathleen McBride, Doug Kopp, and Jeff Patnode of the Washington State Juvenile Rehabilitation Administration.

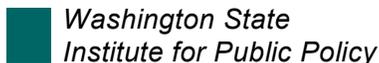
This project could not have been completed without the work of two more CJAA Committee members: Chris Hayes of the Snohomish County Juvenile Court and Dana Phelps of the Juvenile Rehabilitation Administration.

We would also like to thank the members of the Assessment Quality Assurance Committee: Dick Carlson, Tom Davis, Harold Delia, Dan Erker, Greg Grammer, Tom Kearney, Sharon Paradis, and Diana Wavra.

Finally, it is the competent work of the juvenile court staff and the program providers that produced these positive findings.

All reports published by the Institute are available at www.wsipp.wa.gov. For further information, contact Robert Barnoski, (360) 586-2744, barney@wsipp.wa.gov; or Steve Aos (360) 586-2740, saos@wsipp.wa.gov

Document Number: 04-01-1201



The Washington Legislature created the Washington State Institute for Public Policy in 1983. A Board of Directors—representing the legislature, the governor, and public universities—governs the Institute and guides the development of all activities. The Institute's mission is to carry out practical research, at legislative direction, on issues of importance to Washington State.

Beyond Correctional Quackery— Professionalism and the Possibility of Effective Treatment

Edward J. Latessa, University of Cincinnati

Francis T. Cullen, University of Cincinnati

Paul Gendreau, University of New Brunswick at Saint John

LONG-TIME VIEWERS of *Saturday Night Live* will vividly recall Steve Martin's hilarious portrayal of a medieval medical practitioner—the English barber, Theodoric of York. When ill patients are brought before him, he prescribes ludicrous “cures,” such as repeated bloodletting, the application of leeches and boar's vomit, gory amputations, and burying people up to their necks in a marsh. At a point in the skit when a patient dies and Theodoric is accused of “not knowing what he is doing,” Martin stops, apparently struck by the transforming insight that medicine might abandon harmful interventions rooted in ignorant customs and follow a more enlightened path. “Perhaps,” he says, “I've been wrong to blindly follow the medical traditions and superstitions of past centuries.” He then proceeds to wonder whether he should “test these assumptions analytically through experimentation and the scientific method.” And perhaps, he says, the scientific method might be applied to other fields of learning. He might even be able to “lead the way to a new age—an age of rebirth, a renaissance.” He then pauses and gives the much-awaited and amusing punchline, “Nawwwwwww!”

The humor, of course, lies in the juxtaposition and final embrace of blatant quackery with the possibility and rejection of a more modern, scientific, and ultimately effective approach to medicine. For those of us who make a living commenting on or doing corrections, however, we must consider whether, in a sense, the joke is on us. We can readily see the humor in Steve Martin's skit and wonder how those in medieval societies “could have been so stupid.” But even a cursory sur-

vey of *current* correctional practices yields the disquieting conclusion that we are a field in which quackery is tolerated, if not implicitly celebrated. It is not clear whether most of us have ever had that reflective moment in which we question whether, “just maybe,” there might be a more enlightened path to pursue. If we have paused to envision a different way of doing things, it is apparent that our reaction, after a moment's contemplation, too often has been, “Nawwwwwww!”

This appraisal might seem overly harsh, but we are persuaded that it is truthful. When intervening in the lives of offenders—that is, intervening with the expressed intention of reducing recidivism—corrections has resisted becoming a true “profession.” Too often, being a “professional” has been debased to mean dressing in a presentable way, having experience in the field, and showing up every day for work. But a profession is defined not by its surface appearance but by its intellectual core. An occupation may lay claim to being a “profession” only to the extent that its practices are based on research knowledge, training, and expertise—a triumvirate that promotes the possibility that what it does can be effective (Cullen, 1978; Starr, 1982). Thus, medicine's professionalization cannot be separated from its embrace of scientific knowledge as the ideal arbiter of how patients should be treated (Starr, 1982). The very concept of “malpractice” connotes that standards of service delivery have been established, are universally transmitted, and are capable of distinguishing acceptable from unacceptable interventions. The concept of liability for “correctional malpractice” would bring

snickers from the crowd—a case where humor unintentionally offers a damning indictment of the field's standards of care.

In contrast to professionalism, *quackery* is dismissive of scientific knowledge, training, and expertise. Its posture is strikingly overconfident, if not arrogant. It embraces the notion that interventions are best rooted in “common sense,” in personal experiences (or clinical knowledge), in tradition, and in superstition (Gendreau, Goggin, Cullen, and Paparozzi, forthcoming). “What works” is thus held to be “obvious,” derived only from years of an individual's experience, and legitimized by an appeal to custom (“the way we have always done things around here has worked just fine”). It celebrates being anti-intellectual. There is never a need to visit a library or consult a study.

Correctional quackery, therefore, is the use of treatment interventions that are based on neither 1) existing knowledge of the causes of crime nor 2) existing knowledge of what programs have been shown to change offender behavior (Cullen and Gendreau, 2000; Gendreau, 2000). The hallmark of correctional quackery is thus ignorance. Such ignorance about crime and its cures at times is “understandable”—that is, linked not to the willful rejection of research but to being in a field in which professionalism is not expected or supported. At other times, however, quackery is proudly displayed, as its advocates boldly proclaim that they have nothing to learn from research conducted by academics “who have never worked with a criminal” (a claim that is partially true but ultimately beside the point and a rationalization for continued ignorance).

Need we now point out the numerous programs that have been implemented with much fanfare and with amazing promises of success, only later to turn out to have "no effect" on reoffending? "Boot camps," of course, are just one recent and salient example. Based on a vague, if not unstated, theory of crime and an absurd theory of behavioral change ("offenders need to be broken down"—through a good deal of humiliation and threats—and then "built back up"), boot camps could not possibly have "worked." In fact, we know of no major psychological theory that would logically suggest that such humiliation or threats are components of effective therapeutic interventions (Gendreau et al., forthcoming). Even so, boot camps were put into place across the nation without a shred of empirical evidence as to their effectiveness, and only now has their appeal been tarnished after years of negative evaluation studies (Cullen, Pratt, Miceli, and Moon, 2002; Cullen, Wright, and Applegate, 1996; Gendreau, Goggin, Cullen, and Andrews, 2000; MacKenzie, Wilson, and Kider, 2001). How many millions of dollars have been squandered? How many opportunities to rehabilitate offenders have been forfeited? How many citizens have been needlessly victimized by boot camp graduates? What has been the cost to society of this quackery?

We are not alone in suggesting that advances in our field will be contingent on the conscious rejection of quackery in favor of an *evidence-based corrections* (Cullen and Gendreau, 2000; MacKenzie, 2000; Welsh and Farrington, 2001). Moving beyond correctional quackery when intervening with offenders, however, will be a daunting challenge. It will involve overcoming four central failures now commonplace in correctional treatment. We review these four sources of correctional quackery not simply to show what is lacking in the field but also in hopes of illuminating what a truly professional approach to corrections must strive to entail.

Four Sources of Correctional Quackery

Failure to Use Research in Designing Programs

Every correctional agency must decide "what to do" with the offenders under its supervision, including selecting which "programs" or "interventions" their charges will be subjected to. But how is this choice made (a choice that is consequential to the offender,

the agency, and the community)? Often, no real choice is made, because agencies simply continue with the practices that have been inherited from previous administrations. Other times, programs are added incrementally, such as when concern rises about drug use or drunk driving. And still other times—such as when punishment-oriented intermediate sanctions were the fad from the mid-1980s to the mid-1990s—jurisdictions copy the much-publicized interventions being implemented elsewhere in the state and in the nation.

TABLE 1
Questionable Theories of Crime We Have Encountered in Agency Programs

- ✔ "Been there, done that" theory.
- ✔ "Offenders lack creativity" theory.
- ✔ "Offenders need to get back to nature" theory.
- ✔ "It worked for me" theory.
- ✔ "Offenders lack discipline" theory.
- ✔ "Offenders lack organizational skills" theory.
- ✔ "Offenders have low self-esteem" theory.
- ✔ "We just want them to be happy" theory.
- ✔ The "treat offenders as babies and dress them in diapers" theory.
- ✔ "Offenders need to have a pet in prison" theory.
- ✔ "Offenders need acupuncture" theory.
- ✔ "Offenders need to have healing lodges" theory.
- ✔ "Offenders need drama therapy" theory.
- ✔ "Offenders need a better diet and haircut" theory.
- ✔ "Offenders (females) need to learn how to put on makeup and dress better" theory.
- ✔ "Offenders (males) need to get in touch with their feminine side" theory.

Notice, however, what is missing in this account: The failure to consider the existing research on program effectiveness. The risk of quackery rises to the level of virtual certainty when nobody in the agency asks, "Is there any evidence supporting what we are intending to do?" The irrationality of not consulting the existing research is seen when we consider again, medicine. Imagine if local physicians and hospitals made no effort to consult "what works" and simply prescribed pharmaceuticals and conducted surgeries based on custom or the latest fad. Such malpractice would be greeted with public condemnation, lawsuits, and a loss of legitimacy by the field of medicine.

It is fair to ask whether research can, in fact, direct us to more effective correctional interventions. Two decades ago, our knowledge was much less developed. But the science of crime and treatment has made important strides in the intervening years. In particular, research has illuminated three bodies of knowledge that are integral to designing effective interventions.

First, we have made increasing strides in determining the *empirically established or known predictors* of offender recidivism (Andrews and Bonta, 1998; Gendreau, Little, and Goggin, 1996; Henggeler, Mihalic, Rone, Thomas, and Timmons-Mitchell, 1998). These include, most importantly: 1) antisocial values, 2) antisocial peers, 3) poor self-control, self-management, and prosocial problem-solving skills, 4) family dysfunction, and 5) past criminality. This information is critical, because interventions that ignore these factors are doomed to fail. Phrased alternatively, successful programs start by recognizing what causes crime and then *specifically design the intervention to target these factors for change* (Alexander, Pugh, and Parsons, 1998; Andrews and Bonta, 1998; Cullen and Gendreau, 2000; Henggeler et al., 1998).

Consider, however, the kinds of "theories" about the causes of crime that underlie many correctional interventions. In many cases, simple ignorance prevails; those working in correctional agencies cannot explain what crime-producing factors the program is allegedly targeting for change. Still worse, many programs have literally invented seemingly ludicrous theories of crime that are put forward with a straight face. From our collective experiences, we have listed in Table 1 crime theories that either 1) were implicit in programs we observed or 2) were voiced by agency personnel when asked what crime-causing factors their programs were target-

ing. These "theories" would be amusing except that they are commonplace and, again, potentially lead to correctional quackery. For example, the theory of "offenders (males) need to get in touch with their feminine side" prompted one agency to have offenders dress in female clothes. We cannot resist the temptation to note that you will now know whom to blame if you are mugged by a cross-dresser! But, in the end, this is no laughing matter. This intervention has no chance to be effective, and thus an important chance was forfeited to improve offenders' lives and to protect public safety.

Second, there is now a growing literature that outlines what does *not* work in offender treatment (see, e.g., Cullen, 2002; Cullen and Gendreau, 2000; Cullen et al., 2002; Cullen et al., 1996; Gendreau, 1996; Gendreau et al., 2000; Lipsey and Wilson, 1998; MacKenzie, 2000). These include boot camps, punishment-oriented programs (e.g., "scared straight" programs), control-oriented programs (e.g., intensive supervision programs), wilderness programs, psychological interventions that are non-directive or insight-oriented (e.g., psychoanalytic), and non-intervention (as suggested by labeling theory). Ineffective programs also target for treatment low-risk offenders and target for change weak predictors of criminal behavior (e.g., self-esteem). Given this knowledge, it would be a form of quackery to continue to use or to freshly implement these types of interventions.

Third, conversely, there is now a growing literature that outlines what *does* work in offender treatment (Cullen, 2002; Cullen and Gendreau, 2000). Most importantly, efforts are being made to develop principles of effective intervention (Andrews, 1995; Andrews and Bonta, 1998; Gendreau, 1996). These principles are listed in Table 2. Programs that adhere to these principles have been found to achieve meaningful reductions in recidivism (Andrews, Dowden, and Gendreau, 1999; Andrews, Zinger, Hoge, Bonta, Gendreau, and Cullen, 1990; Cullen, 2002). However, programs that are designed without consulting these principles are almost certain to have little or no impact on offender recidivism and may even risk increasing re-offending. That is, if these principles are ignored, quackery is likely to result. We will return to this issue below.

TABLE 2

Eight Principles of Effective Correctional Intervention

1. Organizational Culture

Effective organizations have well-defined goals, ethical principles, and a history of efficiently responding to issues that have an impact on the treatment facilities. Staff cohesion, support for service training, self-evaluation, and use of outside resources also characterize the organization.

2. Program Implementation/Maintenance

Programs are based on empirically-defined needs and are consistent with the organization's values. The program is fiscally responsible and congruent with stakeholders' values. Effective programs also are based on thorough reviews of the literature (i.e., meta-analyses), undergo pilot trials, and maintain the staff's professional credentials.

3. Management/Staff Characteristics

The program director and treatment staff are professionally trained and have previous experience working in offender treatment programs. Staff selection is based on their holding beliefs supportive of rehabilitation and relationship styles and therapeutic skill factors typical of effective therapies.

4. Client Risk/Need Practices

Offender risk is assessed by psychometric instruments of proven predictive validity. The risk instrument consists of a wide range of dynamic risk factors or criminogenic needs (e.g., anti-social attitudes and values). The assessment also takes into account the responsiveness of offenders to different styles and modes of service. Changes in risk level over time (e.g., 3 to 6 months) are routinely assessed in order to measure intermediate changes in risk/need levels that may occur as a result of planned interventions.

5. Program Characteristics

The program targets for change a wide variety of criminogenic needs (factors that predict recidivism), using empirically valid behavioral/social learning/cognitive behavioral therapies that are directed to higher-risk offenders. The ratio of rewards to punishers is at least 4:1. Relapse prevention strategies are available once offenders complete the formal treatment phase.

6. Core Correctional Practice

Program therapists engage in the following therapeutic practices: anti-criminal modeling, effective reinforcement and disapproval, problem-solving techniques, structured learning procedures for skill-building, effective use of authority, cognitive self-change, relationship practices, and motivational interviewing.

7. Inter-Agency Communication

The agency aggressively makes referrals and advocates for its offenders in order that they receive high quality services in the community.

8. Evaluation

The agency routinely conducts program audits, consumer satisfaction surveys, process evaluations of changes in criminogenic need, and follow-ups of recidivism rates. The effectiveness of the program is evaluated by comparing the respective recidivism rates of risk-control comparison groups of other treatments or those of a minimal treatment group.

Note: Items adapted from the *Correctional Program Assessment Inventory—2000*, a 131-item Questionnaire that is widely used in assessing the quality of correctional treatment programs (Gendreau and Andrews, 2001).

Failure to Follow Appropriate Assessment and Classification Practices

The steady flow of offenders into correctional agencies not only strains resources but also creates a continuing need to allocate treatment resources efficaciously. This problem is not dissimilar to a hospital that must process a steady flow of patients. In a hospital (or doctor's office), however, it is immediately recognized that the crucial first step to delivering effective treatment is diagnosing or *assessing* the patient's condition and its severity. In the absence of such a diagnosis—which might involve the careful study of symptoms or a battery of tests—the treatment prescribed would have no clear foundation. Medicine would be a lottery in which the ill would hope the doctor assigned the right treatment. In a similar way, effective treatment intervention requires the appropriate assessment of both the risks posed by, and the needs underlying the criminality of, offenders. When such diagnosis is absent and no classification of offenders is possible, offenders in effect enter a treatment lottery in which their access to effective intervention is a chancy proposition.

Strides have been made to develop more effective classification instruments—such as the Level of Supervision Inventory (LSI) (Bonta, 1996), which, among its competitors, has achieved the highest predictive validity with recidivism (Gendreau et al., 1996). The LSI and similar instruments classify offenders by using a combination of “static” factors (such as criminal history) and “dynamic factors” (such as antisocial values, peer associations) shown by previous research to predict recidivism. In this way, it is possible to classify offenders by their level of risk and to discern the types and amount of “criminogenic needs” they possess that should be targeted for change in their correctional treatment.

At present, however, there are three problems with offender assessment and classification by correctional agencies (Gendreau and Goggin, 1997). First, many agencies simply do not assess offenders, with many claiming they do not have the time. Second, when agencies do assess, they assess poorly. Thus, they often use outdated, poorly designed, and/or empirically unvalidated classification instruments. In particular, they tend to rely on instruments that measure exclusively static predictors of recidivism (which cannot, by definition, be changed) and that provide no information on the criminogenic needs that offenders have. If these “needs” are not iden-

tified and addressed—such as possessing antisocial values—the prospects for recidivism will be high. For example, a study of 240 (161 adult and 79 juvenile) programs assessed across 30 states found that 64 percent of the programs did not utilize a standardized and objective assessment tool that could distinguish risk/needs levels for offenders (Matthews, Hubbard, and Latessa, 2001; Latessa, 2002).

Third, even when offenders are assessed using appropriate classification instruments, agencies frequently ignore the information. It is not uncommon, for example, for offenders to be assessed and then for everyone to be given the same treatment. In this instance, assessment becomes an organizational routine in which paperwork is compiled but the information is ignored.

Again, these practices increase the likelihood that offenders will experience correctional quackery. In a way, treatment is delivered blindly, with agency personnel equipped with little knowledge about the risks and needs of the offenders under their supervision. In these circumstances, it is impossible to know which offenders should receive which interventions. Any hopes of individualizing interventions effectively also are forfeited, because the appropriate diagnosis either is unavailable or hidden in the agency's unused files.

Failure to Use Effective Treatment Models

Once offenders are assessed, the next step is to select an appropriate treatment model. As we have suggested, the challenge is to consult the empirical literature on “what works,” and to do so with an eye toward programs that conform to the principles of effective intervention. At this stage, it is inexcusable either to ignore this research or to implement programs that have been shown to be ineffective. Yet, as we have argued, the neglect of the existing research on effective treatment models is widespread. In the study of 240 programs noted above, it was reported that two-thirds of adult programs and over half of juvenile programs did not use a treatment model that research had shown to be effective (Matthews et al., 2001; Latessa, 2002). Another study—a meta-analysis of 230 program evaluations (which yielded 374 tests or effect sizes)—categorized the extent to which interventions conformed to the principles of effective intervention. In only 13 percent of the tests were the interventions judged to fall into the “most

appropriate” category (Andrews et al., 1999). But this failure to employ an appropriate treatment approach does not have to be the case. Why would an agency—in this information age—risk quackery when the possibility of using an evidence-based program exists? Why not select effective treatment models?

Moving in this direction is perhaps mostly a matter of a change of consciousness—that is, an awareness by agency personnel that quackery must be rejected and programs with a track record of demonstrated success embraced. Fortunately, depending on the offender population, there is a growing number of treatment models that might be learned and implemented (Cullen and Applegate, 1997). Some of the more prominent models in this regard are the “Functional Family Therapy” model that promotes family cohesion and affection (Alexander et al., 1998; Gordon, Graves, and Arbuthnot, 1995), the teaching youths to think and react responsibly peer-helping (“Equip”) program (Gibbs, Potter, and Goldstein, 1995), the “Prepare Curriculum” program (Goldstein, 1999), “Multisystemic Therapy” (Henggeler et al., 1998), and the prison-based “Rideau Integrated Service Delivery Model” that targets criminal thinking, anger, and substance abuse (see Gendreau, Smith, and Goggin, 2001).

Failure to Evaluate What We Do

Quackery has long prevailed in corrections because agencies have traditionally required no systematic evaluation of the effectiveness of their programs (Gendreau, Goggin, and Smith, 2001). Let us admit that many agencies may not have the human or financial capital to conduct ongoing evaluations. Nonetheless, it is not clear that the failure to evaluate has been due to a lack of capacity as much as to a lack of desire. The risk inherent in evaluation, of course, is that practices that are now unquestioned and convenient may be revealed as ineffective. Evaluation, that is, creates accountability and the commitment threat of having to change what is now being done. The cost of change is not to be discounted, but so too is the “high cost of ignoring success” (Van Voorhis, 1987). In the end, a professional must be committed to doing not simply what is in one's self-interest but what is ethical and effective. To scuttle attempts at program evaluation and to persist in using failed interventions is wrong and a key ingredient to continued correctional quackery (more broadly, see Van Voorhis, Cullen, and Applegate, 1995).

Evaluation, moreover, is not an all-or-nothing procedure. Ideally, agencies would conduct experimental studies in which offenders were randomly assigned to a treatment or control group and outcomes, such as recidivism, were measured over a lengthy period of time. But let us assume that, in many settings, conducting this kind of sophisticated evaluation is not feasible. It is possible, however, for virtually all agencies to monitor, to a greater or lesser extent, the *quality* of the programs that they or outside vendors are supplying. Such evaluative monitoring would involve, for example, assessing whether treatment services are being delivered as designed, supervising and giving constructive feedback to treatment staff, and studying whether offenders in the program are making progress on targeted criminogenic factors (e.g., changing antisocial attitudes, manifesting more prosocial behavior). In too many cases, offenders are “dropped off” in intervention programs and then, eight or twelve weeks later, are deemed—without any basis for this conclusion—to have “received treatment.” Imagine if medical patients entered and exited hospitals with no one monitoring their treatment or physical recovery. Again, we know what we could call such practices.

Conclusion—Becoming an Evidence-Based Profession

In assigning the label “quackery” to much of what is now being done in corrections, we run the risk of seeming, if not being, preachy and pretentious. This is not our intent. If anything, we mean to be provocative—not for the sake of causing a stir, but for the purpose of prompting correctional leaders and professionals to stop using treatments that cannot possibly be effective. If we make readers think seriously about how to avoid selecting, designing, and using failed correctional interventions, our efforts will have been worthwhile.

We would be remiss, however, if we did not confess that academic criminologists share the blame for the continued use of ineffective programs. For much of the past quarter century, most academic criminologists have abandoned correctional practitioners. Although some notable exceptions exist, we have spent much of our time claiming that “nothing works” in offender rehabilitation and have not created partnerships with those

in corrections so as to build knowledge on “what works” to change offenders (Cullen and Gendreau, 2001). Frequently, what guidance criminologists have offered correctional agencies has constituted *bad* advice—ideologically inspired, not rooted in the research, and likely to foster quackery. Fortunately, there is a growing movement among criminologists to do our part both in discerning the principles of effective intervention and in deciphering what interventions have empirical support (Cullen and Gendreau, 2001; MacKenzie, 2000; Welsh and Farrington, 2001). Accordingly, the field of corrections has more information available to find out what our “best bets” are when intervening with offenders (Rhine, 1998).

We must also admit that our use of medicine as a comparison to corrections has been overly simplistic. We stand firmly behind the central message conveyed—that what is done in corrections would be grounds for malpractice in medicine—but we have glossed over the challenges that the field of medicine faces in its attempt to provide scientifically-based interventions. First, scientific knowledge is not static but evolving. Medical treatments that appear to work now may, after years of study, prove ineffective or less effective than alternative interventions. Second, even when information is available, it is not clear that it is effectively transmitted or that doctors, who may believe in their personal “clinical experience,” will be open to revising their treatment strategies (Hunt, 1997). “The gap between research and knowledge,” notes Millenson (1997, p. 4), “has real consequences....when family practitioners in Washington State were queried about treating a simple urinary tract infection in women, eighty-two physicians came up with an extraordinary 137 different strategies.” In response to situations like these, there is a renewed evidence-based movement in medicine to improve the quality of medical treatments (Millenson, 1997; Timmermans and Angell, 2001).

Were corrections to reject quackery in favor of an evidence-based approach, it is likely that agencies would face the same difficulties that medicine encounters in trying base treatments on the best scientific knowledge available. Designing and implementing an effective program is more complicated, we re-

alize, than simply visiting a library in search of research on program effectiveness (although this is often an important first step). Information must be available in a form that can be used by agencies. As in medicine, there must be opportunities for training and the provision of manuals that can be consulted in how *specifically* to carry out an intervention. Much attention has to be paid to implementing programs as they are designed. And, in the long run, an effort must be made to support widespread program evaluation and to use the resulting data both to improve individual programs and to expand our knowledge base on effective programs generally.

To move beyond quackery and accomplish these goals, the field of corrections will have to take seriously what it means to be a *profession*. In this context, individual agencies and individuals within agencies would do well to strive to achieve what Gendreau et al. (forthcoming) refer to as the “3 C’s” of effective correctional policies: First, employ *credentialed people*; second, ensure that the *agency is credentialed* in that it is founded on the principles of fairness and the improvement of lives through ethically defensive means; and third, base treatment decisions on *credentialed knowledge* (e.g., research from meta-analyses).

By themselves, however, given individuals and agencies can do only so much to implement effective interventions—although each small step away from quackery and toward an evidence-based practice potentially makes a meaningful difference. The broader issue is whether the *field* of corrections will embrace the principles that all interventions should be based on the best research evidence, that all practitioners must be sufficiently trained so as to develop expertise in how to achieve offender change, and that an ethical corrections cannot tolerate treatments known to be foolish, if not harmful. In the end, correctional quackery is not an inevitable state of affairs—something we are saddled with for the foreseeable future. Rather, although a formidable foe, it is ultimately rooted in our collective decision to tolerate ignorance and failure. Choosing a different future for corrections—making the field a true profession—will be a daunting challenge, but it is a future that lies within our power to achieve.

References

- Alexander, James, Christie Pugh, and Bruce Parsons. 1998. *Functional Family Therapy: Book Three in the Blueprints and Violence Prevention Series*. Boulder, CO: Center for the Study and Prevention of Violence, University of Colorado.
- Andrews, D. A. 1995. "The Psychology of Criminal Conduct and Effective Treatment." Pp. 35-62 in James McGuire (ed.), *What Works: Reducing Reoffending*. West Sussex, UK: John Wiley.
- Andrews, D. A., and James Bonta. 1998. *Psychology of Criminal Conduct*, 2nd ed. Cincinnati: Anderson.
- Andrews, D. A., Craig Dowden, and Paul Gendreau. 1999. "Clinically Relevant and Psychologically Informed Approaches to Reduced Re-Offending: A Meta-Analytic Study of Human Service, Risk, Need, Responsivity, and Other Concerns in Justice Contexts." Unpublished manuscript, Carleton University.
- Andrews, D. A., Ivan Zinger, R. D. Hoge, James Bonta, Paul Gendreau, and Francis T. Cullen. 1990. "Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis." *Criminology* 28: 369-404.
- Bonta, James. 1996. "Risk-Needs Assessment and Treatment." Pp. 18-32 in Alan T. Harland (ed.), *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply*. Thousand Oaks, CA: Sage.
- Cullen, Francis T. 2002. "Rehabilitation and Treatment Programs." Pp. 253-289 in James Q. Wilson and Joan Petersilia (eds.), *Crime: Public Policies for Crime Control*. Oakland, CA: ICS Press.
- Cullen, Francis T. and Brandon K. Applegate, eds. 1997. *Offender Rehabilitation: Effective Correctional Intervention*. Aldershot, UK: Ashgate/Dartmouth.
- Cullen, Francis T. and Paul Gendreau. 2000. "Assessing Correctional Rehabilitation: Policy, Practice, and Prospects." Pp. 109-175 in Julie Horney (ed.), *Criminal Justice 2000: Volume 3—Policies, Processes, and Decisions of the Criminal Justice System*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Cullen, Francis T. and Paul Gendreau. 2001. "From Nothing Works to What Works: Changing Professional Ideology in the 21st Century." *The Prison Journal* 81:313-338.
- Cullen, Francis T., Travis C. Pratt, Sharon Levrant Miceli, and Melissa M. Moon. 2002. "Dangerous Liaison? Rational Choice Theory as the Basis for Correctional Intervention." p. 279-296 in Alex R. Piquero and Stephen G. Tibbetts (eds.), *Rational Choice and Criminal Behavior: Recent Research and Future Challenges*. New York: Routledge.
- Cullen, Francis T., John Paul Wright, and Brandon K. Applegate. 1996. "Control in the Community: The Limits of Reform?" Pp. 69-116 in Alan T. Harland (ed.), *Choosing Correctional Interventions That Work: Defining the Demand and Evaluating the Supply*. Thousand Oaks, CA: Sage.
- Cullen, John B. 1978. *The Structure of Professionalism*. Princeton, NJ: Petrocelli Books.
- Gendreau, Paul. 1996. "The Principles of Effective Intervention with Offenders." Pp. 117-130 in Alan T. Harland (ed.), *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply*. Newbury Park, CA: Sage.
- Gendreau, Paul. 2000. "1998 Margaret Mead Award Address: Rational Policies for Reforming Offenders." Pp. 329-338 in Maeve McMahon (ed.), *Assessment to Assistance: Programs for Women in Community Corrections*. Lanham, MD: American Correctional Association.
- Gendreau, Paul and D. A. Andrews. 2001. *Correctional Program Assessment Inventory—2000*. Saint John, Canada: Authors.
- Gendreau, Paul and Claire Goggin. 1997. "Correctional Treatment: Accomplishments and Realities." Pp. 271-279 in Patricia Van Voorhis, Michael Braswell, and David Lester (eds.), *Correctional Counseling and Rehabilitation*, 3rd edition. Cincinnati: Anderson.
- Gendreau, Paul, Claire Goggin, Francis T. Cullen, and D. A. Andrews. 2000. "The Effects of Community Sanctions and Incarceration on Recidivism." *Forum on Corrections Research* 12 (May): 10-13.
- Gendreau, Paul, Claire Goggin, Francis T. Cullen, and Mario Paparozzi. Forthcoming. "The Common Sense Revolution in Correctional Policy." In James McGuire (ed.), *Offender Rehabilitation and Treatment: Effective Programs and Policies to Reduce Re-Offending*. Chichester, UK: John Wiley and Sons.
- Gendreau, Paul, Claire Goggin, and Paula Smith. 2001. "Implementing Correctional Interventions in the 'Real' World." Pp. 247-268 in Gary A. Bernfeld, David P. Farrington, and Alan W. Leschied (eds.), *Inside the "Black Box" in Corrections*. Chichester, UK: John Wiley and Sons.
- Gendreau, Paul, Tracy Little, and Claire Goggin. 1996. "A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works?" *Criminology* 34:575-607.
- Gendreau, Paul, Paula Smith, and Claire Goggin (2001). "Treatment Programs in Corrections." Pp. 238-263 in John Winterdyk (ed.), *Corrections in Canada: Social Reaction to Crime*. Toronto, Canada: Prentice-Hall.
- Gibbs, John C., Granville Bud Potter, and Arnold P. Goldstein. 1995. *The EQUIP Program: Teaching Youths to Think and Act Responsibly Through a Peer-Helping Approach*. Champaign, IL: Research Press.
- Goldstein, Arnold P. 1999. *The Prepare Curriculum: Teaching Prosocial Competencies*. Rev. ed. Champaign, IL: Research Press.
- Gordon, Donald A., Karen Graves, and Jack Arbuthnot. 1995. "The Effect of Functional Family Therapy for Delinquents on Adult Criminal Behavior." *Criminal Justice and Behavior* 22:60-73.
- Henggeler, Scott W., with the assistance of Sharon R. Mihalic, Lee Rone, Christopher Thomas, and Jane Timmons-Mitchell. 1998. *Multisystemic Therapy: Book Six in the Blueprints in Violence Prevention Series*. Boulder, CO: Center for the Study and Prevention of Violence, University of Colorado.
- Hunt, Morton. 1997. *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage Foundation.
- Latessa, Edward J. 2002. "Using Assessment to Improve Correctional Programming: An Update." Unpublished paper, University of Cincinnati.
- Lipsey, Mark W. and David B. Wilson. 1998. "Effective Intervention for Serious Juvenile Offenders." Pp. 313-345 in Rolf Loeber and David P. Farrington (eds.), *Serious and Violent Juvenile Offenders: Risk Factors and Successful Intervention*. Thousand Oaks, CA: Sage.
- MacKenzie, Doris Layton. 2000. "Evidence-Based Corrections: Identifying What Works." *Crime and Delinquency* 46:457-471.
- MacKenzie, Doris Layton, David B. Wilson, and Suzanne B. Kider. 2001. "The Effects of Correctional Boot Camps on Offending." *Annals of the American Academy of Political and Social Science* 578 (November):126-143.

- Matthews, Betsy, Dana Jones Hubbard, and Edward J. Latessa. 2001. "Making the Next Step: Using Assessment to Improve Correctional Programming." *Prison Journal* 81:454-472.
- Millenson, Michael L. 1997. *Demanding Medical Excellence: Doctors and Accountability in the Information Age*. Chicago: University of Chicago Press.
- Rhine, Edward E. (ed.). 1998. *Best Practices: Excellence in Corrections*. Lanham, MD: American Correctional Association.
- Starr, Paul. 1982. *The Social Transformation of American Medicine: The Rise of a Sovereign Profession and the Making of a Vast Industry*. New York: Basic Books.
- Timmermans, Stefan and Alison Angell. 2001. "Evidence-Based Medicine, Clinical Uncertainty, and Learning to Doctor." *Journal of Health and Social Behavior* 42:342-359.
- Van Voorhis, Patricia. 1987. "Correctional Effectiveness: The High Cost of Ignoring Success." *Federal Probation* 51 (March):59-62.
- Van Voorhis, Patricia, Francis T. Cullen, and Brandon K. Applegate. 1995. "Evaluating Interventions with Violent Offenders: A Guide for Practitioners and Policymakers." *Federal Probation* 59 (June):17-28.
- Welsh, Brandon C. and David P. Farrington. 2001. "Toward an Evidence-Based Approach to Preventing Crime." *Annals of the American Academy of Political and Social Science* 578 (November):158-173.

What Doesn't Work in the "What Works" Approach

By Robert A. Shearer, Ph.D.

For some time, there has been an ongoing research effort to determine which correctional programs work and do not work with offenders. This research has led to evidence to support several principles of intervention with offenders. This "what works" or Canadians' Theory (Gendreau, 1996; Cullen & Gendreau, 2001; and Cullen, 2002) has made an impressive contribution to the study of offender treatment. There will continue to be discussion and empirical study of the "what works" approach, but we are beginning to get a better grasp of what works, what doesn't work, and how effective programs work when they do work. Specifically, Latessa, Cullen, and Gendreau, (2002) have identified three important bodies of knowledge that have improved our ability to design effective interventions in corrections. First, we now have better indicators or predictors of recidivism. Second, we know more about what doesn't work. Third, we know more about what does work.

Latessa and Holsinger (1998) have summarized the original principles identified by Gendreau (1996) and these principles can be found throughout the current literature of correctional treatment. A summary of these principles includes the following:

- Services should be intensive, behavioral in nature, and focused on higher risk offenders.
- Behavioral strategies should be enforced in a firm but fair manner by qualified staff.
- Programs should target criminogenic needs.
- Responsivity should occur between staff, offenders, and programs.
- Programs should disrupt criminal networks.
- Programs should provide relapse prevention in the community.

- High levels of advocacy and brokerage should be provided.

These principles tend to drive the "what works" enthusiasm in current correctional treatment literature. In fact, the suggestion has been made that programs that are not



evidence based, are promoting correctional quackery (Latessa, Cullen, & Gendreau, 2002).

Indeed, strong support for the "what works" approach has emerged in the United Kingdom. The National Probation Service has moved quickly to an evidence-based practice based on the principles of "what works," with a strong emphasis on program evaluation (Underdown, 1998; Chapman & Hough, 1998; Boateng, 1999; Slater, 2002; and Slater, 2003).

An interesting aspect of these developments is the recent discovery of principles that have been known in the counseling profession for quite some time. First, there has to be theoretical integrity: or the counseling must be consistent with theory. It is

not surprising that Latessa and Holsinger (1998) found weak programs to be atheoretical. Second, a base rate of behavior needs to be determined. Third, there must be feasibility and fidelity of implementation. And finally, there needs to be an evaluation of whether the counseling was effective. The reason why it has taken corrections so long to discover these basic principles is an interesting line of inquiry, but it will be reserved for another discussion at another time.

What doesn't work in the "what works" approach can be found in the third principle of fidelity of implementation. From a general standpoint, the "what works" approach typically calls for champagne programs when only beer budgets exist, because the level of technical sophistication carries a high pricetag. It isn't a problem of whether it works, but one of feasibility of implementation across several requirements of the approach.

The discussion that follows identifies some of the specific elements that do not work in the "what works" approach. This is not meant to be a shopping list of excuses for poorly designed, implemented, and evaluated programs. Rather, it is an indication of where fitting the "what works" approach into specific programs would be counter-indicated. Rather than excuses or rationalizations, the "what works" approach must be questioned in light of current inescapable political, economic, and workforce realities. The "what works" approach was never sold as a panacea. It can work for some programs in some places, but it also doesn't work in some situations. McLellan (2002) indicates that the true test of the effectiveness of a program is in real-world conditions where training, acceptability, cost, and cost reimbursement are considered. He indicates that carefully controlled studies of efficiency are necessary to show that a treatment can work. But, these carefully controlled conditions are not sufficient to show that the component will

work under broader, less controlled, and more complicated real-world conditions."

Screening and Assessment

The pivotal element of the "what works" approach is accurate assessment after an offender completes screening to determine if a problem exists. There needs to be a risk assessment to determine the likelihood of recidivism or relapse. Second, a needs assessment is required to target criminogenic factors, such as antisocial attitudes, negative peer associations, and thought patterns supportive of criminal behavior. Finally, there needs to be a responsivity assessment to determine the offender's motivation level, learning style, and amenability for treatment. These assessments then lead to treatment matching so that offenders can be matched to programs and therapists. These assessments also lead to guidelines for developing comprehensive written treatment plans.

Assessment is the cornerstone of the treatment process and it is the primary reason why implementing the "what works" philosophy does not fit for many correctional programs. First, in some correctional agencies, the large number of offenders makes comprehensive assessment impractical. Shearer and Carter (1999) have identified this reality. In some correctional systems, it is not unusual for the agency to be in-processing 3,000 offenders a month. The time and cost required to conduct assessments in large agencies is prohibitively high because it would be difficult to conduct these assessments in a short time for a small cost. They require lengthy interviews or closely monitored self-report instruments. The psychological dynamics of self-reporting criminal and substance abuse behaviors are much too complex for a few questions, completed in a group setting, to provide reliable and valid indicators of the extent and gravity of these behaviors.

Second, a thorough assessment requires the use of a battery of sophisticated psychometric instruments. Where are community and institutional correctional programs to find a sufficient supply of trained professionals who can administer these instruments? A typical undergraduate degree would not be sufficient. In order to accurately interpret assessments, an individual would need to be thoroughly familiar with

sampling procedures, reliability, validity, and statistical norms. In addition to substance abuse assessments, they would also need to be familiar with assessing for sexual deviance, antisocial personality disorders, mental retardation, and psychopathology. Graduate counseling or social work programs rarely prepare their graduates to conduct assessments in these areas. In addition, only a handful of forensic psychology



programs exist in this country and not all of them provide substance abuse training. In addition, training to be a licensed chemical dependency counselor, in most states, doesn't provide the technical sophistication that is required to conduct assessments. Austin (1998) noted that a large state treatment program did not have requisite trained and experienced counseling staff to implement the large programs. The situation in the state he identified has gotten worse since he reported the dilemma.

Assuming that a sufficient number of trained professionals were available to conduct psychometric evaluations and they were compensated at a competitive wage level, staff turnover among the individuals conducting the assessments would need to be minimal. This is because the offenders would continue to arrive whether or not there were sufficient assessment staff. The offenders must be moved through the process because most intake facilities and

community corrections programs have limited space and a backlog would produce chaos in the system. In large community and institutional correctional programs, the assessment task is problematic, at best, and impossible at worst.

Treatment Matching

A thorough assessment of an offender's risk level, needs, and responsivity is designed to lead to treatment matching. Treatment matching also means that, instead of a one-size-fits-all program or a homogeneous treatment approach, the treatment will be heterogeneous in nature. In other words, different offenders need different programs because of their unique characteristics. The "what works" approach focuses on the need for higher risk offenders to receive the most intensive interventions, and so forth.

At first glance, this would seem to be a logical and common sense approach. On further inspection, it is one of the ways that the "what works" approach may not work well at all because of implicit assumptions that are difficult to accept. Matching offenders on risk, need, and responsivity assumes that there are programs and staff to match to the variety of high/low risk, high/low criminogenic tendencies, and high/low motivation levels. This is a highly unlikely staffing luxury. If gender and ethnicity are added to the equation, treatment matching approaches the impossible.

Implementation

Another key element of the "what works" approach is the recommended implementation of behavioral, social learning, and cognitive behavioral therapies (Latessa, Cullen, & Gendreau, 2002). There is strong empirical support for the effectiveness of these interventions. Furthermore, a number of structured cognitive-behavioral interventions have been developed as a response to the enthusiasm for the findings of many studies supporting the effectiveness of the approach. Cognitive-behavioral programs developed by Ross and Fabiano (1985), *Reasoning and Rehabilitation*; Bairo (1999), *Strategies for Thinking Productively*; Little and Robinson (1988), *Moral Reconciliation Therapy*, and Bush, Glick, and Taymans (1997), *Thinking for A Change* have all shown to be effective treatment programs.

Continued on page 12

What Doesn't Work

Continued from page 11

In order for these programs to be implemented with therapeutic integrity, Gendreau (1996) suggests that therapists have at least an undergraduate degree or equivalent with specific training in theories and treatment of criminal behavior. He also suggests that therapists receive several months of formal in-service training in cognitive/behavioral interventions.

What these suggestions lead to is the conclusion that we have structured programs that work and we need a lot of therapists that can implement the cognitive/behavioral approach with an acceptable level of integrity. Where would we find an adequate number of therapists to implement cognitive-behavioral treatment programs? Undergraduate criminal justice and psychology programs are not producing any. Graduate counseling and social work programs are not producing cognitive-behavioral therapists. Para-professional substance abuse counseling programs are not producing counselors with this theoretical orientation. The net result is that, with the exception of a few forensic psychology programs, virtually no counseling programs, graduate, undergraduate or para-professional, teach theories and procedures of treating criminal behavior. Furthermore, they don't teach specific social learning models of treatment such as therapeutic communities, token economies, or community reinforcement programs. Simpson (2002) indicates that contingency management interventions, motivational enhancement therapy, and cognitive-behavioral therapy are more advanced counselor skills that require intensive training, clinical supervision, and possibly specific credentials. Based on his conclusions, it is not clear what the source and incentive for all of this advanced training will be. This implementation dilemma also leads to what doesn't work in the "what works" approach. It is little surprise that Latessa and Holsinger (1998), and Austin (1998), reported staff training to be inadequate and the quality of the staff to vary greatly.

Furthermore, the prospect of a counselor with an undergraduate degree, or less, counseling substance abusing offenders in

prisons, on community supervision, or in other correctional facilities, is troublesome, if not frightening. Undergraduate programs simply do not provide the education and training for this difficult task. Undergraduate students are at a formative time in their lives, working through a maze of unresolved personal issues, not the least of which are issues of drug and alcohol use or abuse and the legal system in this country. Without running the risk of taking an elitist or traditionalist position, only in our wildest imagination can we expect significant personal change in offenders as a result of a therapist with an undergraduate degree conducting counseling from a scripted cognitive-behavioral program.

The final barrier to implementing cognitive behavioral programs is the very low pay for counselors in many local and state correctional programs. This is known as the *dog catcher effect*. In many state and local judicial districts, the dog catchers and other city and county employees may make substantially more money than substance abuse counselors. The effect is that counselors have left the field for better paying positions such as dog catchers or they become counselors in the private sector. The net effect is extremely high turnover rates and shortages of qualified counselors. In this environment, a sophisticated intervention program ceases to be a reality. Farabee, Prendergast, Cartier, Wexler, Knight, and Anglin (1999) have discussed some of these same issues with implementation. They refer to these problems as barriers to implementing drug treatment programs in corrections. The six they have identified are (a) client/identification, assessment, and referral; (b) recruitment and training of treatment staff; (c) redeployment of correctional staff; (d) over-reliance on institutional versus therapeutic sanctions; (e) aftercare; and (f) coercion. They hoped that their identification of these barriers would encourage administrators to implement more effective treatment models. The question that remains is whether these barriers are insurmountable in a climate of declining correctional budgets and extreme staff shortages.

Multiculturalism

During the past decade when corrections was attempting to determine what works with offenders, the field of counseling and

psychotherapy has been formulating the competencies of a multicultural counselor. Effective counseling must take into account the impact of culture (Corey, 2001). Culture includes ethnic or racial heritage, gender, religion, sexual orientation, physical and mental ability, and socioeconomic status. Culturally competent counselors need the beliefs and attitudes, knowledge, skills, and intervention strategies to work with culturally diverse populations.

Sensitivity to ethnic and gender differences may be assumed in the responsiveness principle. Otherwise, the "what works" approach has been largely silent on the issue of multiculturalism. Studies have shown significant differences in the way women and ethnic groups respond to therapy in correctional settings (Shearer, 2001; Shearer & Ogan, 2002; and Shearer, Myers, and Ogan, 2001).

This provides a slightly different twist to the notion of what doesn't work in the "what works" approach. In the case of multicultural considerations, either the "what works" principles don't apply or the approach has failed to address a very important consideration for treatment intervention. In any case, the question of multiculturalism in correctional treatment is largely unanswered. It is an important question because certain cultural groups value being emotionally reserved or being very selective about sharing personal concerns with strangers or professionals. Graham and Miller (1994) explain this difference as a preference to interact in group activities in which more than one cultural orientation is involved. They have identified this important dimension of culture as "context." High context individuals are more collectivistic and low context individuals are more individualistic. The context in which a person is interacting affects how one relates to others, communicates, and interprets information.

The diversity found in offender treatment groups, including cognitive-behavioral groups, would seem to dictate that the "what works" approach focus on cultural and gender issues. The cognitive-behavioral approach, like most counseling approaches, is very individualistic because it focuses on individual choice, feelings,

decisions, and consequences. How well does this approach work for offenders from different cultural groups? The answer to this question is not yet available, but is a very important one.

Technology Transfer and Other Gaps Cullen (2002), indicates that one of the major challenges in offender rehabilitation is technology transfer. The problem with the current technology in the "what works" approach is that the scientific knowledge that has been produced is not getting to the potential consumers of the new technology. This is, without question, the primary reason the "what works" approach hasn't been implemented very widely. The reasons for this technology transfer gap are complex. Latessa, Cullen, and Gendreau (2002) identify three of these complex reasons. First, there has not been much communication between academic criminologists and correctional practitioners. Second, designing and implementing effective programs is more complicated than conducting research. Finally, the technology transfer gap is a greater problem than any individual or agency can breach. These reasons suggest a larger issue of whether the entire field of corrections can change.

Additional reasons why there is a technology transfer gap are related to problems in continuing communication and confusion. Traditionally, academicians and researchers present their findings in esoteric publications that have an extremely limited audience. Few correctional decision makers read these publications and almost no one in the general public reads them. There are exceptions, but they are rare. In addition, corrections continues to have difficulty with the issue of whether offenders are designated as patients or prisoners (Kassenbaum, Ward, & Wilmer, 1971). There are major philosophical, procedural and professional differences in the two designations. For the "what works" approach, the most important difference is that the designation of patient suggests a high level of professional care. A patient is afforded a care provider who has the primary responsibility for care and the development of a therapeutic alliance. But, a more important difference is that the care provider operates under the direction of specific professional ethical guidelines. The designation of "pris-

oner" typically does not suggest any of these concerns. This dilemma creates a host of prickly issues for correctional treatment such as coerced treatment, therapeutic alliance, confidentiality, informed consent, and self-determination. If offenders are considered patients, then these issues are less of a concern because of built-in safeguards. But, they rarely are considered patients. As long as they are prisoners first and patients second, they will be subject to second class interventions and the question of "what works" becomes less important

There is a gap in the number of required trained professionals and a gap in what they are paid and what they need to be paid if they are going to stay in the field.

because of this status. Consequently, as offenders remain in this treatment purgatory, the "what works" approach, with its sophisticated professional requirements and accompanying high pricetag, will not work. But, the challenge is much greater than technology transfer. We also have a financial, human resources, and paycheck transfer gap. There is a gap between what state and local governments are willing to spend on treatment and what is needed. There is a gap in the number of required trained professionals and a gap in what they are paid and what they need to be paid if they are going to stay in the field. Obviously, these gaps are larger in some places than others. The "what works" approach can work for some programs in some places, but it doesn't work in others.

Which programs are more amenable to the "what works" approach? Lehman, Greener,

and Simpson (2002) and Simpson (2002) have attempted to answer this question by studying organizational readiness for change. Based on a process model of change, they have developed assessment instruments for program staff and directors to describe the climate of an organization. Specifically, they have studied motivation for change, institutional resources, staff attributes, and the organizational climate. The research on organizational readiness is still exploratory and preliminary, but as new data in these areas become available, we will have a much better grasp of not only whether the "what works" approach can work, but also we will know where it won't work, because the program in question is not ready to change.

Conclusions

The appeal in the "what works" literature is for long-run efforts of evaluating, credentialing, and professionalism of corrections. The field of corrections has rarely enjoyed long-run efforts because of the changing legislative climate that makes the field subject to political whims and the associated undependable funding realities. How likely is it that the issue of crime control will cease to be a political issue? Prison warehouses are littered with the residue of short-run programs that were supposed to become long-run and long-run programs that never became short-run. Perhaps, our efforts would be more productive if we would approach the problem of "what works" in the context of chaos theory rather than the theory of scientific criminology. The changing political climate seems to keep most large correctional programs in varying states of chaos that few inside or outside the program understand.

On the other hand, the results of ineffective treatment programs may lead to substantial harm. This treatment harm can lead to several results identified by Latessa and Holsinger (1998). First, the public and the various entities in the criminal justice system will be reluctant to support effective programs. Second, there will continue to be a perpetuation of the search for a magic or quick-fix to the problem of crime. Finally, offenders will continue to receive the blame for ineffective programs. They identify this as the greatest harm, because

Continued on page 14

What Doesn't Work

continued from page 13

laming absolves us from designing and operating effective interactive programs.

If this potential harm is the case, then the harm must be weighed against the harm of not having any programs available for offenders. Not only is it unlikely that there isn't a magic solution to criminal behavior, but also it is unlikely that a broad proclamation that we now know what works will magically produce high quality staff and programs.

What has been lost in the enthusiasm of the "what works" approach is the professional override principle described by Hoge and Andrews (1996). Programs need to consider risk, need, and responsivity and then make appropriate decisions under present conditions. It is quite likely that many programs that are identified as quackery are actually programs that are operating the best they can under extremely limited resources. The question that needs to be answered in addition to "what works" is whether doing something, that doesn't meet the ideal standard, is better than doing nothing.

References

- Austin, J. (1998). The Limits of Prison Drug Treatment. *Corrections Management Quarterly*, 2(4), 66-74.
- Batang, P. (1998). *What Works: Reducing Re-Offending: Evidence-Based Practice*. HM Inspectorate of Probation, London: Home Office.
- Bush, J., Glick, B., & Taymans, J. (1997). *Thinking for a Change*. Longmont, CO: National Institute of Corrections, United States Department of Justice.
- Chapman, T., & Hough, M. (1998). *Evidence Based Practice: A Guide to Effective Practice*. HM Inspectorate of Probation, London: Home Office.
- Corey, G. (2001). *Theory and Practice of Counseling and Psychotherapy* (6th ed.). US: Brooks/Cole.
- Cullen, F. T. (2002). Rehabilitation and Treatment Programs. In J. Q. Wilson & J. Petersilia (Eds.), *Crime: Public Policies for Crime Control* (pp. 253-289). Oakland, CA: ICS Press.
- Cullen, F. T., & Gendreau, P. (2001). From Nothing Works to What Works: Changing Professional Ideology in the 21st Century. *The Prison Journal*, 81(3), 313-338.
- Farabee, D., Prendergast, M., Cartier, J., Wexler, H., Knight, K., & Anglin, M. D. (1999). Barriers to Implementing Effective Correctional Drug Treatment Programs. *The Prison Journal*, 79(2), 150-162.
- Gendreau, P. (1996). The Principles of Effective Intervention with Offenders. In A. T. Harland, (Ed.), *Choosing Correctional Options that Work*, (pp. 117-130) Thousand Oaks: Sage Publications.
- Graham, M., & Miller, D. (1995). Cross Cultural Interactive Preference Profile. In J.W. Pfeiffer (Ed.) 121-129, *The 1995 Annual: Volume 7 Training*, San Diego: Pfeiffer and Company.
- Hoge, R. D. & Andrews, D. A. (1966). *Assessing the Youthful Offender; Issues and Techniques*. New York: Plenum Press.
- Isaro, L. M. (1999). Effects of a Cognitive Restructuring Program on Inmate Institutional Behavior. *Criminal Justice and Behavior*, 26(4), 466-484.
- Kassebaum, G., Ward, D., & Wilmer, D. (1971). *Prison Treatment and Parole Survival: An Empirical Assessment*. (p. 4) New York: John Wiley & Sons, Inc. p. 4.
- Latessa, E. J., Cullen, F. T., & Gendreau, P. (2002). Beyond Correctional Quackery — Professionalism and the Possibility of Effective Treatment. *Federal Probation*, 66(2), 43-39.
- Latessa, E. J. & Holsinger, A. (1998). The Importance of Evaluating Correctional Programs: Assessing Outcome and Quality. *Corrections Management Quarterly*, 2(4), 22-29.
- Lehman, W. E. K., Greener, J. M., & Simpson, D. D. (2002). Assessing Organizational Readiness for Change. *Journal of Substance Abuse Treatment*, 22, 197-209.
- Little, G. L. & Robinson, K. D. (1988). Treating Drunk Drivers with Moral Reconciliation Therapy: A One-Year Recidivism Report. *Psychological Reports*, 64(3), 960-962.
- McLellan, A. T. (2002). Technology Transfer and the Treatment of Addiction: What can Research Offer Practice? [Editorial]. *Journal of Substance Abuse Treatment*, 22, 169-170.
- Ross, R. R., & Fabiano, E. A. (1985). *Time to Think: A Cognitive Model of Delinquency Prevention and Offender Rehabilitation*. Ottawa: Institute of Social Sciences and Arts.
- Shearer, R. A. & Carter, C. R. (1999). Screening and Assessing Substance-Abusing Offenders: Quantity and Quality. *Federal Probation* 63(1), 30-34.
- Shearer, R. A. (2001). *Profiles of Change: Adult Male Offenders in Community Correctional Residential Treatment*. Technical Report, Huntsville, TX: Substance Abuse Treatment Program, Texas Department of Criminal Justice.
- Shearer, R. A., Myers, L., and Ogan, G. (2001). Treatment Resistance and Ethnicity Among Female Offenders in Substance Abuse Treatment Programs. *Prison Journal*, 81(1), 55-72.
- Shearer, R. A. & Ogan, G. (2002). Measuring Treatment Resistance in Offender Counseling. *Journal of Addictions and Offender Counseling*, 22(2), 72-82.
- Simpson, D. D. (2002). A Conceptual Framework for Transferring Research to Practice. *Journal of Substance Abuse Treatment*, 22, 171-182.
- Simpson, D. D. (2002). A Conceptual Framework for Transferring Research to Practice. *Journal of Substance Abuse Treatment*, 22, 171-182.
- Slater, M. (2002). *What Works News*, 9. National Probation Service, London: Home Office.
- Slater, M. (2002/2003). *What Works News*, 11. National Probation Service, London: Home Office.
- Underdown, A. (1998). *Strategies for Effective Offender Supervision: Report of the HMIP What Works Project*. HM Inspectorate of Probation, London: Home Office.

Robert A. Shearer is a Professor, College of Criminal Justice, Sam Houston State University, Huntsville, Texas.



Effects of Correctional Boot Camps on Offending

By DORIS LAYTON MACKENZIE, DAVID B. WILSON,
and SUZANNE B. KIDER

ABSTRACT: A systematic review incorporating meta-analytic techniques of correctional boot camps studies was conducted. An intensive search identified 771 documents of which 144 were deemed potentially relevant, located, and evaluated for eligibility. In 37 documents, 29 studies were judged eligible for inclusion in the systematic review. The 29 studies resulted in 44 samples providing the primary unit of analysis. Quasi-experimental and experimental studies evaluated a residential program with a militaristic environment and compared the recidivism of participants to a comparison group receiving another correctional sanction. In 9 studies, boot camp participants had lower recidivism than did comparison groups; in 8, comparison groups had lower recidivism; and in the remaining studies, no significant differences were found. A meta-analysis found no overall significant differences in recidivism between boot camp participants and comparison samples. Further analyses indicated the results cannot be explained by differences in study methodology, offender characteristics, or boot camp program components.

Doris Layton MacKenzie is the director of the Evaluation Research Group and a professor in the Department of Criminology and Criminal Justice, University of Maryland. Her research interests include evaluation of correctional programs, correctional alternatives, correctional boot camps, self-report criminal activities, and offender behavior.

David B. Wilson is an assistant professor of the administration of justice at George Mason University. His research interests include program evaluation research methodology, meta-analysis, crime and general problem behavior prevention programs, and juvenile delinquency intervention effectiveness.

Suzanne B. Kider is a master's student in the Department of Criminology and Criminal Justice at the University of Maryland. Her main research interests include corrections and probation and parole reentry.

CORRECTIONAL boot camps, also called shock or intensive incarceration, are short-term incarceration programs modeled after basic training in the military (MacKenzie and Parent 1992; MacKenzie and Hebert 1996). Participants are required to follow a rigorous daily schedule of activities including drill and ceremony and physical training. They rise 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 intake 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.

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, while others focus on discipline and rigorous physical training. Programs also differ in whether they are designed to be alternatives 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 reentry program designed to assist the participants with adjustment to the community.

Correctional boot camps were first opened in adult correctional systems in the United States in 1983, in Georgia and Oklahoma. Since that time they have rapidly grown, first within adult correctional systems and later in juvenile corrections. Today, correctional boot camps exist in federal, state, and local juvenile and adult jurisdictions in the United States. Juvenile boot camps developed later than the adult camps. However, during the 1990s camps for juveniles rapidly developed, and by 2000, 70 juvenile camps had been opened in the United States (see the Koch Crime Institute Web site at www.kci.org). The camps for adjudicated juveniles differ somewhat from the adult camps. In juvenile camps, less emphasis is placed on hard labor, and as required by law, the camps offer 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 intake procedures, shaved heads, drill and ceremony,

physical training, immediate physical punishment for misbehavior (for example, push-ups), and graduation ceremonies.

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 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 et al. 1990; Gendreau, Little, and Goggin 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 participants' attitudes, attachments to the community, or impulsivity (MacKenzie et al. 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 (MacKenzie 1997).

According to a survey of state correctional officials, the major goals of the camps are to deter future crime,

protect the public, rehabilitate the offenders, reduce costs, and lower recidivism (Gowdy 1996). Thus, except for reducing the costs of corrections, all of the major goals are associated in some way with reducing the criminal activities of participants. Sufficient time has now elapsed since the beginning of these camps so that a body of research examining the impact of the camps on the recidivism of participants has been produced. This systematic review is designed to examine this research in order to draw conclusions regarding what is currently known about the effectiveness of correctional boot camps in reducing recidivism.

METHOD

Search strategy and eligibility criteria

The scope of this review was experimental and quasi-experimental evaluations that examined boot camp and boot camp-like programs for juvenile and adult offenders. To be eligible to be included in the review a study had to (1) examine a residential program that incorporated a militaristic environment (the programs were called by various names such as boot camp, shock incarceration, and intensive incarceration); (2) include a comparison group that received either community supervision (for example, probation) or incarceration in an alternative facility such as jail, prison, or juvenile residential facility; (3) include participants who were convicted or adjudicated; and (4) report a

postprogram 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). The comparison group in a quasi-experimental design had to be selected to be reasonably similar to the experimental group; thus any study that compared the experimental group to a general national or state sample was eliminated from the study. Furthermore the study eligibility criteria eliminated quasi-experimental designs that only compared program dropouts to program completers.

The strategies used to identify all studies, published or otherwise, that met these criteria included a keyword search of computerized databases and contact with authors working in this area. The following databases were searched: Criminal Justice Periodical Index, Dissertation Abstracts Online, Government Publications Office Monthly Catalog, Government Publications Reference File, National Criminal Justice Reference Service, 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 indexed unpublished works. This identified 771 unique documents. Review of the titles and abstracts suggested that 152 might meet the above criteria or were relevant review articles that might contain additional references. Of these 152, 144 were obtained and evaluated for eligibility, resulting in 29 eligible studies reported in 37

documents (see references). The majority of these studies were state or federal technical reports ($n = 22$). Only 9 of these studies were published in peer-reviewed journals. One study was conducted in Canada, and another study was conducted in England. The remaining studies evaluated boot camp programs in the United States.

Data collection and analysis

The coding protocol developed for the synthesis allowed for the coding of multiple samples from a single study (distinct evaluations reported in a single report, different cohorts or data reported for males and females separately). This resulted in 44 distinct samples, and these samples represent the primary unit of analysis for this systematic review. The coding protocol also allowed for the coding of multiple indicators of criminal involvement, such as arrest, conviction, and technical violation, measured at multiple time points following release from the program. A copy of the coding protocol can be obtained from the authors. All studies were double coded, and any discrepancies in the coding between the two coders were resolved.

The protocol captured aspects of the research design, including methodological quality, characteristics of the boot camp program, comparison group condition, study participants, outcome measures, and direction and magnitude of the observed effects. The primary effect of interest was recidivism or a return to criminal activity on the part of the offender after leaving the program. Recid-

ivism data were reported dichotomously across all studies and were based on official records, generally reflected as arrest, reconviction, or reinstitutionalization. As such, the natural index of effectiveness is the odds ratio (see Fleiss, 1994) and was the index of effect (see below). The mean odds ratio and homogeneity of effects across studies was computed using the inverse variance weight method. A random-effects model was assumed, and the random-effects variance component was estimated using the methods outlined by Dersimonian and Laird (1986) and Raudenbush (1994). The computations were performed using macros written by the second author that are available for use with SAS, SPSS, and Stata (Lipsey and Wilson 2001).

A total of 155 recidivism effect sizes were extracted from the studies. Recidivism effects that reflected technical violations only were excluded from the analyses reported below, reducing the set of effect sizes to 142. The recidivism effects were examined in two ways. First, multiple recidivism effects from a single study and sample were averaged prior to analysis, producing a set of 44 recidivism effect sizes for the analysis. The second set of analyses used arrest as the measure of recidivism if it was available; if not, reconvictions were used as the measure, and if neither of these was available, reinstitutionalizations were used. The results from the two methods of measuring recidivism were compared and did not yield any substantive differences in the results. Therefore, results based on the second method of measuring

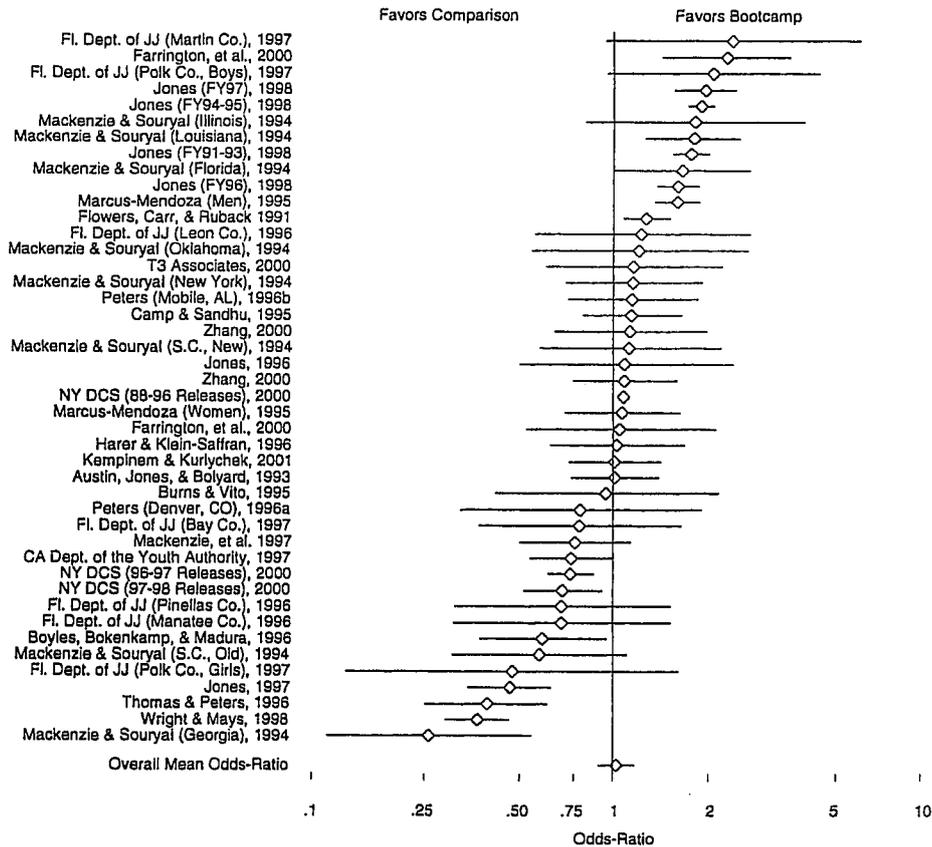
recidivism are reported in the following analyses.

RESULTS

The distribution of recidivism effects across the 44 boot camp versus comparison group samples is shown in Figure 1. Each row of this forest plot represents a distinct sample, identified by the label in the left column. The recidivism odds ratio (effect size) is represented by the small diamonds, and the line spans the 95 percent confidence interval around the odds ratio. The samples are sorted with the largest positive effect at the top and the smallest negative effect (odds ratios between 1 and 0) on the bottom. At the very bottom of the plot is the overall random-effects mean odds ratio.

The effects across these studies ranged from large reductions to large increases in the risk of recidivating for the boot camp participants relative to the comparison groups. The overall mean odds ratio was 1.02 (95 percent confidence interval of 0.90 to 1.17), indicating an almost equal odds of recidivating between the boot camp and comparison groups, on average. Thus there appears to be no relationship between program participation (boot camp or comparison) and recidivism. The equivalent recidivism rates for the average boot camp and comparison group, given this overall odds ratio, would be 49.4 percent for the boot camp and 50 percent for the comparison condition. This is a small difference by most any standard. Thus, overall, the evidence suggests that boot camps do not reduce the risk of recidivism relative to

FIGURE 1
A FOREST PLOT SHOWING THE RECIDIVISM ODDS RATIOS AND
95 PERCENT CONFIDENCE INTERVAL FOR EACH STUDY AND
SAMPLE AND THE OVERALL MEAN ODDS RATIO



other existing criminal justice system forms of punishment and rehabilitation. From the forest plot, it is also evident that 9 studies observed a statistically significant positive benefit of boot camps, whereas 8 studies observed a statistically significant positive benefit of the comparison condition. The remaining 27 studies found no significant differences

between the boot camp samples and the comparisons.

The distribution of odds ratios was highly heterogeneous, $Q = 464.6$, $df = 43$, $p < .0001$, suggesting the presence of moderators of the effects, either methodological or substantive, such as the nature of the boot camp program and comparison conditions and the types of offenders served.

TABLE 1
 CROSS-TABULATION OF QUALITATIVE METHODOLOGICAL
 QUALITY SCORE AND OTHER METHOD DESCRIPTORS (N = 44)

Method Variable	Qualitative Methodological Quality Score		
	4 (n = 19)	3 (n = 17)	2 (n = 8)
Randomly assigned participants to conditions			
Yes	4 (21)	1 (6)	0 (0)
No	15 (79)	16 (94)	8 (100)
Used group-level matching**			
Yes	14 (74)	5 (29)	1 (13)
No	5 (26)	12 (71)	7 (87)
Prospective research design**			
Yes	17 (89)	9 (53)	6 (75)
No	2 (11)	8 (47)	2 (25)
Used statistical controls in analyses**			
Yes	13 (68)	3 (18)	1 (13)
No	6 (32)	14 (82)	7 (87)
Boot camp dropouts in analysis**			
Yes	9 (47)	9 (53)	0 (0)
No	10 (53)	8 (47)	8 (100)
Overall attrition apparent			
Yes	3 (16)	2 (12)	1 (12)
No	16 (84)	15 (88)	7 (88)
Differential attrition apparent			
Yes	3 (16)	3 (18)	2 (25)
No	16 (84)	14 (82)	6 (75)

NOTE: Percentages are in parentheses.

** $p < .05$, based on a chi-square test.

Possible moderating effects are explored below.

Methodological characteristics of the studies

Any conclusion regarding the effectiveness (or ineffectiveness, as the data suggest) of boot camps relative to more traditional correctional approaches in reducing the risk of recidivism is valid only if the methodological quality of this collection of studies is sufficiently high. Table 1 displays the frequency of studies with various methodological characteristics by our qualitative methodological rating scale. This scale was

developed by Sherman and colleagues (1997) and has five levels of methodological rigor. The lowest level of methodological quality was excluded from this synthesis and reflects studies without a comparison group. The highest level of methodological rigor (level 5) represents randomized designs that are not compromised through attrition or other common problems in carrying out a randomized evaluation study.

As can be seen in Table 1, none of the five randomized evaluations included in this synthesis were granted a method quality score of 5. This was generally because the stud-

ies had high attrition or excluded program dropouts from the recidivism analysis, creating a potential threat from selection bias. Thus there were no evaluations of the effectiveness of boot camps that were free from methodological blemishes. That said, however, many of the studies (19 of 44, or 43 percent) were judged to be methodologically solid (method score of 4). These studies were generally the higher-quality quasi-experimental designs that either carefully selected the comparison group so as to maximize similarity with the boot camp group (for example, selecting boot camp eligible offenders and matching the groups on demographic characteristics) or used statistical controls in the analysis of recidivism effects. Only 8 of the 44 evaluations (18 percent) were judged to be of poor methodological quality.

To assess the robustness of the general finding of no effect, a separate mean odds ratio was computed for each category of the different methodological variables (see Table 2). The mean effect size was slightly lower for the studies judged to be of overall higher methodological quality, although the trend was statistically nonsignificant. Studies that used a prospective research design had observed larger positive effects (although not significantly different from a null odds ratio of 1) than did retrospective designs. That is, while the mean odds ratio of prospective and retrospective designs are significantly different from each other, neither design produces an odds ratio that suggests that the experimental and control samples are significantly

different from each other (for example, confidence interval includes 1). In contrast to studies that did not use statistical controls in the analysis of recidivism outcomes, studies that used controls observed smaller effects that were negative in direction. Once again, neither category differed significantly from the null hypothesis. All other methodological variables were unrelated to the observed odds ratios.

Offender characteristics across studies

There was generally little information regarding the characteristics of the offenders in the studies. For 11 of the 44 samples, the authors did not indicate the gender, although it is reasonable to assume that in these cases the samples were all male. Only 3 of the 44 samples were all female, and the mean odds ratio for these samples was 1.06 and statistically nonsignificant. This mean odds ratio is roughly the same as that for the overall sample. Four samples were mixed gender, although they were predominantly male (equal to or greater than 80 percent). Thus there are insufficient data to adequately explore whether boot camps are differentially effective for males and females, as some theorists have hypothesized (Morash and Rucker 1990).

All samples were successfully classified as either juvenile or adult. The adult samples were typically young adults and in some cases included at least a small percentage of juveniles who were adjudicated as adults. As shown in Table 3, the mean odds ratio for the studies evaluating

TABLE 2
 MEAN ODDS RATIO AND 95 PERCENT CONFIDENCE
 INTERVAL BY METHOD VARIABLES (N = 44)

Method Variable	Mean Odds Ratio	95 Percent Confidence Interval		k ^a
		Lower	Upper	
Qualitative methodological quality score				
Random assignment, not degraded				0
High-quality quasi-experiment	0.92	0.73	1.15	19
Standard quasi-experiment	1.07	0.85	1.34	17
Poor-quality quasi-experiment	1.15	0.84	1.59	8
Randomly assigned participants to conditions				
Yes	0.75	0.48	1.17	5
No	1.06	0.91	1.24	39
Used group-level matching				
Yes	1.11	0.88	1.40	20
No	0.97	0.80	1.17	24
Prospective research design**				
Yes	1.13	0.95	1.34	32
No	0.83	0.65	1.06	12
Used statistical controls in analyses**				
Yes	0.85	0.68	1.07	17
No	1.14	0.96	1.37	27
Boot camp dropouts in analysis				
Yes	1.03	0.82	1.28	18
No	1.02	0.83	1.24	26
No overall attrition apparent				
Yes	1.06	0.91	1.24	39
No	0.72	0.46	1.14	5
No differential attrition apparent				
Yes	1.03	0.87	1.21	36
No	0.96	0.67	1.41	8

a. k = number of samples included in analysis.

**p < .05.

the effectiveness of juvenile boot camps was lower than that of the studies evaluating adult (often young adult) boot camps, although this difference was not statistically significant. This difference may reflect a difference in the typical comparison group for juveniles relative to adults. Traditional juvenile detention facilities are qualitatively different from adult prison or adult probation, the common comparison groups for the studies of adult boot camps.

Juvenile detention facilities are more likely, although not guaranteed, to have a greater emphasis on rehabilitation than their adult counterparts. Unfortunately, the availability of rehabilitative treatment within the comparison facilities was not reported by the primary studies.

The racial/ethnic makeup of the offender populations and the offender risk level were often unreported, with no information available for 9 of the 44 samples (20 percent).

TABLE 3
MEAN ODDS RATIO AND 95 PERCENT CONFIDENCE
INTERVAL BY OFFENDER CHARACTERISTICS (N = 44)

Offender Characteristic	Mean Odds Ratio	95 Percent Confidence Interval		k ^a
		Lower	Upper	
Age group of offender				
Juvenile	0.88	0.68	1.14	16
Adult	1.09	0.92	1.30	28
Offender type				
Juveniles				
Nonviolent/nonperson crimes	0.92	0.61	1.38	4
Mixed (violent and nonviolent) crimes	0.85	0.65	1.11	12
Adults				
Nonviolent/nonperson crimes	1.17	0.92	1.50	13
Mixed (violent and nonviolent) crimes	1.01	0.79	1.31	15

a. k = number of samples included in analysis.

For an additional 8 samples, only the percentage of African Americans was reported. Thus roughly half of the samples had complete racial/ethnic makeup information. In general, African Americans were the predominant racial group, representing roughly 52 percent of the samples reporting this information. Caucasians represented 23 percent of the 24 samples, and Hispanics represented roughly 9 percent of the 21 samples reporting these data. The data did not lend themselves to an analysis of the relationship between racial/ethnic makeup of the samples and the observed odds ratios.

*Programmatic differences
across studies*

Boot camps vary in the emphasis placed on rehabilitative treatment relative to physical exercise and military drill and ceremony. It has been speculated that the greater the emphasis on treatments, such as

drug abuse counseling, vocational education, and aftercare transition assistance, the greater the likelihood that boot camps will have positive benefits relative to alternative correctional approaches, such as prison and probation. To assess this issue, we coded whether the evaluation report described the boot camp program as providing various rehabilitative programs listed in Table 4. Mean odd ratios were computed separately for juvenile and adult programs.

The only program characteristic that showed a strong relationship to the effectiveness of the boot camp programs was the presence of an aftercare treatment component for the adult programs. The 11 odds ratios for boot camps with an aftercare component versus comparison group contrasts had a mean of 1.46 with a 95 percent confidence interval that did not include 1, indicating a statistically significant positive effect. This evidence suggests that

TABLE 4
 MEAN ODDS RATIO AND 95 PERCENT CONFIDENCE INTERVAL BY
 PROGRAM CHARACTERISTICS (JUVENILES $n = 16$, ADULTS $n = 28$)

Program Characteristic	Mean Odds Ratio	95 Percent Confidence Interval		k^a
		Lower	Upper	
Aftercare treatment component				
Juveniles				
Yes	0.88	0.70	1.12	14
No	0.79	0.44	1.43	2
Adults***				
Yes	1.46***	1.14	1.87	11
No	0.89	0.72	1.10	17
Academic education				
Juveniles				
Yes	0.88	0.68	1.14	16
No				0
Adults				
Yes	1.13	0.93	1.38	24
No	0.86	0.51	1.43	4
Vocational education				
Juveniles				
Yes	0.98	0.62	1.55	3
No	0.84	0.66	1.08	13
Adults*				
Yes	0.82	0.56	1.20	6
No	1.17*	0.97	1.43	22
Drug treatment				
Juveniles				
Yes	0.90	0.70	1.15	12
No	0.78	0.49	1.24	4
Adults				
Yes	1.08	0.88	1.33	22
No	1.12	0.73	1.72	6
Counseling (group and individual)				
Juveniles				
Yes	0.91	0.70	1.17	10
No	0.79	0.52	1.18	6
Adults				
Yes	1.17	0.95	1.44	21
No	0.85	0.58	1.26	7
Manual labor				
Juveniles				
Yes	1.03	0.73	1.44	7
No*	0.79	0.61	1.02	9
Adults				
Yes	1.07	0.88	1.31	24
No	1.22	0.73	2.04	4

a. k = number of samples included in analysis.

* $p \leq .10$. *** $p \leq .01$.

aftercare may be important in reducing the risk of recidivism, at least for adult samples.

A counterintuitive finding is the negative relationship between vocational education and odds ratio for the adult samples. Study samples with vocational education had a lower mean odds ratio than did those without. The number of boot camp programs with vocational education was small, however, raising the possibility that this relationship is confounded with other study differences.

Multivariate analysis of effect size and study characteristics

The simple univariate analyses of the relationships between odds ratios and study characteristics do not take into account the possible confounding of study features. To assess this possibility, a mixed-effects regression model (see Lipsey and Wilson 2001; Raudenbush 1994) was estimated, regressing the logged odds ratios onto study features. The basic model included the major methodological features, accounting for significant variability in odds ratios across studies, $R^2 = .28$, $Q = 16.19$, $df = 7$, $p = .02$. Significant variability remained, however, after accounting for methodological differences. Building on this basic methods model, separate regression analyses were run for each major program characteristic shown in Table 4. Because of the possibility of an interaction between program characteristics and offender age, these models were run separately for juveniles and adults. The finding of a positive benefit from aftercare for the adult offenders remained statistically significant

after adjusting for methods features. The counterintuitive finding regarding vocational education was not robust to method difference; that is, it was statistically nonsignificant once conditioned on method features. This reinforces our hunch that this finding was the result of a confounding of study features and not due to any negative effects of vocational education. No new significant study characteristics emerged in the multivariate analyses.

DISCUSSION
AND CONCLUSION

In our overall meta-analysis of recidivism, we found no differences between the boot camp and comparison samples. Our analysis predicts that if the comparison sample's recidivism is estimated to be 50 percent, the boot camp sample's recidivism would be estimated to be 49.4 percent, or only 0.6 percent lower. When the individual studies were examined, no significant differences were found between the boot camp samples and the comparisons in the majority of the studies. In only 17 samples out of the total of 44, a significant difference between the experimental and control samples was found; approximately half favored the boot camp while the remaining favored the comparisons. Thus, by whatever criteria are used, there is no evidence that the boot camps reduce recidivism.

The results of this systematic review and meta-analysis will be disappointing for many people. Advocates of the programs expect them to successfully reduce the future

criminal activities of adults and juveniles. Critics argue that the programs are poorly conceived as therapeutic interventions, they will not reduce recidivism, and they may actually have the opposite effect by increasing criminal activities. Our results do not support either side of this argument because we found no differences in recidivism between the 44 boot camp samples and the comparisons. Correctional boot camps are neither as good as the advocates assert nor as bad as the critics hypothesize.

An examination of the forest plot of the individual studies (see Figure 1) and our analysis of the data demonstrated large differences in the studies in terms of the effect of boot camps. Some studies found boot camp participants did better than the comparisons, and others found comparison samples did better. For this reason, we explored whether the differences among studies could be attributed to the methods or design of the studies or to characteristics of the programs or individual participants. In our examination of the methodological variables, we did not find any evidence that differences in the results of studies could be explained by the study methodology.

Our examination of the offender characteristics was disappointing because very few studies reported sufficient information to enable us to code and analyze the possible impact of these characteristics on study outcomes. Few studies even reported on the gender of the samples. The only variables we could examine were (1) whether the studies focused on adult offenders or adjudicated juveniles,

and (2) whether the participants were limited to those convicted or adjudicated for nonviolent/nonperson crimes or mixed violent and nonviolent crimes. Again we found no evidence that differences in these characteristics explained the differences in the results.

We were able to code and analyze the possible impact of six program characteristics, including whether the boot camps had aftercare, academic education, vocational education, drug treatment, counseling, or manual labor components. It is important to note that this information was limited to general information about the characteristics of the programs. We assume the quality and intensity of the programs differed greatly. From our knowledge of the boot camps we know that some programs consider Narcotics Anonymous or Alcoholics Anonymous meetings drug treatment, whereas others provide a more intensive drug treatment experience using a Therapeutic Community-type model. We did not have enough information to code such differences. Almost no information was given about what happened to the comparison samples. The potential impact of these differences on recidivism cannot be overlooked.

When we examined the impact of program characteristics, the only differences we found were for adult studies and, after controlling for methodological differences, the only difference was for boot camps that included an aftercare component. In other words, whereas the odds ratios differed for boot camps with and without aftercare, in neither case did the boot camp samples differ

significantly from the comparisons. While the recidivism of releasees from boot camps with aftercare differed from the recidivism of releasees from boot camps without aftercare, there were no significant differences in recidivism between boot camp releasees and comparisons for either type of boot camp (for example, with or without aftercare). Thus we were unable to identify any characteristic of the methods, offenders, or programs that would explain differences in results of the studies.

Why don't boot camps reduce recidivism when compared to other correctional alternatives? In our opinion, one possible reason boot camps are not any more or less effective than other alternatives is because they may offer no more therapy or treatment than the alternatives. That is, boot camps by themselves have little to offer as far as moving offenders away from criminal activities. Sufficient research currently exists to demonstrate that appropriate correctional treatment with particular characteristics can be effective in changing offenders (Andrews and Bonta 1998; Gendreau and Ross 1987; Lipsey 1992). Some boot camps incorporate this type of treatment and therapy into the regime of the camps, while others do not. Similarly, some comparison facilities or programs provide such treatment. Almost all studies compared offenders or juveniles in boot camps to others in correctional programs within the same jurisdictions. We hypothesize that there are similarities within jurisdictions such that boot camps with therapy and treatment will be located in

jurisdictions that also provide such treatment to those in the comparison programs within the jurisdiction. Thus, in terms of the type of treatment or therapy that has been shown to be effective, correctional programs within the same jurisdictions will be similar. The boot camps may only differ from other correctional programs in the same jurisdiction in the military aspects and not in therapy and treatment. It seems likely that the therapy and treatment are the important components in reducing recidivism. Therefore, since boot camps and other correctional programs provide similar therapy and treatment, the impact on recidivism will be similar.

The research demonstrates that there are no differences in recidivism when boot camp samples are compared to those who receive other correctional sanctions. In our opinion, this can be interpreted to show that a military atmosphere in a correctional setting is not effective in reducing recidivism. However, many questions remain. It would be particularly valuable to have more information about the characteristics of the participants, and the components of the programs, both for the boot camps and for the comparisons. From these studies, we were able to code very little of this information. We anticipate that programs with more treatment and therapy will be more successful in reducing recidivism. The question is whether this would explain some of the differences in results across studies. Future research would greatly benefit by increasing the amount of detailed

information about the programs and the participants.

**APPENDIX
SECONDARY SOURCES USED
IN THE META-ANALYSIS**

1. Burns and Vito (1995)
Burns, Jerald C. 1994. A Comparative Analysis of the Alabama Department of Corrections Boot Camp Program. Unpublished Ph.D. diss., University of Alabama, Tuscaloosa.
2. State of New York Department of Correctional Services (2000)
Courtright, Kevin E. 1991. An Overview and Evaluation of Shock Incarceration in New York State. Unpublished master's thesis, Mercyhurst College, Erie, PA.
3. Marcus-Mendoza (1995)
Holley, Philip D. and David E. Wright. 1995. Oklahoma's Regimented Inmate Discipline Program for Males: Its Impact on Recidivism. *Journal of the Oklahoma Criminal Justice Research Consortium* 2:58-70.
4. Kempinem and Kurlychek (2001)
Kempinem, Cynthia and Mark Motivans. 1998. Who Goes to Pennsylvania's Boot Camp? Paper presented at the meeting of the American Society of Criminology, Washington, DC, Nov.
5. Harer and Klein-Saffran (1996)
Klein-Saffran, Jody. 1991. Shock Incarceration, Bureau of Prisons Style. *Research Forum* 1:1-9.
6. MacKenzie and Souryal (1994)
MacKenzie, Doris L., Robert Brame, David McDowall, and Claire Souryal. 1995. Boot Camp Prisons and Recidivism in Eight States. *Criminology* 33:327-57.
- MacKenzie, Doris Layton, James W. Shaw, and Voncile B. Gowdy. 1990. *Evaluation of Shock Incarceration in Louisiana, Executive Summary*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
7. Peters (1996a), Peters (1996b), and Thomas and Peters (1996)
Peters, Michael, David Thomas, and Christopher Zamberlan. 1997. *Boot Camps for Juvenile Offenders: Program Summary*. Rockville, MD: U.S. Department of Justice, National Institute of Justice.

NOTE: Secondary sources are shown after the primary sources included in the reference list.

References

- Andrews, Donald A. and James Bonta. 1998. *The Psychology of Criminal Conduct*. Cincinnati, OH: Anderson.
- Andrews, Donald A., Ivan Zinger, Robert D. Hoge, James Bonta, Paul Gendreau, and Frances T. Cullen. 1990. Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis. *Criminology* 28:369-404.
- *Austin, James, Michael Jones, and Melissa Bolyard. 1993. *Assessing the Impact of a County Operated Boot Camp: Evaluation of the Los Angeles County Regimented Inmate Diversion Program* (NCJRS document reproduction service no. 154401). San Francisco: National Council on Crime and Delinquency.

*References marked with an asterisk indicate studies included in the meta-analysis. (See appendix for secondary sources used in the meta-analysis showing primary resource.)

- *Boyles, Cecilia E., Eric Bokenkamp, and William Madura. 1996. *Evaluation of the Colorado Juvenile Regimented Training Program*. Golden: Colorado Department of Human Services, Division of Youth Corrections.
- *Burns, Jerald C. and Gennaro F. Vito. 1995. An Impact Analysis of the Alabama Boot Camp Program. *Federal Probation* 59:63-67.
- *California Department of the Youth Authority. 1997. *LEAD: A Boot Camp and Intensive Parole Program; the Final Impact Evaluation* (Report to the California Legislature). Sacramento: Author.
- *Camp, David A. and Harjit S. Sandhu. 1995. Evaluation of Female Offender Regimented Treatment Program (FORT). *Journal of the Oklahoma Criminal Justice Research Consortium* 2:50-57.
- Clark, Cheryl L. and David W. Aziz. 1996. Shock Incarceration in New York State: Philosophy, Results, and Limitations. In *Correctional Boot Camps: A Tough Intermediate Sanction*, ed. Doris Layton MacKenzie and Eugene E. Hebert. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Dersimonian, Rebecca and Nan Laird. 1986. Meta-Analysis in Clinical Trials. *Controlled Clinical Trials* 7:177-88.
- *Farrington, David P., G. Hancock, M. Livingston, Kate Painter, and G. Towl. 2000. *Evaluation of Intensive Regimes for Young Offenders* (Home Office research findings). London: Home Office Research, Development and Statistics Directorate.
- Fleiss, Joseph L. 1994. Measures of Effect Size for Categorical Data. In *The Handbook of Research Synthesis*, ed. Harris Cooper and Larry V. Hedges. New York: Russell Sage.
- *Florida Department of Juvenile Justice. 1996a. *Leon County Sheriff's Department Boot Camp: A Follow-Up Study of the First Five Platoons*. Tallahassee: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *———. 1996b. *Manatee County Sheriff's Boot Camp: A Follow-Up Study of the First Four Platoons*. Tallahassee: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *———. 1996c. *Pinellas County Boot Camp: A Follow-Up Study of the First Five Platoons* (research rep. no. 33). Tallahassee: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *———. 1997a. *Bay County Sheriff's Office Boot Camp: A Follow-Up Study of the First Seven Platoons* (research rep. no. 44). Bay County: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *———. 1997b. *Martin County Sheriff's Office Boot Camp: A Follow-Up of the First Four Platoons* (research rep. no. 43). Martin County: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *———. 1997c. *Polk County Juvenile Boot Camp: A Follow-Up Study of the First Four Platoons*. Tallahassee: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *———. 1997d. *Polk County Juvenile Boot Camp—Female Program: A Follow-Up Study of the First Seven Platoons*. Polk County: Florida Department of Juvenile Justice, Bureau of Data and Research.
- *Flowers, Gerald T., Timothy S. Carr, and R. Barry Ruback. 1991. *Special Alternative Incarceration Evaluation*. Atlanta: Georgia Department of Corrections.
- Gendreau, Paul, Tracy Little, and Claire E. Goggin. 1996. A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works! *Criminology* 34:575-607.

- Gendreau, Paul and Robert R. Ross. 1987. Revivication of Rehabilitation: Evidence from the 1980s. *Justice Quarterly* 4:349-408.
- Gowdy, Voncile B. 1996. Historical Perspective. In *Correctional Boot Camps: A Tough Intermediate Sanction*, ed. Doris Layton MacKenzie and Eugene E. Hebert. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- *Harer, Miles D. and Jody Klein-Saffran. 1996. An Evaluation of the Federal Bureau of Prisons Lewisburg Intensive Confinement Center. Unpublished manuscript, Federal Bureau of Prisons, Research and Evaluation, Washington, DC.
- *Jones, Mark. 1996. Do Boot Camp Graduates Make Better Probationers? *Journal of Crime and Justice* 19:1-14.
- *———. 1997. Is Less Better? Boot Camp, Regular Probation and Rearrest in North Carolina. *American Journal of Criminal Justice* 21:147-61.
- *Jones, Robert J. 1998. *Annual Report to the Governor and the General Assembly: Impact Incarceration Program*. Springfield: Illinois Department of Corrections.
- *Kempin, Cynthia A. and Megan C. Kurlychek. 2001. *Pennsylvania's Motivational Boot Camp* (2000 Report to the Legislature). Quehanna: Pennsylvania Commission on Sentencing.
- Lipsey, Mark. 1992. Juvenile Delinquency Treatment: A Meta-Analytic Inquiry into the Variability of Effects. In *Meta-Analysis for Explanation: A Casebook*, ed. Thomas Cook, Harris Cooper, David S. Cordray, Heidi Hartmann, Larry V. Hedges, Richard J. Light, Thomas A. Louis, and Frederick Mosteller. New York: Russell Sage.
- Lipsey, Mark W. and David B. Wilson. 2001. *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.
- MacKenzie, Doris L. 1997. Criminal Justice and Crime Prevention. In *Preventing Crime: What Works, What Doesn't, What's Promising*, ed. Lawrence W. Sherman, Denise C. Gottfredson, Doris Layton MacKenzie, John Eck, Peter Reuter, and Shawn Bushway. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- MacKenzie, Doris Layton and Eugene E. Hebert, eds. 1996. *Correctional Boot Camps: A Tough Intermediate Sanction*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- MacKenzie, Doris Layton and Dale G. Parent. 1991. Shock Incarceration and Prison Crowding in Louisiana. *Journal of Criminal Justice* 19:225-37.
- MacKenzie, Doris L. and Dale G. Parent. 1992. Boot Camp Prisons for Young Offenders. In *Smart Sentencing: The Emergence of Intermediate Sanctions*, ed. James M. Byrne, Arthur J. Lurigio, and Joan Petersilia. Newbury Park, CA: Sage.
- MacKenzie, Doris Layton and Alex Piquero. 1994. The Impact of Shock Incarceration Programs on Prison Crowding. *Crime & Delinquency* 40:222-49.
- MacKenzie, Doris Layton and James W. Shaw. 1990. Inmate Adjustment and Change During Shock Incarceration. *Justice Quarterly* 7:125-50.
- *MacKenzie, Doris L. and Claire Souryal. 1994. *Multi-Site Evaluation of Shock Incarceration: Executive Summary*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- . 1995. Inmate Attitude Change During Incarceration: A Comparison of Boot Camp with Traditional Prison. *Justice Quarterly* 12:325-54.
- *MacKenzie, Doris L., Claire Souryal, Miriam Sealock, and Mohammed Bin Kashem. 1997. *Outcome Study of the Sergeant Henry Johnson Youth Leadership Academy (YLA)*. Washington, DC: University of Maryland, U.S. De-

- partment of Justice, National Institute of Justice.
- MacKenzie, Doris L., David B. Wilson, Gaylene S. Armstrong, and Angela R. Gover. 2001. The Impact of Boot Camps and Traditional Institutions on Juvenile Residents: Perception, Adjustment, and Change. *Journal of Research in Crime and Delinquency* 38:279-313.
- *Marcus-Mendoza, Susan T. 1995. Preliminary Investigation of Oklahoma's Shock Incarceration Program. *Journal of the Oklahoma Criminal Justice Research Consortium* 2:44-49.
- Morash, Merry and Lila Rucker. 1990. A Critical Look at the Idea of Boot Camp as a Correctional Reform. *Crime & Delinquency* 36:204-22.
- *Peters, Michael. 1996a. *Evaluation of the Impact of Boot Camps for Juvenile Offenders: Denver Interim Report*. Fairfax, VA: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- *———. 1996b. *Evaluation of the Impact of Boot Camps for Juvenile Offenders: Mobile Interim Report*. Fairfax, VA: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- Raudenbush, Stephen W. 1994. Random Effects Models. In *The Handbook of Research Synthesis*, ed. Harris Cooper and Larry V. Hedges. New York: Russell Sage.
- Sechrest, Dale D. 1989. Prison "Boot Camps" Do Not Measure Up. *Federal Probation* 53:15-20.
- Sherman, Lawrence W., Denise C. Gottfredson, Doris Layton MacKenzie, John Eck, Peter Reuter, and Shawn Bushway. 1997. *Preventing Crime: What Works, What Doesn't, What's Promising*. Washington, DC: U.S. Department of Justice, National Institute of Justice.
- *State of New York Department of Correctional Services, Division of Parole. 2000. *The Twelfth Annual Shock Legislative Report (Shock Incarceration and Shock Parole Supervision)*. Albany, NY: Division of Parole.
- *Thomas, David and Michael Peters. 1996. *Evaluation of the Impact of Boot Camps for Juvenile Offenders: Cleveland Interim Report*. Fairfax, VA: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- *T3 Associates Training and Consulting. 2000. *Project Turnaround Outcome Evaluation. Final Report*. Ottawa, Canada: Author.
- *Wright, Dionne T. and G. Larry Mays. 1998. Correctional Boot Camps, Attitudes, and Recidivism: The Oklahoma Experience. *Journal of Offender Rehabilitation* 28:71-87.
- *Zhang, Sheldon X. 2000. *An Evaluation of the Los Angeles County Juvenile Drug Treatment Boot Camp (Final Report)*. Washington, DC: U.S. Department of Justice, National Institute of Justice.

Crime & Delinquency

<http://cad.sagepub.com>

Well-Meaning Programs Can Have Harmful Effects! Lessons from Experiments of Programs Such as Scared Straight

Anthony Petrosino, Carolyn Turpin-Petrosino and James O. Finckenauer

Crime Delinquency 2000; 46; 354

DOI: 10.1177/0011128700046003006

The online version of this article can be found at:

<http://cad.sagepub.com/cgi/content/abstract/46/3/354>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

On behalf of:



National Council on Crime and Delinquency

Additional services and information for *Crime & Delinquency* can be found at:

Email Alerts: <http://cad.sagepub.com/cgi/alerts>

Subscriptions: <http://cad.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 5 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://cad.sagepub.com/cgi/content/abstract/46/3/354#BIBL>

Well-Meaning Programs Can Have Harmful Effects! Lessons From Experiments of Programs Such as Scared Straight

Anthony Petrosino
Carolyn Turpin-Petrosino
James O. Finckenauer

Despite their importance in assessing the impact of policies, outcome evaluations—and in particular randomized experiments—are relatively rare. The rationalizations used to justify the absence of outcome evaluations include such assertions as “we know our programs are working,” “they can’t possibly harm anyone,” and “if they only help one kid they’re worth it.” Using preliminary results from a systematic review of nine randomized experiments of the Scared Straight, or prison visitation program, the authors show that a popular and well-meaning program can have harmful effects. They use these results to argue for more rigorous evaluations to test criminal justice interventions.

Many justice programs, policies, and practices are widely disseminated without pilot testing. Exacerbating this problem is that careful studies are not often done to test these interventions after they are implemented. As Fitz-Gibbon (1999) noted about education and Sherman (1984) about policing,¹ the failure to randomize does not mean the government is not experimenting; instead they are conducting uncontrolled experiments every day across a multitude of policy sectors. Though randomized experiments seem to be increasing in criminal justice and other settings (Boruch, Snyder, &

ANTHONY PETROSINO: Research Fellow, Center for Evaluation, Initiatives for Children Program, American Academy of Arts & Sciences. **CAROLYN TURPIN-PETROSINO:** Assistant Professor, Department of Criminal Justice, University of Massachusetts, Lowell. **JAMES O. FINCKENAUER:** Director, International Center, National Institute of Justice.

Grants from the Mellon Foundation to the Center for Evaluation, the Initiatives for Children Program, the American Academy of Arts & Sciences, and the Smith-Richardson Foundation to the University of Pennsylvania supported the first author’s work. We appreciate the comments of Robert Boruch, Robert Brennan, Iain Chalmers, Jane Dennis, Philip Davies, Lynette Feder, Tim Hacsí, Howard Hiatt, Penny Janeway, Geraldine Macdonald, Frederick Mosteller, and Carol Weiss on various aspects of this project. The preliminary work presented here is being done for the Cochrane Collaboration’s Developmental, Psychosocial and Learning Problems Group. The final report has not been completed and submitted to the Cochrane Group for formal editorial review.

CRIME & DELINQUENCY, Vol. 46 No. 3, July 2000 354-379
© 2000 Sage Publications, Inc.

354

DeMoya, 1999; Petrosino & Boruch, in press), the number of experiments relative to all the outcome studies reported is still quite small (Boruch et al., 1999). The number of outcome evaluations relative to all the programs, policies, and practices implemented in just one state jurisdiction must be very small (Petrosino, 1998).

Despite the millions of dollars of public funds that are invested, few outcome evaluations and experiments seem to be conducted. To understand the barriers to rigorous outcome evaluation, Petrosino (1998) conducted personal interviews with the research and evaluation managers employed by seven distinct agencies in a single state. Each agency had a different area of responsibility, including criminal justice, education, public health, community affairs, and drug prevention. Petrosino found that despite the thousands of programs administered by these offices, only two were subjected to outcome evaluation. None were tested using comparison groups. A randomized experiment had not been carried out on any agency program, according to interview participants, for years. When asked why this was the case, the research managers noted the objections of their bosses—the upper level management personnel—to outcome evaluations. Many were listed, but three are worth paraphrasing:

1. We know our programs work; why evaluate them?
2. We know they are not harming anyone, and see number 1 above.
3. If the program helps a single child, it's worth it. Why evaluate?

Failure to evaluate ignores a long history of admonitions about failed policies and the potential for harmful effects. Not only could ineffective programs divert money and attention from more successful interventions, they could also cause more harm than good. A program may certainly help one child but hurt two in the process. For example, Pallone (1986) writes persuasively about the occasional harmful effects of psychotherapy. Galvin (1979) notes that follow-ups over a 30-year period of participants in the Cambridge-Somerville experiment found that children initially exposed to the benevolent counseling condition did much worse on a variety of outcome measures than the no-treatment control children. Well-meaning programs can be harmful, and rigorous evaluation is often the only way to find this out and correct it. As Chalmers (1999) said, the goal of science in the public sector should be to maximize the good and minimize the harm caused by government-imposed programs, policies, and practices.

One of the more egregious examples in the history of potentially harmful justice programs is Scared Straight (Finckenauer, 1982). It is a lesson, though, that seems to be forgotten in light of a new television documentary that promises much ("Kids and Crooks," 1999) and the reinvention of the program in the United States (Finckenauer & Gavin, 1999) and worldwide (e.g., Hall,

1999).² In this article, we provide a brief summary of Scared Straight, describe an ongoing project to systematically review randomized experiments testing the effect of these programs, and present some preliminary findings. We also present one major lesson. Scared Straight, at least from the data presented here, is likely a harmful program that more often than not leads to increased crime and delinquency in our communities. We conclude that rigorous evaluations are needed to identify harmful interventions.

SCARED STRAIGHT

In the 1970s, a group of inmates serving life sentences at a New Jersey prison conducted the Juvenile Awareness Program to deter at-risk or delinquent children from a future life of crime. The program, known as Scared Straight, brought youths to Rahway State Prison to participate in a realistic and confrontational rap session run by prisoners serving life sentences. As the inmates led the rap sessions, they graphically depicted prison life, including stories of rape and murder (Finckenauer, 1982). Deterrence is the theory behind the program; troubled youths would refrain from lawbreaking because they would not want to follow the same path as the inmates and end up in adult prison. The New Jersey Scared Straight program is the most famous of juvenile delinquency prevention programs involving visits to prisons by delinquents. The name *Scared Straight* is also now used generically to describe all prison aversion programs, including those that involve tours or orientation sessions without formal contact with inmates. Nearly all of the earlier programs involved a confrontational presentation by prison inmates. Lundman (1993) reports, however, that the program is now designed to be more educational and less confrontational.

The television documentary on the New Jersey program, titled "Scared Straight!," which won several television and film awards, aired in 1979. It was claimed in the program that 80% of the more than 8,000 juveniles who had been exposed to the program remained law-abiding (Shapiro, 1978). Following the airing of the program, more than 30 states and several foreign countries created, mandated, or legislated similar types of programs in their jurisdictions (Finckenauer, 1980). Corrigan (1979) summarized the reasons for the program's popularity: its "get tough" deterrent approach, its simplicity, its low cost, and its constructive use of prisoners. Media attention and the fit between program and ideological climate also propelled its popularity (Cavender, 1984; Finckenauer, 1982; Heeren & Shicor, 1984). The rapid diffusion of the program led to careful examination and cautions about Scared Straight, issued by such luminaries as the American Justice Institute

(Berkman & Pearson, 1980), the National Center on Institutions and Alternatives (1979), the National Advisory Committee for Juvenile Justice and Delinquency Prevention (1980), and the House Committee on Education and Labor (*Oversight on Scared Straight*, 1979).

As Finckenauer noted (1980, 1982), Scared Straight fits into the usual pattern of the search for simple cures for difficult social problems such as juvenile delinquency. Governments seek a panacea, adopt an intervention for a short time, and when it fails to live up to expectations, the search for another easy cure begins (West, 1981). Two decades later, Finckenauer concluded that the panacea phenomenon was more complicated (Finckenauer & Gavin, 1999). In its first iteration, the implication was that the ostensibly failed panacea would be discarded and rejected. In the newer thesis, this is not necessarily so. Some failed panaceas will survive in spite of the evidence against them. Finckenauer and Gavin's (1999) newer take on the panacea phenomenon seems particularly accurate with Scared Straight. Despite the intensity with which jurisdictions adopted the program, evaluation research, including several randomized experiments, found the program was not effective in reducing crime (e.g., Finckenauer, 1982; Lewis, 1983; Yarborough, 1979).

Nearly every broad synthesis of the crime prevention literature that has included programs such as Scared Straight usually categorizes them with other types of deterrence-oriented programs (e.g., shock probation). Almost universally, these reviewers report no crime reduction effect for Scared Straight and other deterrence-oriented programs (e.g., Lipsey, 1992). In 1997, University of Maryland researchers completed a report for Congress on the evidence for various crime prevention strategies. Although they found evaluation evidence lacking for many areas of intervention, the researchers had no problem listing Scared Straight as one of the programs that "doesn't work" (Sherman et al., 1997). Though the program continued in use worldwide, the enthusiasm that initially greeted interventions such as Scared Straight has waned since the early 1980s.

In 1999, however, the television program "Scared Straight: 20 Years Later," hosted by noted actor Danny Glover, aired in the United States ("Kids and Crooks," 1999). The program followed up on the 17 delinquent children who were the subject of the original documentary and claimed that only 1 became a career criminal. News coverage of the new show proclaimed the program's success. For example, *USA Today* concluded, "The erstwhile delinquents, now in their 30s, testify that the prison encounter deterred them from a life of crime" ("Kids and Crooks," 1999, p. 4D). Indeed, one prison guard is quoted in the film as saying that only 92 of the 500 kids she sent into the program committed new offenses ("Kids and Crooks," 1999). The program's producer, Arnold Shapiro, is also quoted: "You don't know how many people

have come up to me and said, 'I was juvenile delinquent and when I saw this, I stopped, I changed' " (Eicher, 1999, p. F-05).

Most citizens, unaware of studies questioning such programs, believe the program makes intuitive sense (after all, what kid wants to end up in prison?) and is effective. Given the program and its coverage, it was only natural that policy makers would ask whether the program should be part of a government-supported portfolio of delinquency prevention programs. In keeping with the panacea phenomenon, a new generation of legislators looking for more punitive solutions to crime despite falling crime rates—including the rate of juvenile crime (Zimring, 1999)—continue to be interested in reviving programs such as Scared Straight in their jurisdictions (Finckenauer & Gavin, 1999). This also seems to be the case outside of the United States. For example, Australia's Day in Prison Program appeared to have been initiated due to political pressures rather than consideration of the potential for the program (O'Malley, Coventry, & Walters, 1993). From Germany, there are reports of a popular program similar to Scared Straight recently implemented for young offenders with ties to organized hate groups such as the Neo-Nazis, with plans to expand nationwide (Hall, 1999).

Given the renewed interest in programs such as Scared Straight, it seemed sensible to undertake a systematic review of the randomized experimental evidence on the program. Although some Scared Straight program evaluations were included in prior reviews (e.g., Lipsey, 1992; Sherman et al., 1997), no previous attempt to systematically and exclusively review Scared Straight evaluations has been reported.

During 1999, the first two authors initiated a trial run of a systematic review for the newly initiated Campbell Collaboration, an internationally based group that will prepare, maintain, and make accessible systematic reviews of research on the effects of social and educational interventions (see its Web site at <http://campbell.gse.upenn.edu>). They are using the existing infrastructure provided by the Cochrane Collaboration, an international organization that focuses on reviews in health care (see its Web site at <http://www.cochrane.org>). This is being done to get estimates on costs and the time required for Cochrane-type reviews in the social sector, and to see how well the Cochrane software and editorial process handled reviews conducted in fields such as criminal justice. Given the charge for that project, a systematic review of the Scared Straight experiments seemed to be a natural fit. We report on our preliminary findings, cautioning that our results here have not yet gone through the Cochrane Collaboration's rigorous editorial process.

A SYSTEMATIC REVIEW OF SCARED STRAIGHT EXPERIMENTS

Systematic reviews use explicit and well-established methods in synthesizing the results of separate but similar studies (Chalmers & Altman, 1995). Meta-analysis or quantitative methods are often used in systematic reviews but are by no means appropriate in all circumstances. Systematic reviewing methods are designed to reduce the potential biases that can affect conclusions in the synthesis of findings from multiple evaluations. For example, by collecting unpublished studies, reviewers can reduce the possibility that published studies in peer-review journals are more likely to report statistically significant effects. Systematic reviews are usually reported in the same detail as primary research studies, often including sections on background, methods, and results. In short, a science of reviewing has established that such reviews are themselves important pieces of research that need to follow the same rules of conduct and reporting as original studies. In keeping with the recommendations from the reviewing methods literature, we report below on each stage of our review. Our objective from the outset was to systematically review high-quality evidence on the effects of Scared Straight and similar programs.

Study Eligibility Criteria

There is evidence-based literature indicating that the results from randomized experiments can differ, sometimes dramatically, from findings obtained by nonrandomized methods (e.g., Boruch et al., 1999; Chalmers & Altman, 1995). Because of this evidence, we included only randomized experiments in this review. We made no exclusion on the basis of how well implemented the randomization was, but will examine the influence of breakdowns of random assignment on the results in our future analyses. We excluded all non-randomized or quasi-experimental evaluations.³

We required that the program's focus be on juvenile participants. We included studies that also exposed young adults along with juveniles to the intervention (e.g., ages 14 to 20). The program had to be delivered at a reformatory or prison. Programs involving classroom or other public visits by offenders or ex-offenders, such as Oklahoma's Speak-Outs Program, were not considered (Holley & Brewster, 1996). Programs using other methods for delivery, such as the creation of videos and their mailing to schools, were also excluded. We found no randomized experiments that tested these programs, however. The program could include either confrontational or educational

presentations by the offenders, tours of the facility (Michigan Department of Corrections, 1967), or orientation and counseling (Vreeland, 1981). We did not require confrontational activity on the part of the inmates, though this is the most visible component in the Scared Straight television documentaries. Other eligibility criteria included (a) the study report had to include a clear statement of random assignment of juveniles to experimental or control conditions, (b) the study had to include at least one measure of crime in the community, and (c) the study document had to be published or available through 1999. We imposed no English-language restriction but did not find any abstracts to potentially eligible studies in languages other than English.

Search for Eligible Studies

Randomized experiments were identified from a larger review of randomized trials in crime reduction conducted by the first author (Petrosino, 1997). Petrosino used the following methods to find more than 300 randomized experiments (and analyze 150):

1. Handsearch (i.e., visually scanning the contents) of 29 leading criminology and other journals;
2. Checking the *Registry of Randomized Experiments in Criminal Sanctions* (Weisburd, Sherman, & Petrosino, 1990);
3. Electronic searches of Criminal Justice Abstracts, Sociological Abstracts and Social Development and Planning Abstracts (Sociofile), Education Resource Information Clearinghouse (ERIC), and Psychological Abstracts (PsycInfo);
4. Electronic searches of 18 bibliographic databases, including the National Criminal Justice Reference Service (NCJRS), personally and with information specialists;
5. An extensive mail campaign with more than 200 researchers and 100 research centers;
6. Published solicitations in association newsletters;
7. Tracking of references in more than 50 relevant systematic reviews and literature syntheses;
8. Tracking of references in relevant bibliographies, books, articles, and other documents.

More details about these search methods can be found in Petrosino (1997).

The citations found in Petrosino (1997) cover literature published or available through 1993. We augmented this work with searches of recent literature made available from 1994 through 1999. These methods included the following:

1. Electronic search of the Social, Psychological, Educational & Criminological Trials Register being developed by the U.K. Cochrane Center and the Univer-

- sity of Pennsylvania (Petrosino, Boruch, Rounding, McDonald, & Chalmers, in press);
2. Check of citations from systematic or literature reviews with coverage of more recent studies (e.g., Sherman et al., 1997);
 3. Electronic searches of relevant bibliographic databases, including Criminal Justice Abstracts, NCJRS, Sociofile, PsycInfo, and ERIC.⁴

Many of these databases include unpublished literature such as dissertations and government reports. The first two authors screened relevant abstracts and agreed on 10 citations to investigate further. We rejected one, however, because the evaluation used a matched design and not randomization (Buckner & Chesney-Lind, 1983). We include, therefore, nine randomized experiments in our review.

Data Extraction and Analysis

We extracted information on variables of interest from the original study reports. We supplemented data from the original reports by contacting original investigators when critical data, such as those on outcomes, were missing. This occurred with two studies. Our initial plan was to extract data on outcome measures focusing on changes in educational performance, but only one experiment included information on educational measures (Cook & Spirrisson, 1992). Though several did report data on attitude measures, the scales and analyses reported were so diverse, both within and across studies, as to make synthesis and interpretation inappropriate if not impossible. Given the weak relationship between attitude measures and subsequent criminal activity (e.g., Morris, 1974), we decided not to focus on that information and instead to look only at crime outcomes.⁵

Descriptive Results

As described in Appendix A, the nine experiments were conducted in eight different states, with Michigan the site for two studies (Michigan Department of Corrections, 1967; Yarborough, 1979). No research team conducted more than one experiment. The studies span the years 1967 through 1992. The first five studies located were unpublished and were disseminated in government documents or dissertations; the remaining four were found in academic journals or book. Our searches, therefore, were able to identify and retrieve some documents from the fugitive literature that are generally more difficult for reviewers to take account of (Chalmers & Altman, 1995). None of the prior syntheses of crime prevention programs included all nine Scared Straight-style experiments we review here. For example, the University of

Maryland report concludes that Scared Straight does not work based on negative results in three evaluations (Buckner & Chesney-Lind, 1983; Finckenaue, 1982; Lewis, 1983) and the comparative analysis of program effects reported earlier by Lipsey (1992).

The average age of the juvenile participants in each study ranged from 15 to 17. Only the New Jersey study included girls (Finckenaue, 1982). Racial composition across the nine experiments was diverse, ranging from 36% to 84% White. Most of the studies dealt with delinquent youths already in contact with the juvenile justice system.

The interventions were also diverse. The program components used in any one of these studies did not match any other study in the review. The closest in content were the three studies that implemented single program components: Illinois's realistic rap (Greater Egypt Regional Planning & Development Commission, 1979), New Jersey's confrontational rap (Finckenaue, 1982) and Mississippi's educative rap (Cook & Spurrison, 1992). Nevertheless, these three differed in the intensity of confrontation and graphic depiction by the inmates. All of the experiments listed in Appendix A included a no-treatment control group and all but one were simple two-group experiments. Vreeland (1981) is the exception. He used a factorial design in which juveniles were randomly assigned to four conditions: (a) prison orientation and counseling, (b) prison orientation only, (c) counseling only, and (d) no-treatment control.

Substantive Findings

Programs such as Scared Straight and their derivatives not only show little deterrent effect, but very likely cause more harm than good. They are each summarized below. Appendix B provides more detail on sample sizes and crime outcomes for each of the nine experiments.

The Michigan Department of Corrections reported the first of these experiments in 1967. Unfortunately, the report is remarkably brief and provides little more than the outcome data. Juveniles who attended two tours of a state reformatory were compared with a no-treatment control group. At 6 months, 43% of the experimental group had committed a new delinquent offense, compared to only 17% of the control group. Curiously, more attention is not given to this large negative result in the original document.

The Greater Egypt Regional Planning & Development Commission (1979) examined the effects of a Scared Straight program in Illinois with a no-treatment control group. They examined the percentage of boys in each group who were subsequently contacted by the police. Again, the results are negative in direction, with 17% of the experimental participants failing in

contrast to 12% of the control participants. The authors concluded that “based on all available findings one would be ill advised to recommend continuation or expansion of the juvenile prison tours. All empirical findings indicate little positive outcome, indeed, they may actually indicate negative effects” (Greater Egypt Regional Planning & Development Commission, 1979, p. 19).

Yarborough (1979) reported the second experimental study conducted in Michigan, this time of the Juvenile Offenders Learn Truth (JOLT) program. He compared JOLT participants on a variety of crime outcomes with a control group at 3- and 6-month follow-ups. Although the differences were small and varied across these outcomes, most results were again in the direction of favoring the control group. For example, at 6 months, Yarborough reported that 31% of the experimental group had committed new criminal offenses, compared with 29% of the controls. The average offense rate for program participants was .69, compared with .47 for the control group. Yarborough concluded that “there can be little doubt that the preponderance of the evidence reported here supports the conclusion that JOLT, unfortunately, is not an effective criminal deterrent” (1979, p. 14).

Orchowsky and Taylor (1981) presented the only positive results from the experiments. They compared a group of boys who attended the confrontational Insiders program with a no-treatment control group on a variety of crime outcome measures, at intervals of 6, 9, and 12 months. The percentage of juveniles in each group who failed favored the control group at 6 months (39% of controls had new court intakes vs. 41% of experimental participants). As Appendix B indicates, however, the results favored the experimental participants at 9 and 12 months. The investigators noted, however, that the attrition rates in their experiment were dramatic at both 9 months (42% of the original sample had dropped out) and at 12 months (55% had dropped out).

Vreeland (1981) conducted a factorial experiment to determine the effects of different components of the Texas Face-to-Face juvenile aversion program. He compared boys who had gone through a prison orientation and counseling program with those who attended the orientation only, had counseling only, or were assigned to a no-treatment control group. He examined official court records and self-reported delinquency at 6 months, finding that the control participants outperformed the three treatment groups on official delinquency (28% delinquent vs. 39% for the prison orientation plus counseling, 36% for the prison only, and 39% for the counseling only). The self-report measure, however, showed a reverse pattern. All three treatment groups had similar proportions of participants who self-reported offenses (59%), whereas 69% of the control group self-reported offenses. Vreeland found that there were discrepancies between the self-report and official data; some who were officially charged did not self-report the offense and vice-versa. He

seems to have more confidence that the official data captures more harmful offenses by participants in the study, stating that "official records have been shown to be reasonably accurate with respect to the more serious crimes of persistent delinquents" (Vreeland, 1981, p. 24). Viewing all the data, Vreeland concluded that there was no evidence that Face-to-Face was an effective delinquency prevention program.

Finckenauer (1982) conducted the most visible experiment on the Scared Straight program, comparing the performance of participants with that of a no-treatment control group for 6 months in the community. He reported that 41% of the children who attended Scared Straight committed new offenses, whereas only 11% of controls did. He also found that the program participants committed more serious offenses. Finckenauer (1982) noted that random assignment procedures were violated during the study; only 8 of the 11 participating agencies that referred troubled or delinquent boys to the program correctly assigned their cases. He conducted several additional analyses in an attempt to compensate for violation of randomization. Even when cases that were incorrectly assigned were removed, however, the failure rate for the Scared Straight attendees was 31%, compared with 17% for controls.

Lewis (1983) provided some more evidence of a possible harmful effect in his evaluation of the San Quentin Utilization of Inmate Resources, Experience and Studies (SQUIRES) program. He compared juveniles attending SQUIRES with a no-treatment control group on a variety of crime outcomes at 12 months. Though a number of different measures were used, Lewis reported that 81% of the program participants were arrested, compared with 67% of the controls. He also found that the program did worse with seriously delinquent youths, leading him to conclude that such children could not be "turned around by short-term programs such as SQUIRES . . . a pattern for higher risk youth suggested that the SQUIRES program may have been detrimental" (Lewis, 1983, p. 222).

Locke, Johnson, Kirigin-Ramp, Atwater, and Gerrard (1986) reported little effect of the Juvenile Education Program in Kansas, an intervention designed to be less confrontational and offensive than the New Jersey program. The investigators examined crime outcomes at 6 months for program attendees and a no-treatment control group. Group failure rates were not available, but the investigators concluded that there were no differences between experimental and control groups on any of the crime outcomes measured. Though direction of effect was not provided, the test statistic for the analysis of variance used ($F = .75$) not only indicates that it was not significant but would be very small regardless of direction.

Finally, Cook and Spurrison (1992) compared juveniles who attended Mississippi's Project Aware with a no-treatment control group on a variety of crime outcomes at 12 and 24 months. Most of the findings favored the control participants, but again the differences between the groups were small. For example, the mean offending rate for controls at 12 months was 1.25 versus 1.32 for Project Aware participants. The investigators concluded that "attending the treatment program had no significant effect on the frequency or severity of subsequent offenses" (Cook & Spurrison, 1992, p. 97).

Table 1 provides a summary of results based on the criterion of whether the program increased or decreased officially recorded offenses at first follow-up. Given that most studies report only one follow-up period, reviewers have used a "first effects" approach in summarizing crime and delinquency treatment studies (Lipsey, 1992; Petrosino, 1997). Important information reported in the studies, however, is ignored by this approach, such as if the program reduced the average number of offenses committed by the juveniles or reduced their severity (Orchowsky & Taylor, 1981). Self-report data are not presented in Table 1.

These results, though preliminary, should lead to sobering caution on the part of persons who wish to revive programs such as Scared Straight. Only seven studies reported group failure rates. Examining those data, we find that the program increases the percentage of the treatment group committing new offenses anywhere from 1% to 30%. This is in comparison with a randomly assigned no-treatment control group. If we assume the randomization breakdown in Finckenaue's (1982) experiment rendered that study invalid and exclude it, the remaining six studies increase new offenses in the treatment group anywhere from 1% to 26%. The experiments that did not provide such percentages provide no contradictory evidence of a positive effect for programs such as Scared Straight (Cook & Spurrison, 1992; Locke et al., 1986), and one indeed suggests a slight negative impact (Cook & Spurrison, 1992).

These findings are remarkable in the context of other systematic reviews. Lipsey (1992) reviewed nearly 400 evaluations of juvenile delinquency programs. When looking only at the direction of the first effect reported (the difference between the experimental and the control group), 64% reported a difference in favor of treatment. Thirty percent were negative in direction; that is, they favored the control group. Petrosino (1997) reported that 63% of the first effects in the 150 experiments in his meta-analysis differed between experimental and control groups in favor of treatment. Only 14% of his sample reported effects in a negative direction, favoring the control group (surprisingly, the remaining 23% showed an absolute zero difference). In con-

TABLE 1: Effects of Scared Straight Programs on Participants (official data only, direction of first effect reported, $N = 7$)

<i>Year, Author</i>	<i>Type of Data</i>	<i>Time Interval</i>	<i>Percentage Change</i>
1967, Michigan Department of Corrections	Percent delinquent	6 months	+ 26% increase in failure
1979, Greater Egypt Regional Planning & Development Commission	Percent contacted by police	5 to 15 months	+ 5% increase in failure
1979, Yarborough	Percent committing new offenses	3 months	+ 1% increase in failure
1981, Orchowky and Taylor	Percent with new juvenile court intakes	6 months	+ 2% increase in failure
1981, Vreeland	Percent with officially recorded delinquency	6 months	+ 11% increase in failure
1982, Finckenauer	Percent with new offenses	6 months	+ 30% increase in failure
1983, Lewis	Percent with new arrests	12 months	+ 14% increase in failure

trast, all seven of the experiments shown in Table 1 reported first effects in a negative direction.

DISCUSSION

Galvin (1979) noted that one of the negative consequences of Scared Straight is that it would divert attention and resources from good projects. Our preliminary data show that the consequences are possibly worse. The program likely had harmful effects, leading to increased crime and delinquency in our communities (see Table 1). Why would the program have harmful effects? The reasons have not been explicitly tested, but some rationale is provided by some of the original investigators. For example, one investigation team suggested that some youngsters might find prison attractive, stating, "Many delinquent youths feel alienated . . . delinquents view prison as a place where they can have friends and a community now lacking in their lives. Four walls and bars may, in some way, offer security and a sense of belonging" (Greater Egypt Regional Planning & Development Commission, 1979, p. 19).

Finckenauer also provides some material for why the program had negative results. In the New Jersey study, the program seemed to do worse with those youths not yet officially in contact with the juvenile justice system. Finckenauer suggests that

The controversial possibility also exists that the project actually sets in motion a "delinquency fulfilling prophecy" in which it increases rather than decreases the chances of juvenile delinquency The project may romanticize the Lifers—and by extension other prison inmates—in young, impressionable minds. Or, the belittling, demeaning, intimidating, and scaring of particular youth may be seen as a challenge; a challenge to go out and prove to themselves, their peers and others that they were not scared. (1982, p. 169)

Still, Old Programs Never Seem to Die

Despite negative or harmful effects, the Scared Straight program continued to be run in a number of jurisdictions (Finckenauer & Gavin, 1999), and many similar programs are in operation today (Hall, 1999). Attempts to dismantle programs such as Scared Straight have met resistance. In Michigan, the JOLT program was terminated following the results of the randomized experiment conducted by Yarborough (1979). Yet, despite the results of the experiment, proponents of JOLT argued against termination. They relied on the following themes: (a) The evaluation was flawed, (b) people love the program, (c) it helps the inmates, and (d) it is cost free for the state (Homant, 1981; Homant & Osowsky, 1981). Even Homant (1981) concluded that the program might better have been retooled and modified rather than terminated. Advocates for JOLT also argued that the program had no "statistically significant" harmful effect on juveniles. Finckenauer (1982) noted that after he reported the results of his experiment in New Jersey, the criteria for success changed among some from reducing recidivism to "it's worth it if it only helps one child."

Another reaction was for program supporters to argue that programs such as Scared Straight provided other benefits that were not the target of the experiments. For example, Wormser (1991) talks about its positive impact on the prisoners at East Jersey State Prison (formerly known as Rahway State Prison), who had spoken to more than 35,000 juveniles in an attempt to keep them out of jail. Israel (1980) more vehemently argued his support for the Scared Straight program despite the early results from the Finckenauer (1982) experiment:

The relevant policy question is whether this is an intrinsically valuable experience. There are times when the academic community must take some leader-

ship to encourage a climate of opinion that is willing to take some risks. To see it [the program] ruined by a control group of 35 juveniles . . . is a violation of the sacred values of our discipline, and the social responsibility that should accompany our influence. (Israel, 1980, pp. 16-18)

Cook (1990) speculated that the program could have improved the image of the state's department of corrections. Even the Michigan Department of Corrections report, issued more than 30 years ago, speculated that visits to a reformatory might have inspired more juveniles to formally seek counseling (Michigan Department of Corrections, 1967). Whether these benefits outweigh the apparent harmful effects of programs such as Scared Straight is debatable. Programs such as Scared Straight, as other social interventions, likely have a number of latent goals (Finckenauer & Gavin, 1999). These must be weighed against the manifest aim of the program—to reduce crime and delinquency.

Why the Paradox?

Interestingly, the dubious attitude toward evaluation that is held by some policy makers extends beyond any up-front belief that rigorous evaluation is unnecessary. In those instances when evaluations are carried out, findings are often ignored or rejected by those same policy makers (Finckenauer & Gavin, 1999). Finckenauer and Gavin (1999, pp. 216-217) describe this as a paradox in which programs that have been evaluated and deemed to be ineffective nevertheless continue. Their endurance is seemingly untouched by any credible, empirical evidence of their success or failure.

For example, despite negative findings from the SQUIRES experiment (Lewis, 1983), the program continued. Today, its effectiveness is judged by letters from participating youths (and others), who describe how the program influenced them (Finckenauer & Gavin, 1999). This was the same method that was used to demonstrate the effectiveness of the New Jersey Lifers Program before the randomized experiment was conducted (Finckenauer, 1982). The SQUIRES program has not undergone another rigorous evaluation since the Lewis study (Finckenauer & Gavin, 1999).

These authors point to a number of factors that seem to account for this paradox. The first is a political climate that demands action; in the case of crime and delinquency, often "get tough" action (e.g., Zimring, 1999). Also, in the case of crime control policy, there is a perception that any alternatives to getting tough, such as treating offenders, do not work. With respect to programs more generally, there is an inertia factor among policy makers to account for why programs or policies, once created, take on lives of their own. It is

easier to continue such programs and avoid angering constituents than it is to stop them. There may also be a media factor with visual appeal, compelling stories, and sound bites that help perpetuate certain programs.⁶

Another factor, according to Finckenauer and Gavin (1999), that may account for the lack of impact of evaluations is the information gap that often exists between researchers and policy makers. Practitioners may often be ignorant of research findings because the evaluators have been mostly interested in communication with their peers in the research community. Policy makers may also reject research results because of their suspicion of social science, with its complicated analyses, hedged conclusions, and conflicting findings. Finally, there are administrators and officials who do not try nor care to find information that may be available to them. They know what they want to do and do not wish to be dissuaded. A long history of research on how findings are used by policy makers underscores these and other barriers to the use of knowledge in decision making (Weiss, 1998).

We Need Randomized Experiments and Better Outcome Studies

Some policy makers, practitioners, and researchers, as well as many in the general public, believe that programs are good things that can do no harm. When surveys are undertaken to determine the satisfaction of groups with particular programs, the results are almost always positive, persuading even more that the intervention is a good idea. Even with Scared Straight, whether the original investigators talked with inmates, juvenile participants, parents, corrections personnel, teachers, or the general public, everyone was positive about it (e.g., Greater Egypt Regional Planning & Development Commission, 1979). Almost everyone believed the program was doing good (Finckenauer, 1982). Compounding this was a number of single group before-and-after designs that seemed to indicate the program had dramatic crime reduction effects.

Carefully done evaluation is needed to rule out alternative explanations for changes in outcome measures before we can make causal inferences about a program's impact on crime with much confidence. The literature on the Scared Straight program contains some examples that underscore the need for such careful evaluation. For example, Serpas, Littleton, and Ashcroft (1979) conducted a study of a program similar to Scared Straight in New Orleans. They found a 52% decrease in the absolute number of arrests from pretest to the 1-year follow-up period. How could such a dramatic effect be the result of anything other than the program? There are many who would claim that randomized experiments or quasi-experiments (i.e., comparison

group designs) are not needed with such dramatic effects. Unfortunately, we have no other evaluation data from the Orleans Parish study to understand if the program was responsible for the observed decrease in crime.

Fortunately, two of the Scared Straight experiments suggest that the data from before-and-after studies with a single group must be viewed with extreme caution. In the first experiment, Cook and Spirrison (1992) report substantial decreases for program participants in mean offense rates from the baseline measure at the beginning of the program to the posttest measure at 12 and then 24 months. In the second experiment, Locke and his colleagues (1986) report a comparable finding in their evaluation of the Kansas Juvenile Education Program. Without a control group, the only conclusion, given such large and positive results, would be that the program was successful.

Both randomized experiments, however, underscore the importance of ruling out other threats to internal validity; that is, rival explanations for the observed impact. In both cases, the randomly assigned control group also experiences a sizable and statistically significant decrease in criminality from pretest to posttest! In fact, the postprogram performance of the control group is similar (and in one study, slightly better) to that of the experimental participants. Because of random assignment, we are confident that the groups were comparable and differ only in regard to their participation in the program. The reason for the improvement of both treatment and control groups is speculative at best because they were not implicitly tested in the studies. The authors indicate, across the literature, that the maturation process for juveniles is dramatic during the teen years (when Scared Straight normally selects eligible youths) and naturally leads to a reduction in delinquent activities. The reduction is sometimes mistakenly interpreted as a positive impact for juvenile programs (Langer, 1980). Other researchers have pointed out that juveniles are selected for such programs because they commit offenses at a high rate, but the natural statistical regression back to the mean (i.e., their average offending rate) is wrongly interpreted as a program effect (Finckenauer, 1982).

By including a randomized control group, positive changes in the treatment group's performance were not incorrectly attributed to Scared Straight. We have to ask ourselves whether alternatives to randomization could compensate for the problems of internal validity that particularly hamper before-and-after evaluation designs. There is a long history in evaluation of developing and implementing methods to rule out threats to internal validity when randomization is impossible (Weiss, 1998). Many are underutilized in actual practice. Such alternatives, however, often result in equivocal findings, and leave us wondering whether uncontrolled variables or selection biases were responsible for the observed outcome (Boruch et al., 1999).

CONCLUSION

Although rigorous evaluation is often resisted, the agencies and institutions that facilitated the Scared Straight experiments described here should be credited. It would be difficult to find another justice-related program that has been subjected to nine randomized experiments. On the other hand, only nine experiments were conducted over the 33-year history of a widely disseminated and internationally implemented program. Some may interpret this as even more discouraging evidence that rigorous evaluations are rare and the use of results from sound research rarer still (Finckenauer & Gavin 1999).

The findings reported here are sobering. They do indicate that despite our best intentions, programs can not only fail to reach objectives but can backfire, leading to more harm than good. Few programs were as popular or well intentioned as Scared Straight. Yet, despite such popularity and benevolence, there is little evidence to suggest that the program is a deterrent to subsequent juvenile crime and delinquency. In contrast, the evidence strongly suggests that it leads to more crime by program participants. Given the possibility of harmful effects of interventions, government has an ethical responsibility to rigorously evaluate, on a continual basis, the policies, practices, and programs it implements (Sherman, 1984).

APPENDIX A: Scared Straight Randomized Experiments (control in each study was no-treatment group)

Year, Author	Setting	Document	Program Components	Eligibility Criteria	M Age	% Male	% White	Prior History
1967, Michigan Department of Corrections	Michigan	Unpublished	Two tours of reformatory	Unk.	Unk.	Unk.	Unk.	Unk.
1979, Greater Egypt Regional Planning & Development Commission	Illinois	Unpublished	Realistic rap sessions	Unk.	15	100	84	67% prior court/police contact
1979, Yarborough	Michigan	Unpublished	Prison tour Taken to cell Confrontational rap sessions	Male At least one offense	16	100	43	$\bar{X} = 3$ offenses
1981, Orchowsky and Taylor	Virginia	Unpublished	Locked in cell Confrontational rap sessions Processed out of prison	Male 13 to 20 years old At least two offenses No apparent mental/emotional illness		100		
1981, Vreeland	Texas	Unpublished	1-day orientation with haircut and work detail Group counseling (nine sessions with specially trained probation officers)	On probation	16	100	40	$\bar{X} = 2.4$ offenses

1982, Finckenaue	New Jersey	Published	Confrontational rap	Delinquent or at-risk children	15	80	40	50% had prior offenses
1983, Lewis	California	Published	Guided tour Confrontational rap	16 to 17 years old Early phase of camp or treatment	16	100	Unk.	$\bar{X} = 7$ arrests
1986, Locke, Johnson, Kirigin-Ramp, Atwater, and Gerrard	Kansas	Published	Educational rap Tried to match juvenile with inmate	Record of delinquency On probation 14 to 19 years old	17	100	65	$\bar{X} = 1.5$
1992, Cook and Spurrison	Mississippi	Published	Educational rap	12 to 16 years old Under youth court jurisdiction	15	100	36	All had at least one offense

NOTE: Unk. = unknown.

APPENDIX B: Crime Outcome Data for the Scared Straight Experiments (N = 9)

<i>Year, Author</i>	<i>Number in Experimental Group</i>	<i>Number in Control Group</i>	<i>Outcome Measures</i>	<i>Result</i>
1967, Michigan Department of Corrections	30	28	Percent delinquent at 6 months	43% experimental, 17% control
1979, Greater Egypt Regional Planning & Development Commission	94	67	Percent contacted by police during 5- to 15-month follow-up	17% experimental, 12% control
1979, Yarborough	137	90	Percent with new petitions, percent committing new offenses, mean offense rate, mean weeks to new offense, mean days in detention, type of offense committed at 3 months and at 6 months	Diverse findings, none are statistically significant but most slightly in favor of control group
1981, Orchowsky and Taylor	39	41	Percent with new court intakes, mean intake rate, delinquency involvement score at 6, 9, and 12 months	Diverse findings, most favor experimental group at 9 and 12 months
1981, Vreeland	36 (PC) 39 (P) 36 (C)	40	Official court recorded delinquency and self-report delinquency at 6 months	Official court records in favor of controls (e.g., 28% failure vs. 39% for prison orientation and counseling group), but self-report indicates controls did worse (69% self-reporting offenses vs. 59% for three treatment groups)

1982, Finckenaer	46	35	Percent failure, offense severity at 6 months	41% of experimental group failed vs. 17% of controls; controls also committed less serious offenses ($t = 2.67$, $p = .01$)
1983, Lewis	53	55	Percent arrested, mean arrest rate, percent charged, mean charge rate, length to first arrest, seriousness of offense at 12 months	Diverse findings, most are similar, except percentage arrested favors controls (67% vs. 81% of experimentals). Time to first arrest significantly favors experimentalists (4 months to 3.3 months)
1986, Locke, Johnson, Kirigin-Ramp, Atwater, and Gerrard	18	18	Self-report delinquency, official delinquency at 6 months	No effect, but no indication provided of direction of effect
1992, Cook and Spirrison	97	79	Frequency of offenses, most severe offense, composite severity, mean severity at 12 and 24 months	Diverse findings, most favor control group but actual differences very small

NOTE: PC = prison orientation and counseling. P = prison orientation only. C = counseling only.

NOTES

1. Carol Fitz-Gibbon (1999) noted that teaching represents about 15,000 hours of uncontrolled experimentation into the lives of schoolchildren.

2. At the time of press, this e-mail was received from Correx, the listserv of the National Institute of Corrections (February 7, 2000):

I am an correctional officer for a small detention center. I would like to present a program to my Captain about a program called Scared Straight. I remember it when I was growing up in N.J. I would like to try to start one like it in my detention center. We house state, county, and pre-trial inmates. I would like to use our state inmates in this program to talk to our pre-trials and also to schools. Any info would be much appreciated.

3. For example, we excluded the following studies: Brodsky (1970); Buckner and Chesney-Lind (1983); Chesney-Lind (1981); Langer (1980); Nygard (1980); O'Malley, Coventry, and Walters (1993); Serpas, Littleton, and Ashcroft (1979); Syzanski and Fleming (1971); and Trotti (1980).

4. The exact search terms used can be obtained from the first author.

5. Our future plans include a check of interrater reliability to insure that data extraction was uniform between us. One of us will also enter the data into Review Manager, a software program designed specifically for the production of systematic reviews (Review Manager 4.0, 1999). Though we have yet to conduct more sophisticated meta-analytic procedures on these data, the findings from this preliminary analysis should be sobering to those seeking to revive programs such Scared Straight.

6. This last reason may partially account for why legislation is propelled so quickly by high-profile murders (Petrosino, Hacsí, & Turpin-Petrosino, 2000).

REFERENCES

- Berkman, D. J., & Pearson, S. R. (1980). *Juvenile aversion programs—A status report*. Washington, DC: American Justice Institute, National Juvenile Justice System Assessment Center.
- Boruch, R. F., Snyder, B., & DeMoya, D. (1999, May). *The importance of randomized field trials in delinquency research and other areas*. Paper presented to the American Academy of Arts & Sciences, Cambridge, MA.
- Brodsky, S. L. (1970). The prisoner as agent of attitude change: A study of prison profiles' effects. *The British Journal of Criminology*, 10, 280-285.
- Buckner, J. C., & Chesney-Lind, M. (1983). Dramatic cures for juvenile crime: An evaluation of a prisoner-run delinquency program. *Criminal Justice & Behavior*, 10, 227-247.
- Cavender, G. (1984). Scared Straight—Ideology and the media. In R. Surette (Ed.), *Justice and the media* (pp. 246-259). Springfield, IL: Thomas.
- Chalmers, I. (1999, July). *How can researchers help practitioners and policy makers to maximize the good and minimize the harm they do to others?* Paper presented at the Evidence Based Policies and Indicators Systems Conference, University of Durham, United Kingdom.
- Chalmers, I., & Altman, D. (Eds.). (1995). *Systematic reviews*. London: BMI.

- Chesney-Lind, M. (1981). *Ike Na Pa'ahao—the experience of the prisoners (formerly Stay Straight)—evaluation*. Manoa: University of Hawaii, Center for Youth Research.
- Cook, D. D. (1990). *Effects of a non-confrontational prisoner-run juvenile delinquency deterrence program*. Unpublished master's thesis, Department of Psychology, Mississippi State University.
- Cook, D. D., & Spurrison, C. L. (1992). Effects of a prisoner-operated delinquency deterrence program: Mississippi's Project Aware. *Journal of Offender Rehabilitation, 17*, 89-99.
- Corrigan, J. W. (1979). Scared Straight!—Controversy surrounds inmate program. *Youth Forum, 3*(2), 1-6, 8.
- Eicher, D. (1999, July 31). Stark prison documentary redone for MTV special. *The Denver Post*, p. F-05.
- Finckenauer, J. O. (1980). Scared Straight and the panacea phenomenon. *Annals of the New York Academy of Sciences, 347*, 213-217.
- Finckenauer, J. O. (1982). *Scared Straight and the panacea phenomenon*. Englewood Cliffs, NJ: Prentice Hall.
- Finckenauer, J. O., & Gavin, P. W. (1999). *Scared Straight: The panacea phenomenon revisited*. Prospect Heights, IL: Waveland Press.
- Fitz-Gibbon, C. (1999, January-March). Education: The high potential not yet realized. *Public Money & Management, 33*-40.
- Galvin, J. L. (1979, November). *Scared Straight—The losers lose again*. Paper presented at the national Legal Aid and Defender Association Annual Convention, Albuquerque, NM.
- Greater Egypt Regional Planning & Development Commission. (1979). *Menard Correctional Center: Juvenile tours impact study*. Carbondale, IL: Author.
- Hall, A. (1999, October 26). Jailhouse shock aims to scare youths straight. *The Scotsman*, p. 12.
- Heeren, J., & Shicor, D. (1984). Mass media and delinquency prevention: The case of "Scared Straight." *Deviant Behavior, 5*, 375-386.
- Holley, P. D., & Brewster, D. (1996). Description of speak-outs in Oklahoma prisons. *Journal of the Oklahoma Criminal Justice Research Consortium, 3*, 123-133.
- Homant, R. J. (1981). Demise of JOLT—The politics of being "scared straight" in Michigan. *Criminal Justice Review, 6*, 14-18.
- Homant, R. J., & Osowski, G. (1981). Politics of juvenile awareness programs—A case study of JOLT (Juvenile Offenders Learn Truth). *Criminal Justice and Behavior, 9*, 55-68.
- Israel, M. (1980). *The Rahway Lifer's Group: The politics of being my brother's keeper*. Union, NJ: Kean College.
- Kids and crooks revisited: Some were "Scared Straight!" (1999, April 12). *USA Today*, p. 4D.
- Langer, S. (1980). *Fear in the deterrence of delinquency: A critical analysis of the Rahway State Prison Lifer's Program*. Union, NJ: Kean College, Department of Sociology and Social Work.
- Lewis, R. V. (1983). Scared straight—California style: Evaluation of the San Quentin Squires program. *Criminal Justice and Behavior 10*(2), 209-226.
- Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation* (pp. 83-127). New York: Russell Sage Foundation.
- Locke, T. P., Johnson, G. M., Kirigin-Ramp, K., Atwater, J. D., & Gerrard, M. (1986). An evaluation of a juvenile education program in a state penitentiary. *Evaluation Review, 10*, 281-298.
- Lundman, R. J. (1993). *Prevention and control of juvenile delinquency* (2nd ed.). New York: Oxford University Press.

- Michigan Department of Corrections. (1967, May). *A six month follow-up of juvenile delinquents visiting the Ionia Reformatory* (Research Report No. 4). Lansing: Michigan Department of Corrections.
- Morris, N. (1974). *The future of imprisonment*. Chicago: University of Chicago Press.
- National Advisory Committee for Juvenile Justice and Delinquency Prevention. (1980). *Annual Report, 4th, October 1, 1978-September 30, 1979*. Washington, DC: Government Printing Office.
- National Center on Institutions and Alternatives. (1979). *Scared Straight—A second look*. Alexandria, VA: National Center on Institutions and Alternatives.
- Nygaard, C. H. (1980). *Synopsis of the findings, conclusions and recommendations from the evaluational study of the Department's Youth Assistance Programs*. Albany: New York State Department of Correctional Services, Division of Program Planning, Research and Evaluation.
- O'Malley, P., Coventry, G., & Walters, R. (1993). Victoria's Day in Prison Program: An evaluation and critique. *Australian and New Zealand Journal of Criminology*, 26(2), 171-183.
- Orchowsky, S., & Taylor, K. (1981). *The Insiders juvenile crime prevention program: An assessment of a juvenile awareness program*. Richmond: Virginia Department of Corrections.
- Oversight on Scared Straight—Hearings before the Subcommittee on Human Resources, House of Representatives*, 96th Cong., 1st Sess. (1979).
- Pallone, N. J. (1986). *On the social utility of psychopathology: A deviant majority and its keepers?* New Brunswick, NJ: Transaction.
- Petrosino, A. (1997). "What works?" revisited again: A meta-analysis of randomized experiments in rehabilitation, deterrence and prevention. Unpublished doctoral dissertation, Rutgers University, School of Criminal Justice, Newark, NJ.
- Petrosino, A. (1998). *Improving the evaluation of state-administered programs*. Unpublished manuscript.
- Petrosino, A., Boruch, R. F., Rounding, C., McDonald, S., & Chalmers, I. (in press). Assembling a social, psychological, educational & criminological trials register (SPECTR). *Evaluation Research in Education*.
- Petrosino, A., Hacsí, T., & Turpin-Petrosino, C. (2000). "Policy by murder": Case studies of legislation propelled by high-profile killings involving children (Proposal to the National Institute of Justice). Cambridge, MA: American Academy of Arts & Sciences.
- Review Manager 4.0 [Computer software]. (1999). Oxford, UK: Update Software.
- Serpas, F. R., Littleton, D., & Ashcroft, J. L. (1979). *A closeout report of the New Orleans Parish Juvenile Awareness Program*. Unpublished manuscript, Mayor's Criminal Justice Coordinating Council, Orleans Parish, LA.
- Shapiro, A. (Producer). (1978). *Scared Straight!* Santa Monica, CA: Pyramid Films.
- Sherman, L. W. (1984). Experiments in police discretion: Scientific boon or dangerous knowledge? *Law & Contemporary Problems*, 47(4), 61-81.
- Sherman, L. W., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising* (Report to the United States Congress). College Park: University of Maryland, Department of Criminology and Criminal Justice.
- Syzmanski, L., & Fleming, A. (1971). Juvenile delinquent and an adult prisoner—A therapeutic encounter? *Journal of the American Academy of Child Psychiatry*, 10, 308-320.
- Trotti, T. C. (1980). *Save the Children Program—An analysis of client reaction*. Columbia: South Carolina Department of Youth Services, Research and Evaluation Unit.
- Vreeland, A. D. (1981). *Evaluation of Face-to-Face: A juvenile aversion program*. Unpublished doctoral dissertation, University of Texas, Dallas.

- Weisburd, D. L., Sherman, L., & Petrosino, A. (1990). *Registry of randomized experiments in criminal sanctions, 1950-1983*. Los Altos, CA: Sociometrics Corporation, Data Holdings of the National Institute of Justice.
- Weiss, C. H. (1998). *Evaluation*. Englewood Cliffs, NJ: Prentice Hall.
- West, P. (1981). New directions in juvenile corrections. In B. Olsson & A. Dargis (Eds.), *American Correctional Association—Proceedings* (pp. 103-111). Lanham, MD: American Correctional Association.
- Wormser, R. (1991). *Lifers: Learn the truth at the expense of our sorrow*. New York: Messner.
- Yarborough, J. C. (1979). *Evaluation of JOLT as a deterrence program*. Lansing: Michigan Department of Corrections, Program Bureau.
- Zimring, F. (1999, December). *The new politics of punishment*. Paper presented to the National Institute of Justice, Washington, DC.

UNDER SURVEILLANCE: AN EMPIRICAL TEST OF THE EFFECTIVENESS AND CONSEQUENCES OF ELECTRONIC MONITORING

KATHY G. PADGETT
WILLIAM D. BALES
THOMAS G. BLOMBERG
Florida State University

Research Summary:

This study addresses the effectiveness of electronic monitoring (EM) for serious offenders supervised in the community. Using data on 75,661 offenders placed on home confinement in Florida from 1998 to 2002, we find that both radio-frequency and global positioning system monitoring significantly reduce the likelihood of technical violations, reoffending, and absconding for this population of offenders. Additionally, we find that offenders placed on home confinement with EM are significantly more serious than those placed on home confinement without EM, which casts doubt on the anticipated net-widening effect of this particular intermediate sanction.

Policy Implications:

Given the anticipated increase in the use of EM in the immediate future, policy makers will surely be faced with questions about its effectiveness in preventing or deterring further criminal activity among offenders in the community, as well as concerns about the intensity of surveillance it affords and a subsequent increase in the likelihood of a prison sentence or return to prison for technical violations. The results presented here suggest that such concerns may be overblown and that EM of offenders in the community may prove an effective public safety alternative to prison. Additional implications of this research include decision making regarding which offenders should be placed on EM, which type of monitoring device will be the most cost-effective and efficient, and the potential for front-end net-widening if states adopt the practice of "Got 'em?" Use 'em.'"

KEYWORDS: Intermediate Sanctions, Home Confinement, Electronic Monitoring, Radio-Frequency Monitoring, Global Positioning System Monitoring

During the 1980s, the United States began to experience "exponential growth in incarceration" (Blumstein, 1998). Austin et al. (2003) report that between 1980 and 2000, the prison population more than tripled. In response to this unprecedented growth in imprisonment and its associated costs, various intermediate sanctions were promoted as less costly but still "tough" and effective alternatives to imprisonment (Clear and Braga, 1995; Morris and Tonry, 1990). Despite reservations by penal reform scholars and researchers, intermediate sanctions, including intensive supervision, home confinement with and without electronic monitoring (EM), day reporting centers, and boot camps, proved appealing to both liberal and conservative policy makers and quickly spread across the country.

In theory, intermediate sanction programs were to divert offenders from prison, while providing a greater level of offender accountability and surveillance than would be provided by traditional probation supervision. The end result, therefore, would be less penal control imposed on individual offenders and less expense to the taxpayer, without any compromise to public safety (Baumer et al., 1993; Clear et al., 1998). To date, however, the extent to which intermediate sanctions have fulfilled their formal goals of reducing prison populations and protecting public safety has yet to be established. Despite the absence of empirical proof regarding the effectiveness of electronic surveillance, this strategy is likely to become a national approach for managing high-risk offenders in the community.

The widely publicized sexual battery and murder of nine-year-old Jessica Lunsford in early 2005, *allegedly* by a registered sex offender, prompted legislation in Florida that requires sex offenders who molest children to wear satellite tracking devices (global positioning system, or GPS, monitoring) for the rest of their lives once they leave prison. Despite an estimated fiscal impact on the state of \$3.9 million, the legislation passed by unanimous vote, effectively ensuring that the number of offenders in Florida under electronic surveillance will more than double (from 720 to 1,920) within fiscal year 2005–2006. Whether prompted by the case of Jessica Lunsford and Florida's Jessica Lunsford Act or by the growing awareness of the capabilities and availability of GPS monitoring devices, legislation related to the EM of offenders in the community was proposed in at least 11 additional states and at the Federal level in the Spring 2005 legislative season, all of which provide for its increased use. Even if media attention to sex offenders in the community and the presumed public outcry for closer surveillance of these offenders subside, it seems likely that the use of EM devices will increase dramatically in the very near future. Alongside this anticipated increase, policy makers will surely face questions about their effectiveness in preventing or deterring further criminal activity, as well as concerns about the intensity of surveillance they afford

and a subsequent increase in prison sentences or returns to prison for technical violations.

This study addresses the effectiveness of EM for serious offenders supervised in the community. Using data on 75,661 offenders placed on home confinement in Florida from 1998 to 2002, we estimate the effect of radio-frequency (RF) and GPS monitoring on the likelihood of revocation and absconding from supervision. During this time period in Florida, only a small percentage of offenders placed on home confinement was ordered to wear an EM device as a condition of the home confinement sentence, which allows for a comparison between those and other, like offenders who were not electronically monitored in terms of their likelihood of technically violating, reoffending, or absconding while on home confinement. In doing so, we can test for a potential net-widening effect of EM as well as its potential for protecting public safety. We also test for potential net-widening at the "front end," or the point at which the decision is made to place an offender on EM as a condition of his or her home confinement sentence, with additional data on offender seriousness levels.

PRIOR RESEARCH

Scholarly interest in the intentions and consequences of penal reform emerged in the 1970s with studies focused on the penal control consequences of the juvenile diversion programs that proliferated throughout the United States during the late 1960s and 1970s. Evidence of a net-widening effect of these diversion programs was documented in the late 1970s and early 1980s (Austin and Krisberg, 1981; Blomberg, 1977; Klein, 1979; Lemert, 1981), followed by further evidence of net-widening associated with the "get tough" crime control strategies of the 1980s (Blomberg and Waldo, 1987; Hylton, 1982). By the late 1980s, concern was being expressed for the increasing use of intermediate punishment programs, especially as they were exacerbated by violations of strict sentence conditions that resulted in an eventual prison term anyway.

By the 1990s, evidence of disproportionate increases in prison populations, despite the use of intermediate sanctions (Blomberg et al., 1993; Frazier and Lee, 1992; Mainprize, 1992; Petersilia and Turner, 1990), and of a "piling up of sanctions" related to intermediate punishment programs (Blomberg and Lucken, 1994; Bonta et al., 2000; Clear et al., 1998; Lucken, 1997; Ulmer, 2001), had been well-documented. Even more recently, net-widening has been associated with "new penology" kinds of techniques, such as EM (Bonta et al., 2000; Mainprize, 1992), differentiated case management (Taxman and Elis, 1999), and detention alternatives for juvenile offenders (Frazier and Lee, 1992; Walters, 1996).

This evolution of the "net-widening" concept has resulted in an extension of its scope of reference as well. Although its original reference was specific to the consequence of "increasing the scope of corrections" (Clear and Cole, 2003), Renzema (2003:4) notes that "In practice, 'net-widening' refers to both more harsh dispositions . . . and to the expansion of the total offender processing capacity of a jurisdiction." This recognition of the dual nature of the net-widening phenomenon resonates with Tonry and Lynch's (1996:106) distinction between the "front-end" and "back-end" net-widening consequences of intermediate sanctions. They define "front-end" net-widening as the use of enhanced penalties for offenders who would not otherwise have received a prison sentence and "back-end" net-widening as the increased likelihood of an eventual prison sentence for atechanical violations among offenders subject to more intense surveillance.

With only a handful of empirical studies published, to date, and only two that specifically test for an effect of EM on technical violations, no firm conclusions can be drawn about its potential for "back-end" net-widening. Moreover, in the two studies that *do* examine EM and technical violations, the findings are contradictory. For example, Coopridger and Kerby (1990) find significantly higher rates of technical violation for pre-trial release offenders on EM than for those released into the community with no monitoring, whereas SPEC Associates (2002) find a significant *negative* effect of EM on the likelihood of a parole violation. Two additional studies (Bonta et al., 2000; Finn and Muirhead-Steves, 2002) address technical violations in their examinations of EM and "program completion." Bonta et al. (2000) found no effect of EM on technical violations when the offender's risk score is controlled for. Finn and Muirhead-Steves (2002) report that 76% of their sample of parolees placed on EM completed the EM program with no violations, but no comparable figures for parolees not on EM are provided for comparison.

With regard to "front-end" net-widening as a consequence of EM, some findings have been reported by Berry (1985), Schmidt (1991), and Vaughn (1987). However, the evidence reported to date has been limited to demonstrations of the relative "low risk" of the offenders most often sentenced to EM (Baumer and Mendelsohn, 1992; Bonta et al., 2000; Stanz and Tewksbury, 2000). No studies that we know of have systematically examined the "mechanism that operates to widen the net," which Morris and Tonry (1990:225) contend lies in judicial decision making in the context of newly available intermediate punishments. However, a concern for potential front-end net-widening associated with EM and other intermediate sanctions, especially in terms of their application to "low-risk" offenders, is expressed repeatedly (Baumer and Mendelsohn, 1992; Clear et al., 1998; Gendreau et al., 2000; Morris and Tonry, 1990). These and other issues related to the use of EM have led Renzema (2003:5) to the same

question we attempt to answer here: "Is there at least evidence of reduction of reoffending to counterbalance these concerns [punitiveness, intrusiveness, and system expansion due to program failures]?"

Although some form of home confinement with EM had been implemented in all 50 states by 1990 (Renzema, 1992), there is still little known about its effectiveness as an alternative to incarceration or in protecting public safety by reducing rates of reoffending (Renzema, 2003; Tonry and Lynch, 1996; Vollum and Hale, 2002). Much of the literature on EM has been focused on descriptions of its history and the legal and ethical issues associated with its use (Beck et al., 1990; Berry, 1985; Blomberg et al., 1987; Brown and Elrod, 1995; Corbett and Marx, 1991; Erwin, 1990; Gainey and Payne, 2003; Gowen, 2001; Johnson et al., 1989; Payne and Gainey, 1999, 2000; Renzema, 1991; Schmidt, 1991).

Concerning the notable paucity of empirical evidence regarding EM's effectiveness, Gainey et al. (2000) conclude that research has not kept pace with the rapid implementation of the penal strategy, a conclusion reiterated by Vollum and Hale (2002) and Renzema (2003) in their reviews of that research. A meta-analysis of the effect of various intermediate sanctions on recidivism by Gendreau et al. (2000) includes only six effect size estimates for EM (4% of the total number of estimates included in the analysis), estimated from data on only 1,414 offenders (2.6% of the total number of offenders included in effect size estimates for all intermediate sanction types). Moreover, their findings for the effect of EM across these estimates ranges from a 2% decrease to an 8% increase in the rate of recidivism for offenders on EM when the relative number of estimates is taken into account.

In a more recent review and summary of the literature specific to the effect of EM on recidivism for moderate- and high-risk offenders, Renzema and Mayo-Wilson (2005) present a study search flow chart that illustrates the process by which a total of 154 EM studies is reduced to only three that meet their—somewhat stringent—criteria for selection (Bonta et al., 2000b; Finn and Muirhead-Steves, 2002; Sugg et al., 2001). They report their findings (based on the three qualifying studies) as "grim" (Renzema and Mayo-Wilson, 2005:230), concluding "no overall impact on recidivism." However, a less critical eye might characterize the findings as "mixed" or "inconclusive." Bonta et al. (2000b) find a positive effect of EM on treatment program completion and a negative effect of a combination of EM and treatment on recidivism. Finn and Muirhead-Steves (2002) find no effect of EM on recommitment to prison within three to four years after release from parole, but they report that none of the parolees in their sample returned to prison while on EM, and of the 128 parolees assigned to EM, 97 (75%) completed the EM program with no violations, 25 (20%) completed with "some" violations, and only 6

(5%) failed to complete. The third study reviewed by Renzema and Mayo-Wilson (Sugg et al., 2001) looked at reconviction rates for offenders who received curfew orders with EM in Norfolk, Manchester, and Reading, U.K., and found virtually equal two-year reconviction rates for offenders in the curfew/EM group and those in the control group. Despite the fact that these three studies met the criteria set by Renzema and Mayo-Wilson, some decided limitations and weaknesses are associated with them as well.

In addition to the three studies reviewed by Renzema and Mayo-Wilson (2005), we identified six studies that examined the effect of EM on recidivism and whose findings we think are relevant to this research and similarly mixed and inconclusive. Three of the six studies find a negative effect of EM on recidivism, measured as re-arrest while on supervision (Coopri-der and Kerby, 1990; Gainey et al., 2000; Jolin and Stipak, 1992), two studies report no effect (Bonta et al., 2000; Courtright et al., 1997), and one study (Cadigan, 1991) finds significantly higher rates of re-arrest for pre-trial releasees on EM as compared with national rates.

Another possible "outcome" of EM is absconding, or escape from supervision, which represents another threat to public safety, in that the absconder's whereabouts are unknown to his/her probation officer and his/her activities unmonitored by any device—electronic or human. The literature on EM, however, has yet to address this outcome, with the exception of the two studies of pre-trial release offenders and the effect of EM on the likelihood of their failure to appear (FTA) (Cadigan, 1991; Coopri-der and Kerby, 1990). Like the findings for EM and technical viola-tions, the findings for EM and FTAs contradict each other. Cadigan (1991) finds higher rates for offenders on EM, whereas Coopri-der and Kerby (1990) find lower rates.

In sum, the prior research on EM's outcomes indicates mixed evidence for its effectiveness in reducing the likelihood of "failure" for offenders on community supervision and weak evidence for its effect on widening the net of penal control. In addition, the data and methodological limitations of the prior studies demonstrate the clear need for more rigorous empirical research, including richer data on a larger and more representative sample and an analytical technique that takes into account the complex nature of the experience of convicted offenders on home confinement and EM and their increased likelihood of recommitment to prison, or other incapacitative events, within the span of time from their admission to the follow-up period. The current study overcomes the limitations of previous studies by using data on a five-year cohort of offenders placed on home confinement in Florida ($N = 75,661$). More importantly, this study takes advantage of the precision of coefficient estimation afforded by propor-tional-hazards regression and its accommodation of time-varying indepen-dent and dependent variables to model the supervision failure outcomes of

revocation for a technical violation, revocation for the commission of a new offense, and absconding from supervision.

THE CURRENT STUDY

In 1983, Florida became the first state to legislate and implement a statewide home confinement program specifically designed to address the problem of exponential increases in prison admissions and the need for intermediate sanctions as an alternative to incarceration (Florida Depart-ment of Corrections (FDOC), 2001). As the program developed and admissions increased, it became apparent that even within the narrower category of home confinement, different offenders required different levels of supervision intensity and surveillance while on the program. Vari-ous approaches to case management were tried and revised, but the advent of RF technology as a viable option for closer surveillance of higher risk offenders reframed the issue and ushered in a second phase of home confinement supervision strategy in which EM became the primary differentiating factor in the treatment of offenders.

Since 1987, with legislative approval, the FDOC implemented RF moni-toring as an additional surveillance technique for offenders on home con-finement, and in 1998 the use of EM was expanded to include GPS monitoring for those offenders judged to be of higher risk to public safety and in need of an even higher level of surveillance while in the community. According to the FDOC (2003):

The additional features of inclusive and exclusive boundaries, two-way communication with the victim or the offender, location mapping for archives retrieval, immediate tamper notification and remote laptop tracking with a wireless modem for constant communication with the monitoring center, makes the GPS system the best available. It would seem logical that violations of home confinement would decrease because offenders would know in advance that violations are tracked in "near real time" 24 hours a day.

METHODS

The data for this study were drawn from the FDOC's Offender-based Information System (OBIS). The sample comprises 75,661 offenders placed on home confinement from 1998 to 2002. These "placements" include original sentences to home confinement, split sentences (prison followed by supervision) to home confinement, post-prison sentences (Home Confinement—Parole), and sentences to home confinement for a violation of probation.

As noted, one method for assessing the effectiveness of EM as an alter-native to incarceration versus an enhancement that results in net-widening

has been the comparison of the relative "risk" to public safety of offenders sentenced to EM and offenders sentenced to community supervision without EM (Bonta et al., 2000; Gendreau et al., 2000; Renzema, 2003; Renzema and Mayo-Wilson, 2005). The logic underlying this kind of comparative analysis is that evidence of offenders who are sentenced to this new alternative being of no greater "risk" than offenders sentenced to the previously existing community supervision sanction lends support to a net-widening rather than an alternative-sentencing argument. Although Bonta et al.'s (2000) analysis uses the results of a self-reported questionnaire to measure "risk" as the Level of Service Inventory-Revised (LSI-R) score, they rely on a much broader definition of offender risk in making their case for a "net-widening" effect of EM. In their assessment of the prior research on EM and the relative risk of offenders placed on EM, Bonta et al. (2000) include factors such as prior record (Ball et al., 1988; Cadigan, 1991), violent versus nonviolent primary offense (Baumer et al., 1993; Maxfield and Baumer, 1990), DUI or traffic offenders only (Lilly et al., 1993), first offense or property offenders (Mortimer and May, 1997), or other "low-risk" offenders (Beck et al., 1990; Ontario, 1991; Roy, 1997; Whittington, 1987) as indicators of relative risk. A more recent examination of EM and front-end net-widening (Renzema and Mayo-Wilson, 2005) follows this course as well, defining offender "risk" in terms of prior record and primary offense convicted of. For this study, we use primary offense type (violent/not violent) and Florida's sentencing guidelines scoring system as indicators of offender "risk" and contend that equal levels of risk for offenders on EM and offenders on home confinement without EM supports the net-widening argument as it applies to the imposition of harsher sentences, or "front-end" net-widening.

Although sentencing guidelines are just that—guidelines—from which judges can and do depart, Florida's sentencing guidelines scoring system has been shown to serve as a valid indicator of "offender seriousness" (Burton et al., 2004). The weighted score produced by this system takes into account an offender's primary offense and all additional offenses, his or her prior record and the seriousness of prior offenses, and other circumstances of the criminal event (victim injury, weapon use, supervision violation, etc.). In the absence of risk scores derived from psychological or other such inventories, this indicator of offender seriousness is the best available quantitative measure of the risk an offender poses to public safety.

To test the effectiveness of EM in reducing the likelihood of failure while on home confinement, three outcome measures—revocation for a technical violation, revocation for a new offense, and absconding—were modeled using proportional-hazards regression (survival) analysis. This statistical modeling technique allows for right-censoring and the inclusion

of time-varying independent variables as well as taking into account "time to failure" in the estimation of maximum-likelihood coefficients. The time variable used is weeks from placement on home confinement to release, and cases were right-censored on the week of the release event. When the release event was due to something other than one of the three types of supervision "failure" listed above—death, successful termination of supervision, sentence reduction to regular probation, etc.—the offender was considered "at risk" of failure for each of the weeks before that event and then right-censored, or dropped from the analysis. In the event of an offender remaining on active home confinement beyond 105 weeks (2 years) from placement, the case was censored at week 105.

MEASURES

DEPENDENT VARIABLES

As discussed, three outcomes of a period of supervision are modeled in the multivariate survival analysis—revocation for a new offense,¹ revocation for a technical violation, and absconding from supervision. In Florida, a revocation results from a court decision to terminate supervision for failing to meet the requirements of supervision. It is not, necessarily, an indicator of offender behavior, but an indicator of "getting caught" and the subsequent community supervision officer and judicial response. Both types of revocation—for a new offense or for a technical violation—are considered permanent releases, although many offenders are returned to community supervision with a new sentence. Absconding, on the other hand, does not in and of itself constitute a permanent release. FDOC (2005) defines absconding as follows: "Offender absconds from supervision; the whereabouts are unknown and the court issues a warrant for violation of supervision." Supervision may or may not be terminated upon return from absconding. In this analysis, absconding is treated as a separate "outcome," and an offender is still considered "at risk" for revocation after an absconding event.

For this study, a separate analysis was conducted for each of the three "outcome" measures—revocation for a technical violation, revocation for a new offense, and absconding. These outcome variables were dichotomized so that the value is zero for all weeks that an individual offender is at risk of the unsuccessful outcome but does not experience the event, and the value is 1 for the one week in which he or she does experience the unsuccessful outcome. As noted, for offenders who experience a release

1. For this analysis, we collapsed the categories of "revocation for a new misdemeanor" and "revocation for a new felony" into one outcome variable, "revocation for a new offense." Nearly three-quarters (71%) of revocations for a new offense are for a new felony offense.

event other than one of these "failures," the outcome variable is coded 0 for all weeks up to the week of release, at which point the case is dropped from the analysis.

INDEPENDENT VARIABLES

The variable of primary interest in this analysis is whether the offender was placed on EM while on home confinement. Two dichotomous, time-varying variables were created to indicate time on EM in any given week—one for RF monitoring and one for GPS monitoring—to determine whether one device type has a greater or lesser effect on the outcome variables than the other. Comparing the relative effectiveness of the two types of EM is important because one type—GPS monitoring—involves considerably more intensive and precise surveillance than the other (RF monitoring), which means it should be more effective in deterring and incapacitating the offender and more likely to "catch" offenders violating the conditions of their community supervision. Unlike RF monitoring, which only provides surveillance when the offender is in his or her home, GPS monitoring tracks the location of offenders and maps their whereabouts for retrieval by the community supervision officer (FDOC, 2005).

CONTROL VARIABLES

Several variables were included in the analysis to control for any offender characteristics, criminal history, or current period of supervision circumstances related to the likelihood of an unsuccessful outcome. Table 1 presents descriptive statistics for these control variables, the dependent variables, and the two independent variables of primary interest. Each of the time-varying independent variables, like the dependent variables, was dichotomized so that its value is 1 for any week in which the condition applies to an individual offender and 0 for any week in which it does not. In Table 1, these variables are presented in terms of the proportion of the entire sample to which the condition applied at any time during the risk period.

TABLE 1. DESCRIPTIVES OF VARIABLES INCLUDED IN THE ANALYSIS

Variables	Definition/Coding	Percentage/ Mean
<i>Dependent Variables</i>		
Revocation for a technical violation	Yes = 1 for week in which revocation event occurs; 0 for each week prior	30.4%
Revocation for a new offense	Yes = 1 for week in which revocation event occurs; 0 for each week prior	10.8%

Variables	Definition/Coding	Percentage/ Mean
Absconded from supervision	Yes = 1 for week in which absconding event occurs; 0 for each week prior	15.7%
<i>Electronic Monitoring</i>		
Radio-frequency monitoring	On radio frequency monitoring [Yes = 1 for week(s) on]	4.3%
GPS monitoring	On global positioning system monitoring [Yes = 1 for week(s) on]	3.0%
<i>Offender Characteristics</i>		
Male	Sex (male = 1)	78.8%
White	Race (white = 1)	59.7%
Age at admission	Age in years	30.74
Residency confirmed	Address is confirmed and permanent (yes = 1)	51.6%
Employed while supervised	Employed for at least one day during period of supervision (yes = 1)	66.0%
<i>Offender History/Prior Record</i>		
Habitual offender	Habitual offender (yes = 1)	1.0%
Habitual violent offender	Habitual violent offender (yes = 1)	0.0%
Ever committed to prison	Ever been committed to prison in Florida (yes = 1)	18.9%
Prior commitments to supervision	Number of prior supervision commitments in Florida	1.46
Ever absconded	Ever absconded from supervision (yes = 1)	17.6%
Ever revoked—felony	Supervision ever revoked for a new felony (yes = 1)	17.6%
Ever revoked—misdemeanor	Supervision ever revoked for a new misdemeanor (yes = 1)	7.9%
Ever revoked—technical	Supervision ever revoked for a technical violation (yes = 1)	32.7%
HC placement for VOP	Place on HC for a violation of probation (yes = 1)	40.6%
<i>Court-Ordered Conditions of Supervision</i>		
Domestic violence treatment	Participate in domestic violence treatment (yes = 1)	0.6%
Education/GED	Educational program and/or prepare for GED (yes = 1)	1.4%
Psychological treatment	Receive psychological treatment (yes = 1)	2.5%
Residential drug treatment	Residential drug treatment (yes = 1)	6.1%
Sex offender treatment	Participate in sex offender treatment (yes = 1)	0.9%
Drug testing	Drug testing (yes = 1)	50.6%
Outpatient drug treatment	Participate in outpatient drug treatment (yes = 1)	16.9%
Public service hours	Complete public service hours (yes = 1)	2.3%
<i>Current Sentence</i>		
HC-Parole	HC under parole supervision (yes = 1)	0.1%
HC-Sex Offender	Sex Offender HC (yes = 1)	0.8%
Serving split sentence	Split sentence (yes = 1)	2.1%

Variables	Definition/Coding	Percentage/ Mean
Sentence days	Number of days offender sentenced to HC	1715.2
Mitigated	Sentence mitigated—Scored to prison under guidelines (yes = 1)	34.1%
Circuit of supervision	Judicial circuit offender supervised at time of "failure" or successful completion	*
<i>Current Offense</i>		
Murder	Primary offense is murder/manslaughter (yes = 1)	0.7%
Sex offense	Primary offense is a sex offense (yes = 1)	3.9%
Robbery	Primary offense is robbery (yes = 1)	3.3%
Other violent	Primary offense is other violent (yes = 1)	14.0%
Burglary	Primary offense is burglary (yes = 1)	10.9%
Theft	Primary offense is theft (Yes = 1)	19.2%
Drug offense	Primary offense is a drug offense (reference category)	34.2%
Weapons offense	Primary offense is a weapon offense (yes = 1)	2.5%
Other offense	Primary offense is "other" offense (yes = 1)	9.4%
Principal	Offender was principal in a completed act (yes = 1)	97.7%
Counts	Number of offenses convicted of	1.1
<i>Time-Varying Factors in Current Period of Supervision</i>		
Abscond event	Offender in absconder status [yes = 1 for week(s) in]	17.4%
Treatment event	Offender in residential drug treatment [yes = 1 for week(s) in]	14.1%
Drug court event	Offender participating in drug court [yes = 1 for week(s) in]	4.1%
Non-reporting event	Offender in non-reporting status [yes = 1 for week(s) in]	46.7%

*Overall, 19 of the 20 dummy variables were included in the models. Distributions of cases by circuit available on request from the authors.

In all, 62 independent variables are included in each proportional-hazards regression model, not counting the reference categories for the multinomial variables. The control variables include indicators of community supervision success or failure in the following categories: sociodemographic characteristics of the offender and his/her criminal history and prior record; factors related to the offender's current term of community supervision, including any conditions and provisions of supervision or other sentence-event differences, the judicial circuit in which the offender was being supervised, and the primary offense for which the offender was convicted; and time-varying events that occurred within the period of supervision being examined that resulted in the offender avoiding surveillance by the community supervision officer (i.e., absconding), being subjected to a greater level of surveillance (i.e., participating in drug court), or

being incapacitated (i.e., in a residential drug treatment facility or in the county jail) for a certain amount of time.

Sociodemographic Characteristics of the Offender. Age, race, and sex are included to control for the well-established relationship between these demographic characteristics and success or failure on community supervision. The two additional sociodemographic variables, permanence of residency and employment status, are included as measures of the offender's lifestyle stability.² Marital status, another measure of lifestyle stability, was not included in our models. Although data on marital status are collected for offenders committed to prison in Florida (however, it is worth noting that, on average, 36% of those data is missing), they are not collected for offenders placed on community supervision. It is expected that if data were available, the variable for marital status would have an effect on the likelihood of an offender being revoked or absconding; however, it does not necessarily follow that this effect would prove a source of spuriousness for the EM—revocation or EM—absconding relationship. Further research on the differences between offenders placed on EM and those not placed on EM is needed to determine whether lifestyle and stability measures have an effect on both the likelihood of placement on EM and the likelihood of revocation or absconding.

Current Primary Offense. Three aspects of the offender's current primary offense³ were included in the multivariate models: primary offense category, whether the offender was the principle in a completed act (not an attempt), and the number of counts for which he or she was convicted. For this analysis, the "primary" offense was coded according to the nine-group categorization of offenses established and used by the FDOC: (1) murder/manslaughter, (2) sex offenses, (3) robbery, (4) other violent/personal offenses, (5) burglary, (6) theft, (7) drug offenses, (8) weapons offenses, and (9) "other" offenses. Current primary offense category dummy variables are included in the multivariate analyses to control for the known association between offense type and community supervision outcome.

Current Sentence. In addition to current offense, we include aspects of the current sentence as control variables in the multivariate models. Within

2. Data on offender marital status and educational level were not available, nor were data on arrest history, prior county jail incarceration, or prior convictions in other states.

3. Each placement on home confinement is associated – by a unique (to the offender) "prefix" code – to a particular sentencing event. An offense is designated as primary for that sentencing event by means of a formula that takes into account the seriousness of the offense (according to the offense code), the level of the charge (Capital Life, Life, 1st, 2nd, 3rd –degree felony, etc.), and the associated sentence length.

the home confinement program as a whole, separate conditions are mandated for offenders placed on sex-offender home confinement and for post-prison releasees on home confinement—parole. The particular circumstances and characteristics of these offenders are controlled for with dummy variables for the type of home confinement to which the offender is sentenced. Similarly, offenders serving a split (prison then home confinement) sentence and those originally sentenced to probation and later placed on home confinement for a violation are taken into account with dummy variables for “split sentence” and “home confinement placement for VOP,” respectively. Sentence length is controlled for with a continuous variable measuring the number of days the offender was sentenced. If the offender’s sentence was mitigated, meaning that he or she “scored” to prison according the Florida’s Sentencing Guidelines but was sentenced to home confinement instead, we take that into account with a dummy variable where mitigated is 1.

Conditions of Supervision. In addition to the standard conditions of home confinement, offenders can be held to several special provisions stipulated by the sentencing judge. These include participation in a treatment program (domestic violence, psychological, drug, and/or sex offender treatment), participation in an educational program, regular drug testing, and/or the completion of public service hours, all of which are controlled for with dummy variables to indicate whether the provision was court-ordered. Dates of attendance for outpatient treatment programs were not available, but time-varying variables reflecting weeks in which an offender was in residential drug treatment or participating in drug court were included to control for the incapacitation effect of residential drug treatment, at which point an offender would be at lower risk for reoffending or absconding, and for the more stringent conditions imposed on offenders in drug court, who would be at greater risk for technically violating during that time. A time-varying variable for “non-reporting status” is also included, as this status is assigned to offenders who are temporarily incarcerated in a county jail and, therefore, at lower risk for reoffending or absconding.

Circuit of Supervision. Finally, the judicial circuit in which the offender was being supervised is included to control for local-level discretion in “violation” policies and judicial decision making. If and when an offender was transferred from one circuit to another, the circuit variable was recoded to reflect that change for the week in which it took place and the weeks thereafter. The distribution of home confinement placements by circuit is available from the authors upon request.

FINDINGS

TABLE 2. OFFENDER SERIOUSNESS AND LEVEL OF CONTROL—PERCENTAGES AND MEAN VALUES FOR OFFENDERS WITH AN ORIGINAL SENTENCE TO HOME CONFINEMENT

	No Electronic Monitoring	Electronic Monitoring		
		Total EM	Radio Frequency	GPS
Primary offense was violent	19.4%	38.7%*	30.9%	53.8%**
Scored to prison under sentencing guidelines	30.1%	46.2%*	41.8%	54.6%**
Scored to prison for primary offense only	21.0%	38.5%*	33.3%	48.5%**
Mean sentencing guidelines points for primary offense	36.5	59.7*	49.7	79.0**
Mean total sentencing guidelines points	42.4	64.4*	54.3	83.8**
N	45,475	3,347	2,203	1,144

*Difference between EM and No EM (RF and GPS combined) is statistically significant ($p < 0.001$).

**Difference between RF and GPS is statistically significant ($p < 0.001$).

Table 2 presents the results of an analysis comparing the relative “risk” of EM offenders and offenders sentenced to home confinement without EM as indicated by the nature of their primary offense (violent vs. not violent) and their calculated sentencing guidelines scores. Specifically, the sentencing guidelines score sheet data are used in four different ways to measure the level of risk: whether offenders scored to a recommended prison sentence (total points of 45 or greater), whether the primary offense points alone resulted in a recommended prison sentence, the mean points for the primary offense, and the mean total guidelines points.

The comparisons of risk, or seriousness, levels for home confinement offenders with and without EM in Table 2 show that for all five measures, EM offenders have statistically significant ($p < 0.001$) higher levels than those for offenders not on EM. Additionally, offenders on the higher level surveillance EM modality of GPS have risk levels significantly higher ($p < 0.001$) than those under the less controlling RF monitoring mechanism on all five risk measures. Specifically, EM offenders are more likely to have committed a violent offense and more likely to have “scored” to prison, and their mean sentencing guidelines points scores, in terms of the total points and points for the primary offense alone, are significantly higher than those for offenders sentenced to home confinement without EM ($p < 0.001$).

Whether these offenders would have received a sentence to prison in the absence of the EM alternative is impossible to determine using secondary data. However, these findings do indicate that offenders on EM are, on average, more serious offenders, and their perceived risk to the community makes them more likely than offenders not on EM to be sentenced to prison in the absence of the EM alternative. Additionally, offenders sentenced to supervision under the enhanced level of offender control through GPS instead of RF are clearly more serious offenders and considered more of a risk to public safety.

TABLE 3. OFFENDER SERIOUSNESS AND LEVEL OF CONTROL—PERCENTAGES AND MEAN VALUES BY PRIMARY OFFENSE CATEGORY

	No Electronic Monitoring	Total EM	RF	GPS
<i>Primary Offense = Violent</i>				
Scored to prison under sentencing guidelines	58.5%	72.7%*	69.3%	76.6%**
Scored to prison for primary offense only	50.3%	68.2%*	64.1%	72.7%**
Mean sentencing guidelines points for primary offense	62.5	100.0*	85.6	116.0
Mean total sentencing guidelines points	66.6	103.1*	88.3	119.4**
N	8,798	1,295	680	615
<i>Primary Offense = Property</i>				
Scored to prison under sentencing guidelines	28.0%	38.9%*	39.9%	36.2%
Scored to prison for primary offense only	21.0%	31.4%*	32.2%	29.4%
Mean sentencing guidelines points for primary offense	33.4	41.3*	41.2	41.5
Mean total sentencing guidelines points	39.0	46.3*	45.7	47.7
N	13,771	831	596	235
<i>Primary Offense = Drug</i>				
Scored to prison under sentencing guidelines	21.0%	23.5%	23.4%	23.6%
Scored to prison for primary offense only	10.3%	12.7%*	12.2%	14.1%
Mean sentencing guidelines points for primary offense	30.2	32.0*	31.4	33.7
Mean total sentencing guidelines points	36.3	37.4	36.7	39.4
N	17,038	844	624	220

*Difference between No EM and EM (RF and GPS combined) is statistically significant at $p < 0.001$.

**Difference between RF and GPS is statistically significant $p < 0.001$.

To further address the issue of whether EM is, in fact, an alternative to prison at the "front-end" of the sentencing process, Table 3 presents the same four sentencing guidelines indicators of the seriousness of the offender's current and past criminal actions, separately, within each of three primary offense categories (violent, property, and drug). Within the

violent and property crime categories, the results are similar to those reported in Table 2 for all offenders. Specifically, offenders on home confinement with EM of either type exhibit significantly higher risk scores on all four of the sentencing guidelines measures ($p < 0.001$). However, mixed results are found in comparisons of those offenders monitored under GPS versus RF. For violent offenders, GPS offenders are found to be significantly more serious and pose a greater risk to the community ($p < 0.001$) than those under RF surveillance for the same four measures as for all offenders combined. For property offenders, however, the differences virtually disappear, indicating no greater risk to the community from offenders placed on the more intensive GPS monitoring than from those placed on RF monitoring.

For drug offenders, the picture is slightly different when comparing EM with non-EM offenders and RF with GPS offenders. Although drug offenders on EM are significantly more likely to have scored to prison for their primary offense and to have higher mean sentencing guidelines points for their primary offense, the differences disappear when the total sentencing points are compared. Furthermore, regardless of statistical significance, differences for all four measures are considerably smaller for drug offenders than for violent or property offenders, both for EM vs. non-EM offenders and RF vs. GPS offenders.

This analysis of the differences in risk levels, or offender seriousness, between offenders on home confinement with and without EM and between those on RF versus the more intrusive GPS provides no clear evidence that, overall, the decision to monitor offenders on home confinement with enhanced electronic control mechanisms results in "front-end" net-widening. In other words, offenders sentenced to home confinement with EM seem to have posed a significantly higher risk to public safety and would have had a higher likelihood of receiving a prison sentence if not for the availability of EM as an enhanced control mechanism. However, possible "front-end" net-widening for drug offenders is suggested by findings that show that non-EM drug offenders exhibit almost equivalent levels of risk to the public as those placed on the more controlling EM program.

Although the question of "front-end" net-widening remains somewhat elusive, relying on proxies to indicate the likelihood of a sentence to prison in the absence of qualitative data on judicial decision making, the question of "back-end" net-widening is more easily addressed and answered. With the introduction of intermediate sanctions into the continuum of punishment alternatives came the concern that these more intensive forms of community supervision, with their stricter conditions and closer surveillance, would increase the likelihood of an offender violating

those conditions and getting caught doing so. The fact that EM of offenders, on home confinement constitutes the last option before prison, it seems likely that a violation while on EM would result in a sentence to prison, therefore, widening the net.

Table 4 presents the results of the proportional-hazards regression modeling of the likelihood of revocation for a technical violation within two years (104 weeks) of the first day of placement on home confinement for the 75,661 offender placements in our sample.⁴ The parameter estimates tell us the direction of the effect of the independent and control variables on the likelihood of revocation and whether those effects are statistically significant. The hazard ratios indicate the relative likelihood of revocation and time to revocation across categories or values of the independent and control variables. These ratios can be converted to percentage differences with the formula: $(1 - \text{Hazard Ratio}) * 100$.

TABLE 4. PROPORTIONAL-HAZARDS MODEL RESULTS FOR LIKELIHOOD OF REVOCATION FOR A TECHNICAL VIOLATION

Variables	Total Sample		Violent Offenders		Property Offenders		Drug Offenders	
	Parameter Estimate	Hazard Ratio						
RF monitoring	-3.135***	0.043	-3.430***	0.032	-2.905***	0.055	-3.268***	0.038
GPS monitoring	-2.322***	0.098	-2.444***	0.087	-2.549***	0.078	-1.875***	0.153
<i>z-score for difference in coefficients</i>	2.962***		2.052**		0.637		2.501**	
Murder	-0.492***	0.612						
Sex offense	-0.004	0.996						
Robbery	-0.057	0.944						
Other violent	-0.154***	0.857						
Weapons offense	-0.168***	0.846						
Other offense	-0.139***	0.870						
N	74,276		16,586		22,801		25,885	

NOTE: Models shown include all offender demographic and prior record variables, as well as current term of supervision control variables, as displayed in Table 1.

** $p < 0.01$.

*** $p < 0.001$.

4. Due to space limitations, the full models, including all the control variables presented in Table 1, are not presented in Tables 4, 5, and 6, and only the variables of interest are displayed. The results of the full models are available from the authors.

Unlike previous findings of a "surveillance" effect, our findings indicate that offenders on EM are *less* likely to be revoked for a technical violation. In fact, and surprisingly, offenders on RF monitoring are 95.7% less likely and offenders on GPS monitoring are 90.2% less likely than offenders on home confinement without EM to be revoked for a technical violation. However, the difference in the magnitude of these effects is statistically significant (z -score = 2.962), which partially supports the "surveillance effect" hypothesis, in that offenders on the more intense form of electronic surveillance are more likely than those on the less intense form to get caught violating the conditions of their home confinement sentence. Secondly, these findings show that the prohibitory effect of EM on technically violating holds true for offenders in all three primary offense categories and has virtually the same degree of effect across the three categories. This finding is of particular significance given that EM is used at a considerably higher rate for violent offenders (12.3%, compared with 5.7% and 4.7% for property and drug offenders, respectively) and that violent offenders are significantly less likely than property and drug offenders (the reference category) to be revoked for a technical violation, whether they are placed on EM.

Although one set of concerns related to the addition of EM to home confinement sentences is that it will widen the net of control, another set of concerns has to do with public safety and the effectiveness of EM in deterring or incapacitating offenders living in the community. Table 5 presents the results of the proportional-hazards regression modeling of the likelihood of revocation for a new offense, our primary measure of risk to public safety, and Table 6 presents those same results for the likelihood of absconding from supervision, a second measure of offender risk to public safety.

For both outcomes, the results show that EM significantly reduces the likelihood of failure and that the degree to which that likelihood is reduced is about the same for revocation for a new offense and absconding. For the total sample of offenders, the hazard ratio of 0.053 for both forms of EM and revocation for a new offense indicates a 94.7% reduction in the likelihood of revocation for offenders on RF or GPS versus no form of electronic surveillance. For violent, property, and drug offenders, the percent reduction ranges from 89.8 to 98.1 for RF and from 91.4 to 95.5 for GPS (note that within the category of drug offenders, too few offenders were placed on GPS monitoring and committed a new offense to produce a valid parameter estimate). Although these figures indicate that RF monitoring is slightly more effective than GPS monitoring and slightly more effective for violent than for property or drug offenders, the overall range in the rate of reduction for EM versus no EM is very small. Where public

TABLE 5. PROPORTIONAL-HAZARDS MODEL RESULTS FOR LIKELIHOOD OF REVOCATION FOR A NEW OFFENSE

Variables	Total Sample		Violent Offenders		Property Offenders		Drug Offenders	
	Parameter Estimate	Hazard Ratio						
RF monitoring	-2.933***	0.053	-3.947***	0.019	-2.283	0.102	-3.146***	0.043
GPS monitoring	-2.929***	0.053	-2.461***	0.085	-3.097**	0.045	-15.150+	0.000
<i>z-score for difference in coefficients</i>	0.007		1.327		0.753		0.024	
Murder	-0.882***	0.414						
Sex offense	-0.593***	0.552						
Robbery	-0.089	0.915						
Other violent	-0.223***	0.800						
Burglary	-0.003	0.997						
Other property	0.073*	1.076						
Weapons offense	-0.177*	0.838						
Other offense	0.114**	1.121						
	74,276		16,586		22,801		25,885	

NOTE: Models shown include all offender demographic and prior record variables, as well as current term of supervision control variables, as displayed in Table 1.

+ Too few cases to produce a valid parameter estimate.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

safety is concerned, either form of electronic surveillance seems to significantly reduce the likelihood of reoffending for all three "types" of offender.

Absconding from supervision is an outcome measure that has not, to date, been addressed in the literature on EM and home confinement. However, its implications for public safety and the relative frequency of its occurrence suggest that absconding and the potential for EM to reduce the rate of absconding should be considered in any test of the effectiveness of EM. As of December 30, 2004, more than 40,000 of the 114,891 offenders on community supervision in Florida were classified as absconders, their "whereabouts unknown" (FDOC, 2005). In our sample of 75,661 offenders on home confinement, 11,857 (15.7%) absconded from supervision at some point within two years of placement, and 1,911 (16.1%) of those absconders were subsequently revoked for a new offense. This rate of reoffending is considerably higher than that for the offenders who did not

TABLE 6. PROPORTIONAL-HAZARDS MODEL RESULTS FOR LIKELIHOOD OF ABSCONDING

Variables	Total Sample		Violent Offenders		Property Offenders		Drug Offenders	
	Parameter Estimate	Hazard Ratio						
RF monitoring	-2.426**	0.088	-3.149**	0.043	-1.869**	0.154	-2.571**	0.076
GPS monitoring	-2.325**	0.098	-2.432**	0.088	-1.899**	0.150	-2.203*	0.111
<i>z-score for difference in coefficients</i>	0.260		0.855		0.051		0.424	
Murder	-1.247*	0.287						
Sex offense	-0.538*	0.584						
Robbery	-0.198*	0.821						
Other violent	-0.289*	0.749						
Burglary	0.012	1.012						
Other property	-0.045	0.956						
Weapons offense	-0.497*	0.608						
Other offense	-0.198*	0.820						
N	74,276		16,586		22,801		25,885	

NOTE: Models shown include all offender demographic and prior record variables, as well as current term of supervision control variables, as displayed in Table 1.

* $p < 0.01$.

** $p < 0.001$.

abscond (9.8%) and indicates an increased risk to public safety of offenders who escape surveillance. The results presented in Table 6 indicate that EM also has a prohibitive effect on the likelihood of absconding, with statistically significant parameter estimates and hazard ratios of 0.088 and 0.098 for RF and GPS monitoring, respectively. Again, this effect is the same for both types of EM (z -score for difference in coefficients is 0.260) and varies only slightly for the three categories of primary offense type.

These findings consistently demonstrate that either form of EM significantly reduces the risk to public safety from offenders living in the community. Moreover, our findings for the effect of EM on the likelihood of revocation for a technical violation indicate that rather than widening the net of penal control, the addition of electronic surveillance to a home confinement sentence may actually reduce the probability of eventual imprisonment and, therefore, effectively serve as a useful alternative sanction. Notwithstanding the limitations of using official data to represent the complex circumstances of offenders serving a sentence to home confinement and the complexities of officer discretion and judicial decision making,

these findings suggest that the dual goals of reducing the number of admissions to prison while protecting public safety may, in fact, be achieved via the introduction of newer and more refined means of offender surveillance.

SUMMARY AND DISCUSSION

The findings reported here have addressed two questions related to the net-widening effect and the public safety effectiveness of EM for offenders on home confinement. With regard to net-widening, the findings provide only scant support for a net-widening effect resulting from the addition of EM into Florida's home confinement program. Using primary offense type (violent or not) as the measure of offense seriousness, it was shown that those offenders on home confinement with EM were significantly more likely to have committed a violent offense as compared with those offenders on home confinement without EM. Additionally, those offenders on home confinement with GPS monitoring were even more likely than those with RF monitoring to have committed a violent offense. Furthermore, using sentencing guideline scores as an indicator of the likelihood of a prison sentence, offenders on home confinement with EM had a greater likelihood of a prison sentence than did offenders on home confinement without EM. Similarly, offenders on home confinement with GPS had a greater likelihood than those offenders on home confinement with RF of receiving a prison sentence in the absence of some form of home confinement. However, when all home confinement offenders were divided into primary offense type groups, the positive relationship between relative levels of control (i.e., No EM, RF, and GPS) and the likelihood of a prison sentence held true for violent and, to a lesser extent, property offenders, but significantly decreased for drug offenders, demonstrating that the net may, in fact, have widened for this group of offenders. With regard to "back-end" net-widening, EM was found to decrease rather than increase the likelihood of revocation for a technical violation, which contradicts the expectation of a surveillance effect.

In relation to public safety effectiveness, EM was found effective in reducing the likelihood of reoffending and absconding while on home confinement. Both RF and GPS significantly reduced the likelihood of revocation for a new offense and absconding from supervision, even when controlling for sociodemographic characteristics of the offender, current offense, prior record, and term of supervision factors and conditions. The use of GPS monitoring compared with the use of RF monitoring was found to be no more likely to reduce revocations or incidents of absconding. However, the use of either GPS or RF monitoring had virtually the

same inhibiting effect on revocations and absconding for violent, property, and drug offender groups on home confinement.

POLICY IMPLICATIONS

As for the policy implications of this research, it is important to note that the statistical modeling technique used in our analysis, proportional-hazards regression (survival analysis), takes into account the timing of an event and its occurrence in relation to the timing of placement on and removal from electronic monitoring. Therefore, our findings are limited to the effect of EM while the offender is actually being monitored, not after he/she completes the program. Although we agree with Renzema (2003:9) that "Many agencies using EM neither build rehabilitation components into their programs nor expect an enduring impact," and the informational literature related to EM that has been produced by the FDOC indicates the same, further study of the long-term effects of EM is needed before drawing conclusions about a rehabilitative effect or basing policy decisions on such an effect. However, our findings do indicate that home confinement with EM can effectively serve an incapacitation and/or deterrence role in protecting public safety.

Additional policy implications of this research include decision making regarding which offenders should be placed on EM, which type of monitoring device will be the most cost-effective and efficient, and the potential for front-end net-widening if states adopt a practice of "Got 'em?" Use 'em.'" The first two of these policy issues are addressed directly by the results of our analysis, whereas the third calls for further research and some monitoring of our own. Regarding decisions about which offenders should be placed on EM, our findings show that:

- 1 EM works for serious offenders—Much of the previous research has looked at less serious offenders, whereas we find an effect of EM on technical violations, reoffending, and absconding for a cohort of offenders judged too serious to be placed on regular probation. This overall finding bodes well for EM's anticipated use for sex offenders and other, more serious, offenders.
- 2 EM works equally well for all "types" of serious offenders, when offender type is defined as the category of the offender's primary offense (violent, property, or drug). Assuming that EM devices will not be available for every offender placed on community supervision, the decision about which offenders should be electronically monitored will need to be based on more than his or her primary offense. Further research should address the factors associated with success on EM.

As much of the new legislation related to EM is specific to sex offenders

on EM after release from prison, it should be noted that our findings also indicate that sex offenders are less likely than all other types of offenders to have their supervision revoked for a new offense or to abscond and no more or less likely to have their supervision revoked for a technical violation, even when controlling for EM status, and that less than 3% of the offenders in our sample were on parole or serving the second half of a split sentence.

Regarding decisions about which type of monitoring device is most cost-effective and efficient, our findings show that RF is just as effective as GPS in reducing the likelihood of an offender absconding or being revoked for a new offense and slightly more effective than GPS in reducing the likelihood of revocation for a technical violation. Given these findings, policy makers should consider whether GPS monitoring is worth its price. Although this study did not include an in-depth cost analysis of RF versus GPS or EM versus imprisonment, raw cost figures for EM in the State of Florida indicate considerable differences per diem. According to the FDOC (Brooks, 2005), the current (as of July 1, 2005) per diem cost for active GPS monitoring is \$8.97, as compared with \$1.97 for RF monitoring. The per diem cost for prison is \$51.22. At more than four times the cost, policy makers may want to reconsider their commitment to GPS over RF monitoring.

Finally, policy makers need to consider the potential for front-end net-widening as a result of states procuring great numbers of devices to meet the mandates of recent legislation and then keeping all of the devices "in service" regardless of real need. Lawmakers would do well to consider amendments to sentencing guidelines legislation that would specify a point range for which home confinement with EM would be the recommended sentence. In Florida, where there is no pre-trial risk-assessment instrument to guide the sentencing decision, such a point-range guideline would be preferable to blanket policies targeted at specific offender "types" or unlimited judicial discretion. Another option would be to limit the sentence of home confinement with EM to offenders who "score to prison" under the sentencing guidelines, thereby ensuring the use of EM only as a true alternative to incarceration. The adoption of the latter policy is unlikely, and in the end, we have to agree with Morris and Tonry (1990:218) that "all one can hope for is that the important desideratum of parsimony in punishment will restrain enthusiasms [for electronic 'tracking'] and respect autonomy."

THEORETICAL IMPLICATIONS

Such theoretical abstractions as net-widening, dispersal of discipline, transcarceration, carceral society, maximum security society, and culture of control have been proposed to capture and account for the reported

negative and unintended consequences of various penal reforms. Whether concerned with the disparity between the ideas and the policies of penal reform strategies or what these patterned disparities have meant in terms of larger or master penal control shifts, the focus has been on what was believed to be ever expanding penal control. Moreover, it is important to acknowledge that these theoretical abstractions have not only been useful in understanding certain aspects and potentials of penal reforms but were informed by some degree of empirical support for their negative and unintended consequences.

However, and as demonstrated by this study's findings for Florida's statewide home confinement and EM program, also salient *intended* outcomes are associated with this particular penal reform. The question, therefore, is do these findings mean "bad news" for the leading theoretical interpretations of penal reform, which have been largely focused on negative and unintended consequences? Or, alternatively, do these findings provide evidence that must and can be successfully confronted and interpreted by modifying and refining some of the existing penal reform theories? We believe it is the latter, and the task at hand is to reconcile unexpected findings of intended consequences with the larger, theoretical issues surrounding the concept of social control. Lianos's (2003:412) observation that "the question of control presents itself inevitably in the light – or should one say in the shadow?—of its social utility" is especially relevant here. Although EM seems to effectively thwart offenders from reoffending or otherwise threatening public safety, it simultaneously affords a degree of surveillance that would likely offend the sensibilities of the average, "free" citizen of the Western world.

EM presents a new challenge for both theorists and policy makers. Not only is the EM of offenders in the community an intermediate sanction and, therefore, a subject for discussion and debate in that context, it is also a means of surveillance made possible only by recent advances in computer and electronic technology. As such, the EM of offenders falls within the broader discussion of electronic surveillance in general. It seems likely that current and future developments in technology will result in ever increasing levels of personal transparency for both offenders and citizens alike. Certainly this possibility poses an increasingly urgent and important mandate—we need to confront these current and future control strategies and technologies with comprehensive and rigorous empirical, theoretical, and public policy scrutiny. The traditional "great divide" among research, theory, and public policy must be routinely bridged if responsible penal and public policies are to be implemented in this era of a technology-driven "culture of control."

In conclusion, this emerging technological culture of control poses both

positive and negative outcome potentials. Various new forms of technology can be used to produce more refined forms of control and regulation for offenders and citizens alike that are capable of not only negative and unintended consequences but also of being used in a manner that produces maximum desired results with minimum imposition. Technology makes it possible to control subjects in more discerning, less heavy-handed ways. Greater technological control capacities need not always result in more control. Rather, it depends on the uses to which these technologies are put, which depends, in turn, on the pressure that is placed on the control agents and technological methods to be used so that they are in accord with our civil liberties and social values. Consequently, systematic and responsible research is fundamental if we are to maximize the positive and minimize the negative potentials associated with the uses and impacts of these technologies in our fast changing culture of control (Garland, 2004).

REFERENCES

- Austin, James and Barry Krisberg
1981 Wider, stronger and different nets: The Dialectics of criminal justice reform. *Journal of Research in Crime and Delinquency* 18:165-196.
- Austin, James, John Irwin, and Charis E. Kubrin
2003 It's about time: America's imprisonment binge. In Thomas G. Blomberg and Stanley Cohen (eds.), *Punishment and Social Control*. 2d ed. New York: Aldine de Gruyter.
- Ball, Richard A., Ronald C. Huff, and J. Robert Lilly
1988 *House Arrest and Correctional Policy: Doing Time at Home*. Newbury Park, Calif.: Sage.
- Baumer, Terry L. and Robert I. Mendelsohn
1992 Electronically monitored home confinement: Does it work? In James M. Byrne, Arthur J. Lurigio, and Joan Petersilia (eds.), *Smart Sentencing: The Emergence of Intermediate Sanctions*. Newbury Park, Calif.: Sage.
- Baumer, Terry L., Michael G. Maxfield, and Robert I. Mendelsohn
1993 A comparative analysis of three electronically monitored home detention programs. *Justice Quarterly* 10:121-142.
- Beck, James, Jody Klien-Saffran, and Harold B. Wooten
1990 Home confinement and the use of electronic monitoring with federal parolees. *Federal Probation* 54:22-31.
- Berry, Bonnie
1985 Electronic jails: A new criminal justice concern. *Justice Quarterly* 2:1-22.
- Blomberg, Thomas G.
1977 Diversion and accelerated social control. *Journal of Criminal Law and Criminology* 68:274-282.
- Blomberg, Thomas G. and Karol Lucken
1994 Stacking the deck by piling up sanctions: Is intermediate punishment destined to fail? *The Howard Journal* 33:62-80.
- Blomberg, Thomas G. and Gordon P. Waldo
1987 Criminal justice reform and social control: Are we becoming a minimum security society? In John Lowman, Robert J. Menzies, and T. S. Palys (eds.), *Transcarceration: Essays in the Sociology of Social Control*. England: Gower Press.
- Blomberg, Thomas G., William Bales, and Karen Reed
1993 Intermediate punishment: Redistributing or extending social control? *Crime, Law, and Social Change* 19:187-201.
- Blomberg, Thomas G., Gordon P. Waldo, and Lisa C. Burcroff
1987 Home confinement and electronic surveillance. In Belinda R. McCarthy (ed.), *Intermediate Punishments: Intensive Supervision, Home Confinement and Electronic Surveillance*. Issues in Criminal Justice, Vol. 2. London: Willow Press.
- Blumstein, Alfred
1998 U.S. criminal justice conundrum: Rising prison populations and stable crime rates. *Crime & Delinquency* 44:127-135.
- Bonta, James, Suzanne Wallace-Captretta, and Jennifer Rooney
2000a Can electronic monitoring make a difference? An evaluation of three Canadian programs. *Crime & Delinquency* 46:61-75.
2000b A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior* 27:312-329.
- Brooks, Murray
2005 Personal correspondence.
- Brown, Michael P. and Preston Elrod
1995 Electronic house arrest: An examination of citizen attitudes. *Crime & Delinquency* 41:332-346.
- Burton, Susan E., Matthew Finn, Debra Livingston, Kristen Scully, William D. Bales, and Kathy Padgett
2004 Applying a crime seriousness scale to measure changes in the severity of offenses by individuals arrested in Florida. *Justice Research and Policy* 6:1-18.
- Cadigan, Timothy P.
1991 Electronic monitoring in federal pretrial release. *Federal Probation* 55:26-30.
- Clear, Todd and Anthony A. Braga
1995 Community corrections. In James Q. Wilson and Joan Petersilia (eds.), *Crime: Twenty-eight Leading Experts Look at the Most Pressing Problem of our Time*. San Francisco, Calif.: Institute for Contemporary Studies.
- Clear, Todd and George F. Cole
2003 *American Corrections*. 6th ed. Belmont, Calif.: Thomson-Wadsworth.
- Clear, Todd, Matthew White, and Kristen Presnell
1998 The offender in the community: Implications of the experience in the United States for the New Czech Republic. *Crime, Law and Social Change* 28:3-4,243-268.
- Coopridge, Keith W. and Judith Kerby
1990 Practical application of electronic monitoring at the pretrial stage. *Federal Probation* 54:28-35.

- Corbett, Ronald and Gary T. J.
1991 Critique: No soul machine: Technofallacies in the electronic monitoring movement. *Quarterly* 8:399-414.
- Courtright, Kevin E., Bruce L. Robert J. Mutchnick
1997 Effects of house electronic monitoring on DUI offenders. *Journal of Offender Behavior* 24:35-51.
- Erez, Edna, Peter R. Ibarra, and A. Lurie
2004 Electronic monitoring violence cases—A study of two bilateral programs. *Journal of Offender Behavior* 68:15-20.
- Erwin, Billie S.
1990 Old and new tools for probation officer. *Crime & Delinquency* 36:61-70.
- Finn, Mary A. and Suzanne Mes
2002 The effectiveness of monitoring with violent male parolees. *Justice Quarterly* 19:1-15.
- Florida Department of Corrections
2001 A Report on Control, Radio Frequency (RF) Monitoring and Global Positioning System (GPS) Monitoring: October 2001. Tallahassee, FL: Florida Department of Corrections, Bureau of Research and Data Analysis.
2003 A Report on Control, Radio Frequency (RF) Monitoring and Global Positioning System (GPS) Monitoring: October 2003. Tallahassee, FL: Florida Department of Corrections, Bureau of Research and Data Analysis.
2005 Monthly Status Report. FL: Florida Department of Corrections, Bureau of Research and Data Analysis.
- Frazier, Charles E. and Soon F.
1992 Reducing juvenile delinquency or expanding the official control nets: An evaluation of a reform effort. *Crime & Delinquency* 38:204-220.
- Gainey, Randy R. and Brian K.
2003 Changing attitudes and arrest with electronic monitoring: the impact of a single program. *Journal of Offender Therapy and Comparative Criminology* 17:209-229.
- Gainey, Randy R., Brian K. Pake O'Toole
2000 The relationship between jail, time on electronic monitoring, and recidivism: An analysis of a jail-based program. *Justice Quarterly* 17:733-75.
- Garland, David
2004 Personal correspondence.
- Gendreau, Paul, Claire Goggin-Cullen, and Donald A. Andrews
2000 The effects of community and incarceration on recidivism. *Forum on Corrections Research* 12:10-13.
- Gowen, Daren
2001 Remote location monitoring supervision strategy to enhance. *Federal Probation* 65:38-41.

- Hylton, John
1982 Rhetoric and reality: A critical appraisal of community corrections programmes. *Crime & Delinquency* 28:341-373.
- Johnson, Byron R., Linda Haugen, Jerry W. Maness, and Paul P. Ross
1989 Attitudes toward electronic monitoring of offenders: A study of probation officers and prosecutors. *Journal of Contemporary Criminal Justice* 5:153-164.
- Jolin, Annette and Brian Stipak
1992 Drug treatment and electronically monitored home confinement: An evaluation of a community-based sentencing option. *Crime & Delinquency* 38:158-170.
- Klein, Malcolm W.
1979 Deinstitutionalization and diversion of juvenile offenders: A litany of impediments. In Norval Morris and Michael Tonry (eds.), *Crime and Justice: An Annual Review of Research*. Chicago, Ill.: The University of Chicago Press.
- Lemert, Edwin M.
1981 Diversion in juvenile justice: What hath been wrought. *Journal of Research in Crime and Delinquency* 18:34-46.
1993 Vision of social control: Probation considered. *Crime & Delinquency* 39:447-461.
- Lianos, Michalis
2003 Social control after Foucault. *Surveillance and Society* 1:412-430.
- Lilly, J. Robert, Richard A. Ball, G. David Curry, and John McMullen
1993 Electronic monitoring of the drunk driver: A seven-year study of the home confinement alternative. *Crime & Delinquency* 39:462-484.
- Lucken, Karol
1997 Privatizing discretion: "Rehabilitating" treatment in community corrections. *Crime & Delinquency* 43:243-259.
- Mainprize, Stephen
1992 Electronic monitoring in corrections: Assessing cost effectiveness and the potential for widening the net of social control. *Canadian Journal of Criminology* 34:161-180.
- McMahon, Maeve
1990 Net-Widening: Vagaries of the use of a concept. *The British Journal of Criminology* 30:121-149.
- Morris, Norval and Michael Tonry
1990 *Between Prison and Probation: Intermediate Punishments in a Rational Sentencing System*. New York: Oxford University Press.
- Mortimer, E. and C. May
1997 *Electronic Monitoring in Practice: The Second Year of the Trials of Curfew Orders*. London: Home Office Research and Statistics Directorate.
- Ontario
1991 An Evaluation of the Electronic Monitoring Pilot Project: Mimico Correctional Centre, April 1989-October 1990. North Bay, Canada: Ministry of Correctional Services of Ontario.

- Payne, Brian K. and Randy R. Gainey
 1999 Attitudes toward electronic monitoring among monitored offenders and criminal justice students. *Journal of Offender Rehabilitation* 29:195-208.
 2000 Electronic monitoring: Philosophical, systemic, and political issues. *Journal of Offender Rehabilitation* 31:93-111.
- Petersilia, Joan and Susan Turner
 1990 Comparing intensive and regular supervision for high risk probationers. *Crime & Delinquency* 36:87-111.
- Renzema, Marc
 1991 The scope of electronic monitoring today. *Journal of Offender Monitoring* 4:6-11.
 1992 Home confinement programs: Development, implementation, and impact. In James M. Byrne, Arthur J. Lurigio, and Joan Petersilia (eds.), *Smart Sentencing: The Emergence of Intermediate Sanctions*. Newbury Park, Calif.: Sage.
 2003 *Electronic Monitoring's Impact on Reoffending*, Revised March 24, 2003. Retrieved August, 2005. Available online: <http://www.campbellcollaboration.org/doc-pdf/elecmon.pdf>.
- Renzema, Marc and Evan Mayo-Wilson
 2005 Can electronic monitoring reduce crime for moderate to high-risk offenders? *Journal of Experimental Criminology* 1:215-237.
- Roy, Sudipto
 1997 Five years of electronic monitoring of adults and juveniles in Lake County, Indiana: A comparative study on factors related to failure. *Journal of Crime and Justice* 20:141-160.
- Schmidt, Annesley K.
 1991 Electronic monitors: Realistically, what can be expected? *Federal Probation* 55:47-53.
- SPEC Associates
 2002 Final evaluation report: Michigan department of correction's GPS pilot phase II. Detroit Michigan.
- Stanz, Robert and Richard Tewksbury
 2000 Predictors of success and recidivism in a home incarceration program. *The Prison Journal* 80:326-344.
- Sugg, Darren, Louise Moore, and Philip Howard
 2000 *Electronic monitoring and offending behaviour—Reconviction results for the second year of trials of curfew orders*. London: Home Office Research, Development and Statistics Directorate.
- Taxman, Faye S. and Lori Elis
 1999 Expediting court dispositions: Quick results, uncertain outcomes. *Journal of Research in Crime and Delinquency* 36:30-55.
- Tonry, Michael and Michael J. Lynch
 1995 Intermediate sanctions. In Michael Tonry (ed.), *Crime and Justice: A Review of Research*, Vol. 20. Chicago, Ill.: University of Chicago Press.
- Ulmer, Jeffrey T.
 2001 Intermediate sanctions: A comparative analysis of the probability and severity of recidivism. *Sociological Inquiry* 71:164-193.

- Vaughn, Joseph P.
 1987 Planning for change: The use of electronic monitoring as a correctional alternative. In Belinda R. McCarthy (ed.), *Intermediate Punishments: Intensive Supervision, Home Confinement and Electronic Surveillance*. Issues in Criminal Justice, Vol. 2. London: Willow Press.
- Vollum, Scott and Chris Hale
 2002 Electronic monitoring: A research review. *Corrections Compendium* 27:1-4, 23-26, 27.
- Walters, R.
 1996 Alternatives to youth imprisonment: Evaluating the Victorian youth attendance order. *The Australian and New Zealand Journal of Criminology* 29:166.
- Whittington, Marie
 1987 Supervised Electronic Confinement Pilot Program October 1986-September 1987: Final Report. Santa Ana, Calif.: Orange County Probation Department.

Kathy G. Padgett is a research associate in the College of Criminology and Criminal Justice at Florida State University. Her research and publications have been in the areas of racial threat and social control, fear of crime and punitive attitudes, and the evaluation of the effectiveness and consequences of punishment and correctional strategies.

William D. Bales is an associate professor in the College of Criminology and Criminal Justice at Florida State University. His research and publications have been in the areas of evaluation of the effectiveness and consequences of punishment and correctional strategies, research methodologies, and criminal justice policy analysis.

Thomas G. Blomberg is Dean and Sheldon L. Messinger Professor of Criminology in the College of Criminology and Criminal Justice at Florida State University. Since 1998, he has served as Principal Investigator of the Juvenile Justice Educational Enhancement Program. His recent books include *Punishment and Social Control, 2d ed.* (2003), *Data Driven Juvenile Justice Education* (2001), and *American Penology* (2000).

Can electronic monitoring reduce crime for moderate to high-risk offenders?

MARC RENZEMA* and EVAN MAYO-WILSON

Criminal Justice Program, Kutztown University, Kutztown, PA 19530-0730, USA

**corresponding author.*

E-mail: renzema@kutztown.edu

Abstract. Electronic monitoring (EM) of offenders has been in use for just over two decades and motives for using it remain diverse. Some agencies that use EM attempt to deliver humane and affordable sanctions while others seek to relieve jail crowding or to avoid the construction of new jails. Nonetheless, all EM programs aim to suppress the criminal behavior of offenders being monitored and its advocates have always hoped EM could be instrumental in reducing long-term recidivism. This review investigates the history of EM and the extent to which EM empirically affects criminal behavior in moderate to high-risk populations. All available recidivism studies that included at least one comparison group between the first impact study in 1986 and 2002 were considered for the review. Although variants such as GPS tracking and continuous testing for alcohol in perspiration have recently emerged, no studies of these technologies were found that met the review's inclusion criteria. Studies are examined and combined for meta-analysis where appropriate. Given its continued and widespread use and the dearth of reliable information about its effects, the authors conclude that applications of EM as a tool for reducing crime are not supported by existing data. Properly controlled experiments would be required to draw stronger conclusions about the effects of EM.

Key words: electronic monitoring, evaluation, house arrest, house detention, meta-analysis, offender, parole, probation, recidivism, re-offending, review, systematic review, tagging

Introduction

Electronic monitoring (EM) is either in routine use or has been piloted on every inhabited continent. Overwhelmingly, prison overcrowding and the cost of building new prisons are cited as reasons for using EM. But today there are about 100,000 people in the United States being electronically monitored (Conway 2003: 5) and Europe is currently experiencing a wave of EM growth akin to that which swept the United States in the late 1980s (albeit with more attention to planning, quality of implementation, and attention to evaluation). By mid-2004, the number of offenders in Europe who had experienced EM exceeded 150,000. The daily caseload was just under 9,000, of whom 77% were in England and Wales.¹

For some people on EM, monitoring represents a true alternative to prison; without EM, some people who are monitored would be incarcerated. But children are on EM, people who refuse to pay child support are on EM, and so are tax cheats, drunk drivers, child molesters, and paroled killers.

While EM has been implemented in similar ways around the world, its use varies consistently between low-risk and high-risk offenders. In low-risk populations, EM may be used by itself or in conjunction with other forms of

low-contact monitoring. In moderate to high-risk populations, EM is more likely to be one part of a program that involves human contact and supervision, drug treatment, or other services. For these offenders, EM might be used as a true prison diversion program, thereby addressing overcrowding. But it is not known if EM is the best way to address the precipitating issues for this population. EM seems to be included as a solution to prison crowding largely because the public tolerates it.

This review is the first of two reviews of EM. In this work, the authors examine the impact of EM on recidivism for moderate to high-risk offenders. Most of the offenders in this review have been arrested several times, arrested at an early age and convicted of a serious offense. For these offenders, EM may serve its purported role as an alternative to incarceration. A second review will examine low-risk offenders whose crimes and characteristics differ from the offenders included in this review and for whom incarceration is not a likely sanction.²

As this and other reviews find, EM has not demonstrated superiority to options such as penal code reform, intensive probation, or psychotherapy in reducing the burden of imprisonment or in reducing recidivism among moderate to high-risk offenders.

Background

History of the intervention

In the 1960s, a research group at Harvard worked on the development of medical telemetry and tracking systems. As part of their experiments, a few volunteer offenders were electronically monitored and one of the investigators proposed that the equipment could be used as an adjunct to psychotherapy and to enhance accountability (Schwitzgebel 1967). Although there was discussion of the constitutional implications of such tracking during the 1970s, no new applications were attempted during the decade.

House arrest *without* electronically-aided enforcement, used since biblical times, underwent significant expansion in the late 1970s and early 1980s, largely consequent to institutional population pressures. Despite large-scale use, many agencies were uneasy about offender compliance with what was also known as “home detention” or “home confinement.”

Continuous signaling (CS) technology

By the early 1980s, three companies were experimenting with monitoring systems that consisted of ankle-worn radio transmitters and programmable receivers placed in offenders’ homes connected to hardwired telephone lines. Because the devices worn by offenders were constantly monitored, these were frequently called “continuous signaling” systems (CS). At defined intervals and whenever an unauthorized absence or other suspect event occurred, the receivers would automatically place calls to monitoring agencies. The agency could be either a public criminal justice agency or a private contractor that would relay violation results to the responsible public agency. Violating offenders could be taken into custody or

otherwise sanctioned. Although called “continuous signaling” technology, the devices usually monitored only presence/absence at a single location. Sporadic use was made of dual home/work monitoring units as well as “drive-by” units that could pick up the ankle transmitter’s signal at work, educational institutions, or treatment programs. Although offenders could cut the transmitters off from their ankles, various “tamper alert” systems assured that such violations were discovered. Over time, drive-by units were adopted by more and more agencies and used on a regular basis not only to check compliance at scheduled locations outside of the home, but also during sweeps of “hot” violation zones, such as bars and areas known for drug sales.

It is important to note that early systems frequently needed repair and generated abundant “false positives” of offender curfew violations. In many instances, it is impossible to know whether a “monitored” group actually received monitoring to the extent intended. It is also impossible to specify when technical improvements and increased agency competence resulted in acceptable program integrity. While later research is not exempt from technical problems or user competence problems, according to Peggy Conway, editor of *The Journal of Offender Monitoring*, by the late 1990s technical problems had become tertiary to cost and workload issues. All EM research, but particularly that done before 1990, should be examined for treatment delivery problems; the degree to which EM was used as it was meant to be used must be considered.

Random calling (RC) technology

Other machines were not in continuous contact with a device worn by an offender but, instead, used random calling (RC) to track offenders. To verify that the offender was answering the telephone, a variety of systems were used. Marketed first and most popular was an ankle-worn locked band that contained a magnetic key, which had to be mated with a wand connected to a telephone attachment. Identity verification systems included slow-scan picture phones, electronic voice analysis, and code emitting wristwatches. Remote breath-testing for alcohol was developed by the late 1980s and is a variant of RC technology.

Recent developments

In late 1997, two vendors began marketing systems that mated CS, wireless phone, and Global Positioning System (GPS) technologies. Although GPS tracking is limited by cellular network coverage and blockage of satellite coverage by structures, agencies were attracted by the ability to track offenders in real-time. As of 2004, GPS tracking appears to be gaining market share at the expense of CS systems. In 2001, a demonstration project began on a system that linked GPS tracking with police crime-mapping databases. If applied to large numbers of offenders, police could identify offenders in proximity to a reported crime or provide an “electronic alibi” for offenders who were not in the vicinity of the crime. GPS-based loggers that record offender movements but do not relay

movements in real-time to a monitoring agency have also been tried; data from these systems are typically uploaded daily through a modem.

Over the years, several types of home-installed RC systems that test for alcohol use have been introduced with mixed results. In 2003, CS equipment was introduced that can perform up to two tests per hour for alcohol emitted through the skin. Research is underway that may result in remote testing, with or without instant agency notification, for other drugs via traces found in sweat, characteristic eye movements, voice changes, or muscle tremors.

Although all of the emergent technologies have found marketplace acceptance, as of the cutoff for this review, none had been studied relative to reoffending using minimally acceptable methodologies. All of the studies reviewed in this report used either RC or CS monitoring.

Applications of electronic monitoring

In moderate to high-risk populations, EM is often intended as a diversion program; it is used in lieu of jail or prison to relieve overcrowding or to reduce the need for new prisons and jails. EM may also be used at the end of a prison sentence with the intent of helping prisoners transition into their communities. But other prison diversion programs exist; while the impacts of EM on reoffending might be compared to the impacts of prison, EM must also be compared to other programs.

No definitive reports of EM's effects on crime exist, yet it is extremely important to examine the effects of EM on crime for several reasons.

First, EM may have positive, negative or neutral effects on offending during its use. Compared to unsupervised release, EM might suppress crime during the monitored period, but when it is applied to offenders who would otherwise be incarcerated, EM might expose communities to risk during the period of monitoring.

Second, EM may have positive, negative or neutral effects on criminal behavior after its completion.³ Again, EM must be considered relative to other options. For example, compared to EM, prison might be relatively criminogenic while drug treatment might reduce recidivism.

Finally, because the use of EM varies by population and because the impact of EM in low-risk populations may differ from its impact in high-risk populations, it is critical that researchers examine the effects of EM in each group and, if it is to be used at all, determine how EM is most effectively used with particular populations of offenders.

While EM may reduce spending on prisons and jails and while it may affect criminal behavior, EM might be applied in other innovative ways. In moderate to high-risk populations, EM could be used to reduce the burden of monitoring on probation and parole officers. Although other monitoring would continue, some parts of routine monitoring could become "automated" through the use of EM. EM could also be used as an early warning system to distinguish offenders able to function in the community from offenders for whom reincarceration is needed. In such a system, breaches of EM protocol would result in the return of recidivists (or people expected to recidivate) to prison and the release of reformed offenders into the community. Some recidivism would be expected among the EM completers, but

one would expect their rate of recidivism would be lower than the rate of recidivism for EM dropouts or comparable offenders not subjected to a period of EM.

Prior review results

Corbett and Marx (1991), Mainprize (1996), MacKenzie (1997), Schmidt (1998), Gendreau et al. (2000), and Whitfield (2001) all have done careful reviews of the literature about EM's effects. MacKenzie focused on two studies using random assignment while Gendreau et al. did a meta-analysis of 140 studies that included six studies of EM and a total of 1,414 offenders. No positive effects on recidivism for EM were claimed by any of the reviewers. In fact, Gendreau et al. (2000) noted a 6% recidivism rate for EM studies as compared to 4% for the comparison group, a difference not statistically significant. Gendreau et al. did note a 10% recidivism reduction for studies that included a "modicum" of treatment in addition to the primary interventions of intensive supervision programs, arrest, fines, restitution, boot camps, scared straight, drug testing, and electronic monitoring. Unfortunately, they found insufficient information in the studies to address issues of treatment quality.

In addition to the review articles, several research reports contain excellent syntheses of prior work, notably works by Klein-Saffran (unpublished data), Bonta et al. (1999), Gainey et al. (2000), and Finn and Muirhead-Steves (2002).

None of the reviews that examined the methodology of the reviewed studies were able to substantiate any general effect on post-EM recidivism.

The authors of this review improve upon previous efforts in seven aspects:

1. Following the Campbell Collaboration approved protocol (Renzema 2003), the search strategy is both more clearly defined and intensive than most previous reviews. In particular, efforts have been made to obtain agency reports and other unpublished studies in order to minimize publication bias.
2. Inclusion/exclusion criteria are specified and transparent.
3. Program integrity issues are considered in inclusion/exclusion decisions.
4. The extension of the review period through 2002 allows consideration of large studies and studies that are methodologically superior to previously reviewed work.
5. Where possible, outcomes are assessed at both the termination of EM and during a longer follow-up period in recognition of the hypothesis that EM might suppress crime during its application but not in the long run.
6. The authors code the presence/absence of several treatment elements that may co-occur with EM.
7. Although the work resulting in this review is ongoing and includes the evaluation of all applications of EM, given the work summarized in Cullen and Gendreau (2000) on the futility of diffuse interventions with low-risk offenders, the authors focus their initial analysis on moderate to high-risk populations.

Objectives

Considering the number of EM programs around the world and the wide range of potential EM outcomes, it is urgent that we understand what actually happens

when an offender is given EM rather than another intervention. In this review, the authors examine the effect of EM on crime both for the duration of EM and after the discontinuation of monitoring in moderate to high-risk populations.

Criteria for including studies in this review

Types of interventions

For the purpose of this review, electronic monitoring was defined as any technology that “records the location of an offender within the community at particular places and times without human observation and transmits these data electronically to a central monitoring station, or uses an electronic device to detect the presence of a prohibited substance in the body (or to monitor other physiological functions) of an offender living in the community and transmits those data to a central location” (Renzema 2003). This definition excludes ignition interlocks but includes GPS tracking, logging, and emerging drug-testing technologies.

Types of offenders

This review investigates the effectiveness of EM for moderate to high-risk adult (18+) offenders.

Developing a criterion for “moderate to high-risk” proves a bit troublesome in the absence of standard risk assessment instrument scores for most of the samples examined here. Included as “moderate to high-risk” are probationers and others for whom recidivism measures exceed 30% during the study’s criterion period, typically one to three years. This is arbitrary given the variety of recidivism definitions, follow-up periods, offender mixes, and policy variations across jurisdictions.

Offenders at the “back end” of the criminal justice system, i.e., parolees, early releasees, and divertees who have served some institutional time are also defined here per se as “moderate to high-risk.”⁴

Comparison groups

To be considered, a study must have included one or more appropriate comparison groups receiving:

1. Traditional probation or parole
2. Intensive supervision probation or parole
3. Incarceration
4. An intervention other than parole or incarceration

Group assignment

To be considered, studies must have used one of the following methods of group assignment:

1. Random allocation—offenders in the EM group and control group are placed in groups without any attempt by researchers, judges, prosecutors, etc., to match them with offenders in another condition or to otherwise influence assignment.

2. Matching—offenders in the EM group are matched with a contemporary group of subjects that has the same risk of recidivism and is highly similar in most recent crime committed, criminal history, and demographic variables.
3. Historical matching—offenders in the EM group are compared to matched subjects from a comparable time period before EM was implemented in the area where the study takes place.

Outcome measures

To be considered for the review, a study must have included at least one primary outcome measure or one secondary outcome measure.

Primary outcomes:

1. Release condition violations resulting in reincarceration
2. Arrest for a new crime
3. Conviction of a new crime

Secondary outcomes:

1. Violations not resulting in a return to prison
2. Employment
3. Restitution
4. Substance abuse as measured by testing

Search strategy

The lead author attempted to obtain all research, published and unpublished, concerning the impact of EM on offender behavior. Electronic searches were conducted, reference lists and conference reports were examined, government agencies in the U.S., Canada, England, France, the Netherlands, and Sweden were contacted, equipment producers were surveyed, and leading researchers were asked for leads. No language restrictions were applied; studies were found in English, French, Dutch, Swedish, and German. For a more detailed description of the search, see Renzema (2003: 12–16).

In all, 381 articles or abstracts on EM were reviewed. Of these, 154 appeared to include evaluations. At this writing, one of the 154 remains fugitive but would probably *not* be included as the abstract makes no reference to a comparison group (Schafer and Martin 2001). The lead author designed a spreadsheet that includes the key characteristics of the 119 studies that were accurately classified as evaluations of EM.⁵ Of the 119 studies, 100 were immediately and clearly eliminated as not meeting the inclusion criteria. The remaining studies were independently examined by both authors with the inclusion/exclusion decisions reached jointly. Those selected for inclusion were independently coded; coding differences were reconciled in conference (Figure 1).

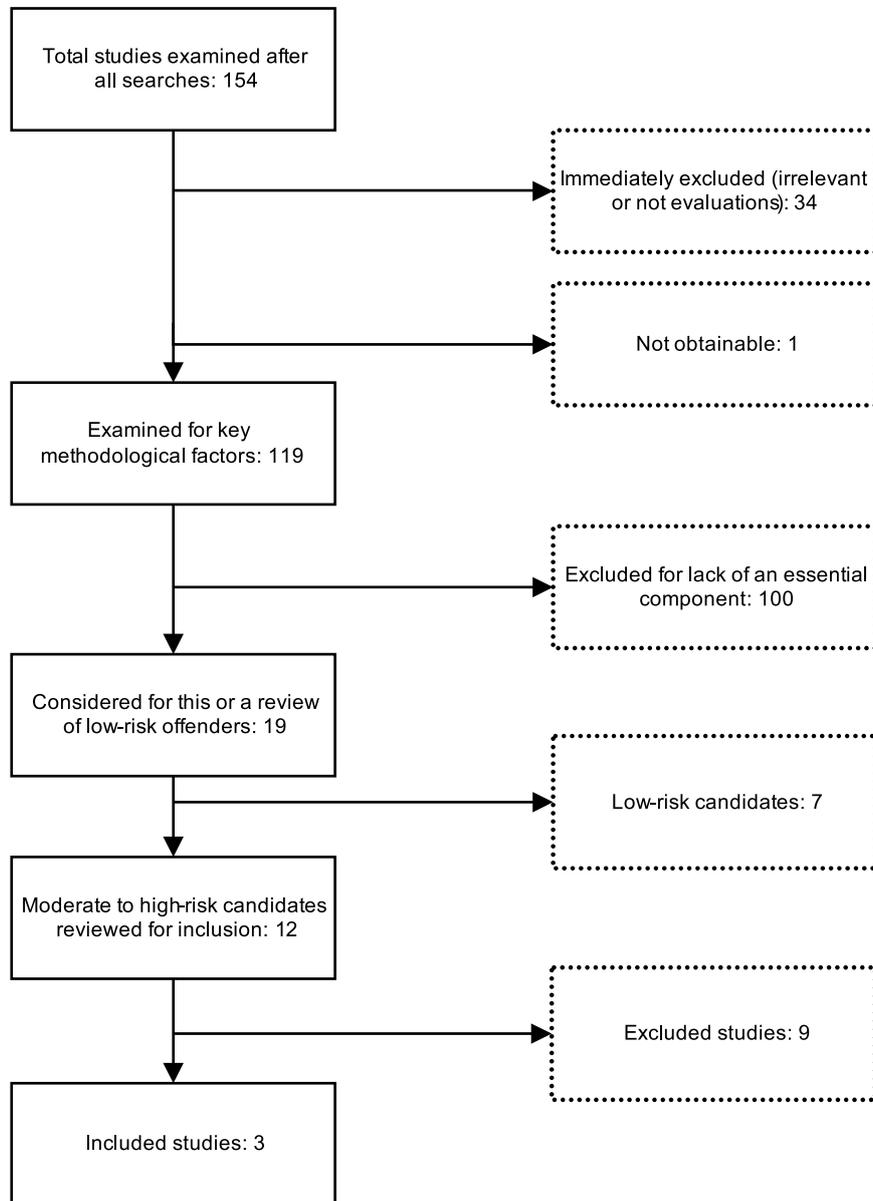


Figure 1. Study search flow chart.

Results

Excluded studies

The nine studies described below were considered for inclusion in this meta-analysis but were eventually excluded.

Petersilia and Turner (1990) conducted one of only four randomized trials discovered during the search. It was one of two randomized controlled trials (RCTs) that considered recidivism as an outcome variable. The authors attempted to study the marginal value of EM as an adjunct to probation; that is, they tried to understand the impact of EM on offenders already under intense human supervision.⁶ While many reports detail the cost and purported benefits of electronic monitoring, few studies examine its value as an instrument used in conjunction with other crime suppressing tools that can be used independently. Petersilia and Turner aimed to do exactly this in Los Angeles, but they encountered barriers to implementation that made their results impossible to interpret with any confidence. As a result of those barriers, only 44% of the offenders assigned to the EM group were ever monitored electronically. Additionally, offenders in EM and intense supervision groups were supposed to receive ten contacts per month during the follow-up period; most received four. Poorly implemented human surveillance resulted in a difference between the groups. The authors estimate that 40% of the offenders receiving EM received a medium (32%) or high (8%) level of face-to-face contact whereas only 28% of the intense supervision group received a medium (22%) or high (6%) level of face-to-face contact. These contacts, a slightly higher number of telephone and collateral contacts, and law enforcement checks of the EM group likely caused the resultant differences in recorded violations.

Though not included in the selection criteria above, one post-hoc hypothesis, that EM needs to be delivered to be effective, seemed reasonable. The study was excluded because treatment integrity was considered insufficient (both by Petersilia and Turner and by the reviewers)⁷ to support any conclusions about EM.

The present authors regret that neither Petersilia and Turner (1990) nor this review add to our immediate understanding of EM's costs and benefits as an adjunct to parole or probation. One can only conclude that high-quality research of the marginal value of EM (the value of EM as an addition to existing methods of supervision) is desperately needed.

Austin and Hardyman (unpublished data) studied the early release of prisoners in Oklahoma through the Pre-Parole Conditional Supervision Program. Between 1989 and 1991, EM was tested as an additional component of an established pre-release program.

Unlike other studies that assign offenders who can meet the conditions of monitoring to EM and assign those offenders who cannot meet these conditions to other groups, Austin and Hardyman (unpublished data) screened participants for their abilities to participate in EM (they were required to have a residence with a phone) and then randomly assigned only those subjects who met the inclusion criteria. Methodologically, this design is laudable and exceptionally rare, but some

of those who were randomized did not actually meet the criteria as anticipated. Offenders who were randomized to the EM group but could not receive EM were put into a "No Phone" group. The study would have benefited by similarly screening the control participants after randomization, but this problem was unforeseen. Still, data collection continued and the authors presented all of the data. Unfortunately, the experimental group may have been creamed despite an otherwise superb effort to obtain comparable groups. That is, those offenders in the experimental group may have been more likely to succeed than offenders in the comparison group.

Despite the potential bias, the study is of comparable or better quality than other studies included in this review. It is not included in the meta-analysis because, apparently as a result of chance, the follow-up periods differed greatly between the EM group and the control group. On average, offenders in the control group were followed for 105.4 days and offenders in the EM group were followed for 126.6 days (20% longer). Rates of recidivism in the EM group were higher than rates of recidivism in the control group, but the difference may be related to the follow-up period; the re-arrest rate for offenders receiving EM (13.9%) was 25% higher than the rate of re-arrest in the control group (11.2%). Accounting for the follow-up period, more technical violations were recorded in the EM group than in the control group, but this did not appear to be related to new crimes.

Austin and Hardyman conclude that EM ought not to be used with all offenders and that the ability of EM to assist in monitoring those parolees who are at highest risk of parole violation is worth examining. If one's goal is to detect technical violations, this appears to be correct. If one wants to reduce crime, this study shows no value for EM in addition to another form of monitoring.

Dodgson et al. (2001) considered EM as an early release program in the U.K. The authors examined a group of 118 prisoners released to home detention curfew (HDC), which was intended to ease the transition of prisoners into society and to reduce recidivism. During the period studied, an additional 558 prisoners were eligible for HDC based on statistical indicators but were denied release during a subjective evaluation.

The study was considered because the offenders studied had already served a custodial sentence and because the risk of recidivism for the group as a whole was moderate. But those offenders who were granted HDC had a much lower likelihood of recidivism than their peers who were denied HDC.

Recognizing that the released group had been creamed, Dodgson and her colleagues decided not to compare the results of the group granted HDC to another population. Instead, they combined the HDC group with the group not granted HDC, and compared the results to a historical control that would have been eligible for consideration had HDC been operating at the time. This resulted in a group of 676 offenders who had been granted (118) or denied HDC (558), and a historical comparison of 6,723 offenders.

Logically, if HDC had a strong effect on recidivism, the more recent group would have demonstrated a reduced rate of recidivism. The group granted HDC, however, represented the lowest risk group in the sample and, at six months, had a

rate of recidivism of only 9.3%. Hypothetically, if this represented a 50% decline in recidivism (11 recidivists rather than 22), the rate of recidivism in the larger group would have declined by only 1.6%. Even the most optimistic proponent of EM would expect a much smaller decrease in recidivism. Lowering the rate of recidivism in the treated group from 11.6% to 9.3%, or decreasing it by 20%, would represent a raw decrease from 14 recidivists to 11 recidivists. By diluting the treatment group, this very significant decline would become undetectable.

The study was excluded because there is no reliable way to determine the effect of EM on the treated group. While it would be inappropriate to include such a design in a meta-analysis, the design used by Dodgson et al. is not without merit. The design is logically defensible, but a much larger sample would be required to detect even a strong effect of EM on recidivism in this population during a short period.

One study (Florida 1987), which appeared superficially to be an RCT, did not appear, on closer examination, to be of high enough methodological quality to be included in the review.

Five studies (Jolin 1987; Jolin and Stipak 1992; Jones and Ross 1997; Klein-Saffran, unpublished data; Quinn and Holman 1991) were excluded because the reviewers concluded that the comparison groups were inadequately matched. The reviewers considered the potential for bias in judging the degree to which control groups matched experimental groups, but decided that potential biases in study selection were outweighed by biases in studies thought to be severely flawed. Reasons for exclusion are listed in Table 1.

In the studies excluded due to poorly matched comparison groups, some variables (e.g., age, number of prior convictions, risk scores) could be coded and considered through statistical analysis, but other variables that influenced group assignment could not be quantified and appeared to affect recidivism. For example, Klein-Saffran (unpublished data) considered two groups that differed on important variables, but were similar in most respects. However, offenders assigned to halfway houses could not find suitable accommodation on their own. Despite other statistical similarities, the reviewers believed that this difference would be impossible to control through any amount of statistical adjustment.

The reviewers note that all matched studies in this field are likely to include groups that are different in some way. Even well-matched historical controls may differ on one important variable. Still, only those studies that met the relatively strict inclusion criteria outlined above were included because the reviewers felt that only studies with the specified characteristics would provide real evidence of the effect of EM on recidivism.⁸ While some might argue that other studies should have been included in this review and meta-analysis despite the objections outlined here, the reviewers doubt that biases or errors in judgment influenced the final result. It is noteworthy that among the studies excluded for poorly matched control groups, results exist both in favor of EM and in favor of the comparison group. It is also worth noting that the outcomes of these studies are consistently in the direction one would predict at baseline given the characteristics of the groups.

Table 1. Excluded studies.

<i>Study</i>	<i>Summary of results</i>	<i>Reason for exclusion</i>
Austin and Hardyman (1991)	Accounting for differences in length of monitoring, EM showed no effect as an adjunct to another form of monitoring. EM did, however, appear to increase the detection of parole violations.	As a result of chance, follow-up periods between groups were too different to allow reasonable comparisons. While the study was well designed and included a safeguard against "creaming" that was not present in any other study examined, the experimental group may have been "creamed."
Dodgson et al. (2001)	After six months, those selected for release to EM had a low rate of recidivism (9.3%) compared to eligible offenders not granted release (40.5%) and an historical comparison group (30.0%).	Offenders placed on EM were carefully selected based on statistical and subjective analyses. Consequently, an appropriate comparison group could neither be found nor formed post hoc for inclusion in meta-analysis.
Florida (1987)	Apparent prison divertees on two types of monitoring were compared to a no-EM condition. EM paid more restitution.	The study was initially misclassified as an RCT. It says that assignment was "generally random," but no details concerning group assignment are offered.
Jolin (1987)	The EM group, a subset of current work releasees, was matched with past work releasees. The EM group had a lower rate of re-arrest than the comparison group.	The sentence length and follow-up periods differed between groups (6–18 <i>versus</i> 18–36 months) and the EM group may have been "creamed."
Jolin and Stipak (1992)	EM was compared to work release and a drug treatment program. Rates of re-arrest were compared. Even controlling for those differences that made inclusion in meta-analysis inappropriate, no reliable conclusions are possible due to baseline differences.	The EM and work release groups differed greatly in convictions for (1) felonies and (2) drug-related offenses. Comparisons to the drug treatment program were impossible for those and other reasons, notably a significant difference in age, another known predictor of recidivism.
Jones and Ross (1997)	The rate of fingerprinted re-arrest within two years after assignment to EM was compared to a group assigned to boot camp. While the EM group showed a much higher rate of recidivism than the boot camp group, this difference reflects baseline differences in risk. Evidence presented here does not support any conclusions concerning the relative merits of the programs.	The groups differed in previous and current convictions for (1) violent or sexual felonies, (2) violent misdemeanors, and (3) property offenses. EM participants were at risk of recidivism while offenders in boot camps were confined. Data were not available for failure on EM and failure after its completion. Approximately 40% of subjects were under 16 years old and nearly all (98%) were less than 23 years old.

Table 1. Continued.

<i>Study</i>	<i>Summary of results</i>	<i>Reason for exclusion</i>
Klein-Saffran (unpublished data)	Offenders were assigned halfway houses or EM as part of Southern Florida's Community Control Project. Within one year of release, those assigned to EM were less likely than offenders placed in halfway houses to be arrested or revoked. This difference likely resulted from baseline differences.	Offenders placed in halfway houses were refused EM by community corrections managers or parole officers, most often because they did not have suitable accommodation or because they "had a need for halfway house services." Offenders placed in halfway houses were convicted of their first offenses four years before the EM offenders, were thrice as likely to be black, and had a higher risk of recidivism.
Petersilia and Turner (1990)	Offenders assigned to EM were more likely than intensely supervised offenders to be jailed or arrested during their probation. Given the failure to actually implement EM, this difference probably resulted from higher levels of contact with probation officers in the EM group. No differences in recidivism appeared after one year.	This was a well-designed trial with random assignment. Implementation was so poor, however, that the reviewers judged that it had to be excluded. Of 52 subjects assigned to EM, only 23 (44%) were ever monitored. Furthermore, there was poor and highly dissimilar implementation of intense supervision probation (ISP), which was intended to be a common feature of the EM group and the comparison group.
Quinn and Holman (1991)	Violating probationers and parolees placed on EM were matched with demographically similar non-violating probationers and parolees. Violators failed twice as often.	In this study, the comparison group was "creamed." Consequently, offense and occupational status differences between the two groups were both logically and statistically significant.

Included studies

Only three studies of moderate to high-risk offenders met the inclusion criteria for the review. All three studies had unique methodologies. While comparisons are informative, the studies merit individual examination; the authors urge caution in interpreting the combined results, except insofar as one may conclude that there are virtually no data supporting the use of EM.

Finn and Muirhead-Steves (2002)

Of the included studies, only Finn and Muirhead-Steves (2002) reported outcomes at multiple times. They compared EM to an historical control for high-risk, violent male offenders in Georgia. As demonstrated in Table 2, Finn and Muirhead-Steves (2002) suggest that EM has a modest impact for its duration, but its effect is

Table 2. EM outcomes over three time periods among male parolees with violence history.

<i>Outcome</i>	<i>Proportions recommitted</i>	<i>Percent C recommittal >E (%)</i>	<i>Odds ratio (fixed)</i>	<i>Lower limit</i>	<i>Upper limit</i>	<i>Z</i>	<i>p Value</i>
Recommitted within 150 days	E: 0/128; C: 4/158	2.53	0.134	0.007	2.505	-1.346	0.178
Recommitted within one year	E: 4/128; C: 15/125	6.93	0.308	0.099	0.951	-2.047	0.041
Recommitted within three years	E: 30/128; C: 37/158	-0.02	1.001	0.577	1.736	0.004	0.997

Limits are for 95% CI; mean duration of EM was 87.4 days with a range of 6–153 days.
Source: Finn and Muirhead-Steves (2002: 303–304).

transient; after EM ends, monitored offenders “catch up” to those who did not experience it. Within three years of release, 23.4% of the EM group ($n = 128$) and 23.4% of an historical comparison group ($n = 158$) were returned to prison.⁹

For one subgroup in Georgia, sex offenders, EM may have reduced recidivism; however, there is another plausible explanation for the observed effect and the study design precludes any definitive conclusions about the unique effects of EM.

Although program details are sketchy, during the study period Georgia was beginning implementation of the “containment model” of sex offender management, an empirically-based highly intensive treatment and surveillance approach described by English et al. (1996). In an e-mail to the first author, John Prevost, Associate Director of Research and Technology at the Georgia Board of Pardons and Parole, described sex offender treatment during the study period as “scattered local initiatives in selected parts of the state.”¹⁰ He noted that the addition of a planned minimum of 90 days of EM was a new element in the treatment package but that treatment was not universal for sex offenders in either time period. Most of the offenders in the control group probably did not receive polygraph exams; a few of the offenders in the EM group may have. Prevost also indicated that there were early concerns about the quality and availability of contracted psychotherapeutic services and contracted polygraph examiners.

In other words, later released (the EM group) sex offenders may have received more extensive and more competent overall treatment than the historical controls released during the previous year. It is reasonably clear from the agency’s 1998 annual report¹¹ that by the end of the study period there was a high level of program integrity, but there may be some historical bias that would tend to reduce later-released sex offenders’ recidivism with or without EM.

The reviewers also caution readers who accept the hypothesis that either EM or the improved treatment of sex offenders reduced recidivism in that group. For this to be true, one must also accept that to result in the identical overall results that were

Table 3. Return to prison with three years for Georgia sex offenders.

Outcome	Treatment			
	EM + other		Comparison (other without EM)	
	Percentage	n	Percentage	n
Not returned	94.3	33	70.4	31
Returned	5.7	2	29.6	13

$p=0.0088$ (Fisher's exact test), $C=0.29$.

Source: Finn and Muirhead-Steves (2002), additional data supplied by Finn.

observed, EM may have *increased* the rate of recidivism among the remaining offenders.

Of sex offenders in the EM group, two of 35 were returned to prison; 13 of 44 sex offenders in the comparison group were returned to prison. As shown in Table 3, this percentage difference is statistically significant ($p=0.0088$) using Fisher's exact test.

Bonta, Wallace-Capretta and Rooney (2000)

Bonta and his colleagues found that EM combined with court orders tended to improve compliance with a treatment program (Bonta et al. 2000b). They also found that the combination was associated with significantly lower recidivism for a group of moderate to high-risk prison divertees, but the same program failed to produce results for lower risk offenders.

Table 4, from Bonta et al. (2000b), illustrates the strength of the relationship but includes both EM + treatment prison divertees (54) and treated probationers (17) and compares them with an untreated matched group of prisoners.¹²

In Newfoundland, Bonta et al. examined EM in conjunction with a treatment program, which was also offered to control subjects. Members of the experimental group averaged 71.4 days of EM and were required to attend a cognitive behavioral

Table 4. Bonta et al.'s recidivism as a function of risk level and treatment.

Risk level	Treatment			
	Yes (IRS) ^a		No (Prison)	
	Percentage	n	Percentage	n
Low	32.3	10	14.5	8
High	31.6	12	51.1	23

^aIncludes participants in "LDP," a CBT program of whom 54 were divertees with EM and 17 were probationers without EM.

Source: Bonta et al. (2000b: 324).

Table 5. Longest term outcome of EM discussion and policy implications.

Study or sub-category	EM n/N	Comp n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
Bonta 1999	17/54	6/17		7.46	0.84 [0.27, 2.66]
Finn 2002	30/128	37/158		30.24	1.00 [0.58, 1.74]
Sugg 2001	190/261	192/261		62.30	0.96 [0.65, 1.42]
Total (95% CI)	443	436		100.00	0.96 [0.71, 1.31]
Total events: 237 (EM), 235 (Comp)					
Test for heterogeneity: $\text{Chi}^2 = 0.07$, $\text{df} = 2$ ($P = 0.97$), $I^2 = 0\%$					
Test for overall effect: $Z = 0.23$ ($P = 0.82$)					

0.1 0.2 0.5 1 2 5 10
Favours treatment Favours control

therapy program for nine hours per week. Control subjects, probationers without EM, were not subject to revocation for failure to attend the program. Only 52.9% of the unmonitored probationers completed the therapeutic program. Of the divertees who were required to attend and also on EM, 87% completed the program.

Although suggestive, it is impossible to gauge whether higher completion rates were due to EM or due to the threat of revocation. Given these data, it is impossible to determine whether EM had an independent contribution to the lowered recidivism of the higher risk offenders. Still, whether or not it occurred in this case, this study does demonstrate one application of EM as a means of increasing participation rates in other programs.

Sugg, Moore and Howard (2001)

Evaluating an emerging EM program in Manchester, Reading and Norfolk, Sugg et al. examined EM compared to combination and community service orders because “previous research has shown that, had curfew orders not been available, offenders would have received community sentences seen by sentencers as an alternative to custody.” The report published by the Home Office offers few specific details about the program. Within two years of being “curfewed” and subjected to EM, 72.8% of the offenders in the study had been reconvicted.

Combined results

Given the results of the individual studies, it should not be surprising that the combined results are equally grim. As Table 5 shows, there was no overall impact on recidivism at the longest follow-up period for each study, periods which ranged from one to three years.

Discussion and policy implications

After 20 years, it is clear that EM has been almost desperately applied without adequate vision, planning, program integration, staff training, and concurrent research. It has punished, perhaps more humanely and cheaply than otherwise possible, and it has been an element in the avoidance of prison crowding and prison construction,¹³ but it is not free and it is not without unintended effects.

Is EM simply another fad, another example of what Latessa et al. (2002) call “correctional quackery?” If one looks at gross recidivism rates for moderate to high-risk offenders, it would seem so. Through this review, the authors failed to identify any methodologically sound evaluation comparing EM to incarceration and they failed to find any convincing evidence that EM is superior to other prison diversion programs.

Yet there *may* be a depression of the rate of offending during the monitored period. Could some of the lessons in “relapse prevention” learned by those who treat substance abuse be applied here? Would an extension of the monitoring period for some offenders so that they “age out” be useful? The programs to test these ideas have not been evaluated and, for the most part, evaluations are not being done.¹⁴

The authors of this review found only two studies in which EM effects were plausible, but in both cases, effects were only observed in small subpopulations of offenders: Georgia sex offense parolees and Newfoundland prison divertees (Finn and Muirhead-Steves 2002; Bonta et al. 2000a). The reviewers considered the possibility that programs for these subgroups might be what Sherman and Strang (2004) call “light bulbs” and that the reviewers should “look for outliers rather than averages.”¹⁵ But in both cases, the reviewers found evidence that EM may not have caused the observed differences. The reviewers caution readers to consider other causal variables and to remember that systematic reviews may identify statistically different subgroups that differ only as a result of chance (Counsell et al. 1994).

Using the results of the long line of treatment impact studies that began in 1979 with Gendreau and Ross’s *Effective Correctional Treatment: Bibliotherapy for Cynics*, general principles of “what works” have been distilled, refined, and published repeatedly. The two EM programs in which effects were noted had several of the “what works” characteristics listed by Latessa et al. (2002). Paraphrasing Latessa, there appeared to be appropriate organizational culture, research-based programs, and client risk and needs assessments. Both programs had several components that addressed offender needs or traits directly related to criminal behavior and had a cognitive behavioral component.¹⁶ By contrast, EM did not appear to reduce recidivism among the remainder of the Georgia parolees or the offenders in the study by Sugg et al. (2001), who received minimal non-EM supervision and services.

One can only speculate why the two programs in which EM was coupled with another treatment did better than the relevant comparison groups. In Newfoundland, it is conceivable that EM was a useless addition to an effective treatment package that would have produced an impact even if it had not been included and that the divertees had a relatively high program completion rate because of the threat of return to prison. For the Georgia parolees, the chronology of treatment implementation is hazy; perhaps the non-EM sex offenders paroled in 1995 simply encountered a less-effective, less-organized treatment package than those who were released in 1996. There is some evidence that, even in those studies where it appeared to have some impact, EM was not the variable responsible for change.

Given the theoretical rationales for EM enumerated elsewhere (Renzema 2003: 6–8) and the meta-analytic studies of “what works” in corrections of Bonta, Cullen, Gendreau, Latessa, Ross, Sherman, and others over the past two decades, it is hardly surprising that recidivism has not been reliably reduced by an intervention that is typically quite short, applied in a standard fashion, and applied to a diverse group of offenders for whom it may or may not have any relevance to their motives for offending.

Practical advice for politicians and policy makers

What should policy makers do given the paucity of good information about the impact of EM? The reviewers have a few suggestions:

- Consider other options. If governments continue to use EM as they have for the past 20 years, EM will not reduce demands on parole officers nor will EM make our communities safer. Although fewer prisons may be built and filled because of EM’s use, EM is not the only prison diversion program. Other paths may be more effective in lowering costs and securing public safety.
- Treat underlying problems. Odds of success improve when EM is used as part of an evidence-based correctional package. Although EM may suppress crime for its duration, EM is *not* a “treatment” that directly changes values or teaches skills. Used in isolation, EM should not be expected to produce enduring effects for moderate to high-risk offenders. If EM is going to be used to address a budget crisis, to relieve prison crowding, or to increase offender accountability, EM should be coupled with programs that are likely to reduce recidivism.
- Use EM logically to accomplish realistic goals. Rather than as a knee-jerk reaction to crime, overcrowding, and high costs of running correctional systems, EM ought to be used in a sensible manner to accomplish clearly defined and realistic objectives. One might use EM to facilitate evidence gathering or to quickly return high-risk offenders to custody with the hope of minimizing risk to communities. One might use GPS technology to disrupt criminogenic associations. One might use EM in lieu of contact with parole officers. But one must use EM in a manner that is logically related to the objective at hand. EM will not necessarily lead to any desirable outcomes. Though the reviewers are uncertain of EM’s full effects, they are certain that it is not a panacea.
- Do not make it impossible for offenders to “succeed.” Technical violations and failure to pay the fees associated with EM and probation can, in some cases, result in incarceration. The costs and benefits of incarcerating people for such offenses should be weighed. Policy makers should also consider what offenders are meant to do while on EM. How will they spend their time? With whom will they interact? EM is necessarily a part of a larger program that should encourage lawful behavior and create opportunities for reform.
- Study the effects of EM. Little evidence about the impact of EM is available and, if governments continue to use it, they have an obligation to show that it creates public value. Even when one cannot randomly allocate offenders to EM or another program, one should find records from a group of similar offenders

and one should either invite outside research (preferred) or undertake research within the agency that oversees the EM program.

Conclusion

All studies of EM in moderate to high-risk populations have serious limitations and matched studies of EM in moderate to high-risk populations are of very low quality. After 20 years of EM, we have only a few clues as to its impact—we should know more by now. Government-approved experimental research may be the only way to determine if EM achieves its goals.

If EM continues to be used as it has been used, shortsighted governments will continue to waste taxpayer dollars for ideological reasons and political gain. Governments that choose to use EM in the future ought to use it to enhance other services that have a known effect on crime reduction. Those governments *must* test the marginal effects of EM, publish the results, and discontinue use of EM if it fails to provide quantifiable public benefits. Money spent on EM could be spent on empirically-tested programs that demonstrably protect our communities.

Acknowledgements

The work of Marc Renzema on this analysis was facilitated by a grant of reassigned time by Kutztown University and by financial support from the Campbell Collaboration. Evan Mayo-Wilson's work on this project was supported by Grant No. 2004-DD-BX-0003 awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the US Department of Justice.

Notes

- 1 E-mail to Marc Renzema on 21 August 2004 from Dick Whitfield, EM consultant in England.
- 2 The reviewers also felt that a single review of EM could result in misleading statistical analyses because a small effect on a rare outcome would be very difficult to detect. That is, the failure to detect a decrease in recidivism in a sample of offenders of whom 50% are expected to recidivate could provide evidence that EM does not work; failure to detect an effect in an equally sized sample of offenders with a 5% rate of recidivism might say very little about the true effects of EM.
- 3 In a Campbell Collaboration protocol for a review of EM, Renzema (2003) surveys several criminological and psychological theories and finds some support for expecting

- crime suppression during the monitoring period and less for expecting post-monitoring crime-free behavior. He found less theoretical support for the hypothesis that EM would increase recidivism.
- 4 The reviewers are aware of jurisdictions where fewer than a fifth of probationers “fail” and places where more than eight in ten have at least one violation of probation rules during their terms. A federal study of the outcomes of 1994 parolees in 15 states (Bureau of Justice Statistics 2002) showed parolee re-arrest rates at 6, 12, and 24 months from release of 29.9%, 44.1%, and 49.1%. The same study showed reconviction rates of 10.6%, 21.5%, and 36.4% in the same time periods. Thus, setting a mean failure rate minimum of 30% for inclusion as “moderate to high” risk accomplishes the primary goal here, which is to segregate the lowest risks for a separate analysis.
 - 5 This updated version of this spreadsheet is available at <http://www.renzema.net/META-DOCS/C2REVIEWCANDIDATES.pdf>.
 - 6 It seems logical that increased supervision of offenders should aid monitoring of drug use, criminal activity, and probation violations; a finding that closely supervised offenders on EM are more likely than virtually unsupervised offenders to be recalled during the period of monitoring would shed little light on the true effects of EM.
 - 7 Petersilia and Turner are clear about their methods and are transparent throughout their statistical analysis. They give an honest assessment of the data and suggest ways to improve future research in the field. While it had to be excluded from this review, this study provides valuable insights for anyone interested in doing experimental research of EM.
 - 8 Compared to criteria used for meta-analyses in medicine and psychology, these criteria are not strict at all. However, because experimental research is rarely done in social sciences other than psychology, previous meta-analyses have included large numbers of studies and sought to handle low-quality studies through sophisticated statistical analyses. The present authors suspected from the outset that a medical approach would produce an empty review (it would have) and that more open inclusion criteria would have generated more heat than light. The authors hoped that the criteria employed would return studies with some value while excluding those with more substantial sources of bias.
 - 9 Return to prison is one way of estimating reoffending, but the reviewers note that it probably underestimates the actual number of offenses committed.
 - 10 E-mail to Renzema, 24 August 2004.
 - 11 See http://www.pap.state.ga.us/results_driven_supervision.html.
 - 12 In our recidivism analysis (see Table 5), we did not use the prison group, which was not comparable to the experimental group. Instead, we considered the possibility that EM has a marginal impact on a reasonably intensive treatment program.
 - 13 Many studies (mostly outside the universe considered for this review) suggest that prison costs may be reduced and construction costs may be avoided because jurisdictions are able to divert offenders to EM in lieu of incarceration. To the extent that diversions have been possible only because the public will tolerate diversions to EM more than they will tolerate other prison diversion programs, these studies make sense. But most such studies fail to consider EM as one of many diversion programs, some of which may be cheaper, less intrusive, and/or of proven utility in reducing recidivism. Further, many analyses fail to consider the costs and benefits of EM as an addition to other forms of monitoring.
 - 14 Several states in the U.S. authorize lifetime probation or indeterminate civil commitment post prison for certain offenders, primarily those who have committed sex crimes. Under these statutes, EM could be used for a long time, but there is no

- evidence that it is being used for periods beyond six months except in the most unusual and extreme cases. Renzema and Skelton (1990) found an average duration of 80 days nationwide and the work reviewed here does not suggest that this has changed much.
- 15 Sherman and Strang suggest that, in some cases, social scientists should think like inventors who embrace outliers and try to replicate them. "Thomas Edison was not interested in the average life of all previous versions of the lightbulb..." (Sherman and Strang 2004). The reviewers find this idea compelling, but as a tool for reducing recidivism, the reviewers believe that EM remains unproven and not very promising.
 - 16 Information about the type (and availability) of psychotherapy received by Georgia sex offenders at the beginning of the study period is incomplete; however by its end it was based on cognitive-behavioral principles.

References

- Bonta, J., Wallace-Capretta, S. & Rooney, J. (1999). *Electronic monitoring in Canada*. Ottawa: Solicitor General.
- Bonta, J., Wallace-Capretta, S. & Rooney, J. (2000a). Can electronic monitoring make a difference? An evaluation of three Canadian programs. *Crime and Delinquency* 46(2), 61–75.
- Bonta, J., Wallace-Capretta, S. & Rooney, J. (2000b). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior* 27(3), 312–329.
- Conway, P. (2003). Survey of agencies using electronic monitoring reveals a promising future. *Journal of Offender Monitoring* 16(2), 5, 18–23.
- Corbett, R. & Marx, G. T. (1991). Critique: No soul in the new machine: Techno fallacies in the electronic monitoring movement. *Justice Quarterly* 8(3), 399–414.
- Counsell, C., Clarke M., Slattery, J. & Sandercock, P. (1994). The miracle of DICE therapy for acute stroke: Fact or fictional product of subgroup analysis? *British Medical Journal* 309, 1677–1681.
- Cullen, F. T. & Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In J. Horney (Ed.), *Policies, processes, and decisions of the criminal justice system: Criminal justice 2000*, vol. 3 (pp. 109–175). Washington, DC: U.S. National Institute of Justice.
- Dodgson, K., Goodwin, P., Howard, P., Llewellyn-Thomas, S., Mortimer, E., Russell, N., et al. (2001). *Electronic monitoring of released prisoners: An evaluation of the home detention curfew scheme (Home Office Research Study No. 222)*. London: Home Office Research Development and Statistics Directorate.
- English, K., Pullen, S. & Jones, L. (1996). *Managing sex offenders: A containment approach*. Lexington, KY: American Probation and Parole Association.
- Finn, M. A. & Muirhead-Steves, S. (2002). The effectiveness of electronic monitoring with violent male parolees. *Justice Quarterly* 19(2), 293–312.
- Florida Department of Corrections, Probation and Parole Services. (1987). *Verifier wristlet project evaluation report (agency report)*. Miami–Fort Lauderdale: Florida Department of Corrections.
- Gainey, R. R., Payne, B. K. & O’Toole, M. (2000). The relationships between time in jail, time on electronic monitoring, and recidivism: An event history analysis of a jail based program. *Justice Quarterly* 17(4), 733–752.

- Gendreau, P. & Ross, R. (1979). Effective correctional treatment: Bibliotherapy for cynics. *Crime and Delinquency* 25(4), 463–489.
- Gendreau, P. L., Goggin, C., Cullen, F. T. & Andrews, D. A. (2000). The effects of community sanctions and incarceration on recidivism. *Forum on Corrections Research* 12(2), 10–13.
- Jolin, A. (1987). *Electronic surveillance program: Clackamas County Community Corrections Oregon Evaluation (agency report)*. Oregon City, OR: Clackamas County Community Corrections.
- Jolin, A. & Stipak, B. (1992). Drug treatment and electronically monitored home confinement: An evaluation of a community-based sentencing option. *Crime and Delinquency* 38(also cited as 39)(2), 158–170.
- Jones, M. & Ross, D. L. (1997). Electronic house arrest and boot camp in North Carolina: Comparing recidivism. *Criminal Justice Policy Review* 8(4), 383–404.
- Latessa, E. T., Cullen, F. & Gendreau, P. (2002). Beyond correctional quackery—professionalism and the possibility of effective treatment. *Federal Probation* 66(2), 43–49.
- MacKenzie, D. L. (1997). Chapter 9: Criminal justice and crime prevention. In L. W. Sherman, D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter & S. Busway (Eds.), *Preventing crime: What works, what doesn't, what's promising: A report to the United States Congress* (pp. 9-1–9-83). Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Mainprize, S. (1996). Elective affinities in the engineering of social control: The evolution of electronic monitoring. *Electronic Journal of Sociology* 2(2), 26.
- Petersilia, J. & Turner, S. (1990). Comparing intensive and regular supervision for high-risk probationers: Early results from an experiment in California. *Crime and Delinquency* 36(1), 87–111.
- Quinn, J. F. & Holman, J. E. (1991). The efficacy of electronically monitored home confinement as a case management device. *Journal of Contemporary Criminal Justice* 7(2), 128–134.
- Renzema, M. (2003). *Electronic Monitoring's Impact on Reoffending, Revised March 24, 2003*. Retrieved 20 August, 2004, from <http://www.campbellcollaboration.org/doc-pdf/elecmon.pdf>.
- Renzema, M. & Skelton, D. T. (1990). *The use of electronic monitoring by criminal justice agencies 1989: A description of extent, offender characteristics, programmatic issues, and legal aspects*. Bloomington, IN: National Institute of Justice.
- Schafer, N. E. & Martin, P. (2001). *Evaluation of a JAIBG-Funded Project: Voice and location telephone monitoring of juveniles (grant report for BJS grant 1999-JR-VX-005)*. Anchorage, AK: University of Alaska Anchorage.
- Schmidt, A. K. (1998). Electronic monitoring: What does the literature tell us? *Federal Probation* 62(2), 10–15.
- Schwitzgebel, R. (1967). Electronic innovation in the behavioral sciences: A call to responsibility. *American Psychologist* 22(5), 364–370.
- Sherman, L. & Strang, H. (2004). Verdicts or inventions?: Interpreting results from randomized controlled experiments in criminology. *American Behavioral Scientist* 47, 575–607.
- Sugg, D., Moore, L. & Howard, P. (2001). *Electronic monitoring and offending behavior—reconviction results for the second year of trials of curfew orders (No. Research Findings No. 141)*. London: Home Office Research, Development and Statistics Directorate.

- U.S. Department of Justice, Bureau of Justice Statistics. (2002). *Recidivism of Prisoners Released in 1994*. Washington, DC.
- Whitfield, D. (2001). *The magic bracelet*. Winchester, UK: Waterside Press.

About the authors

Marc Renzema is a Professor of Criminal Justice at Kutztown University. His work on this article was facilitated by a grant of reassigned time by Kutztown University and by financial support from the Campbell Collaboration.

Evan Mayo-Wilson completed his work on this project while a graduate student at the University of Pennsylvania, Fels Institute of Government. He currently works in the Department of Social Policy and Social Work at the University of Oxford.

A QUANTITATIVE REVIEW OF STRUCTURED, GROUP-ORIENTED, COGNITIVE-BEHAVIORAL PROGRAMS FOR OFFENDERS

DAVID B. WILSON

George Mason University

LEANA ALLEN BOUFFARD

North Dakota State University

DORIS L. MACKENZIE

University of Maryland

Prior reviews and meta-analyses have supported the hypothesis that offender rehabilitation programs based on cognitive-behavioral principles reduce recidivism. This article quantitatively synthesizes the extant empirical evidence on the effectiveness of structured cognitive-behavioral programs delivered to groups of offenders. The evidence summarized supports the claim that these treatments are effective at reducing criminal behavior among convicted offenders. All higher quality studies reported positive effects favoring the cognitive-behavioral treatment program. Specifically, positive reductions in recidivism were observed for moral reconnection therapy, reasoning and rehabilitation, and various cognitive-restructuring programs. The evidence suggests the effectiveness of cognitive skills and cognitive restructuring approaches as well as programs that emphasize moral teachings and reasoning.

Keywords: cognitive-behavioral; recidivism; meta-analysis; group therapy

The debate surrounding the effectiveness of rehabilitation efforts for criminal offenders has moved from the rather pessimistic perspective of the 1970s and 1980s, exemplified best by Martinson

AUTHOR NOTE: *Correspondence concerning this article should be addressed to David B. Wilson, Administration of Justice, George Mason University, 10900 University Blvd, MS 4F4, Manassas, VA 20110-2203. E-mail: dwilsonb@gmu.edu. This work was supported in part by the Jerry Lee Foundation.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 32 No. 2, April 2005 172-204

DOI: 10.1177/0093854804272889

© 2005 American Association for Correctional Psychology

(1974), to a more optimistic perspective driven by research from the 1980s and 1990s. The effectiveness of some rehabilitation approaches has renewed such optimism (e.g., Andrews et al., 1990; Lipsey, 1992; Lipsey & Wilson, 1998; MacKenzie, 2002; Whitehead & Lab, 1989). A consistent theme in numerous reviews of the rehabilitation literature is the positive effects of cognitive and cognitive-behavioral approaches for treating the offender population (e.g., Cullen & Gendreau, 1989; Gendreau & Ross, 1987; Husband & Platt, 1993). For example, Andrews et al. (1990) concluded from a meta-analysis of adult and juvenile correctional treatment that cognitive and behavioral methods are critical aspects of effective correctional treatment (see also Losel, 1995). Similarly, Gendreau and Andrews (1990) concluded that the most effective interventions are those that use cognitive-behavioral techniques to improve cognitive functioning. Research reviews of cognitive-behavioral programs for offenders have also drawn favorable conclusions (Allen, MacKenzie, & Hickman, 2001; MacKenzie & Hickman, 1998). This is not entirely a surprise because cognitive-behavioral treatments have become a dominant, if not the dominant, paradigm in clinical psychology (Dobson & Khatri, 2000).

Cognitive-behavioral therapies include a wide variety of clinical interventions. Research has repeatedly demonstrated their effectiveness with both youths and adults in the field of mental health services (e.g., Berman, Miller, & Massman, 1985; Dobson, 1989; Durlak, Fuhrman, & Lampman, 1991). According to Dobson and Khatri (2000), the common element of these approaches is "an emphasis on broad human change, but with a clear emphasis on demonstrable, behavioral outcomes achieved primarily through changes in the way an individual perceives, reflects upon, and, in general, thinks about their life circumstances" (p. 908). Cognitive-behaviorism assumes that cognitions affect behaviors, that we can monitor and alter our cognitive activity, and that changes in cognitions lead to changes in behaviors (Dobson & Block, 1988). Cognitive-behavioral therapies are designed to help clients become aware of thought processes that lead to maladaptive behavioral responses and to actively change those processes in a positive way (Meichenbaum, 1995).

Cognitive-behavioral therapies used with correctional populations have been conceptualized as either cognitive-restructuring, coping-

skills, or problem-solving therapies (Mahoney & Arnkoff, 1978). The cognitive-restructuring therapies view mental health problems as a consequence of maladaptive or dysfunctional thought processes, including cognitive distortions, misperceptions of social settings, and faulty logic. The coping-skills approaches focus on improving deficits in the ability to adapt to stressful situations. For example, Fabiano, Porporino, and D. Robinson (1991) argued that offenders "lack interpersonal problem-solving skills, critical reasoning skills, and planning skills" (p. 104). According to Mahoney and Arnkoff (1978), the problem-solving therapies view clients' behaviors as ineffective and maladaptive. This framework is consistent with Henning and Frueh's (1996) observation that the cognitive-behavioral programs developed for criminal offenders tend to focus on either cognitive deficits or cognitive distortions or what Kendall and Hollon (1979) called "deficits" and "excesses."

Numerous studies have been conducted in correctional settings to test the effectiveness of cognitive-behavioral techniques at reducing recidivism. This article provides both narrative and quantitative reviews of these studies. The scope of the review is limited to structured programs delivered in groups. Overall, cognitive-behavioral therapies in correctional settings consist of highly structured treatments that are detailed in manuals (Dobson & Khatri, 2000) and typically delivered to groups of 8 to 12 individuals in classroom-like settings. Highly individualized one-on-one cognitive-behavioral therapy, provided by clinical psychologists or other mental health workers, is simply not practical on a large scale within our prison system.

QUANTITATIVE SYNTHESIS METHOD

STUDY SELECTION AND RETRIEVAL

In November 1999, we searched the following computerized bibliographic databases: Dissertation Abstracts, ERIC, NCJRS, PSYCInfo, Social SciSearch, Sociological Abstracts, and Wilson Social Sciences Abstracts. Our goal was to identify all relevant evaluations that met specific inclusion criteria. The search terms were

extensive and included variations on cognitive-behavioral, cognitive-restructuring, cognitive-therapy, cognitive-rehabilitation, moral re-education, reasoning and rehabilitation (R & R), and moral treatment. These terms were crossed with terms restricting the search to offenders, criminals, and delinquents, and to studies with indicators such as recidivism, reoffense, and arrest. We identified additional studies by examining recent narrative reviews (e.g., Allen et al., 2001; MacKenzie, 2002), including works already known to us. We devoted attention to finding unpublished evaluations; the omission of unpublished studies can upwardly bias the findings of a review (Hedges, 1990; Lipsey & Wilson, 2001).

To be included in this review, a study had to meet the following inclusion criteria. First, the study must have evaluated an intervention based on a cognitive-behavioral model administered in a group setting with a structured or semi-structured treatment protocol designed to reduce criminal behaviors (e.g., cognitive life skills, moral reasoning, and cognitive restructuring). We excluded studies if the intervention focused only on social life skills or religious or spiritual concepts, or if the treatment included individual counseling. Second, the study must have included a comparison group that received either no treatment, a non-cognitive-behavioral intervention, or a minimal treatment intervention that was clearly hypothesized to be less effective. Third, the study participants must have been under the supervision of the criminal or juvenile justice system (i.e., incarcerated or on probation or parole) or directly referred to treatment from the criminal justice system. We excluded studies that provided treatment primarily to sex offenders. Fourth, the study must have reported a post-program measure of criminal behavior. Fifth, the study must have evaluated a treatment delivered in North America, Great Britain, Western Europe, or Australia (nonaboriginal) after 1979. And finally, the study must have been reported in the English language. We judged as meeting our criteria a total of 31 documents reporting on the results from 20 distinct studies.

CODING OF STUDIES

From each study, we extracted information describing the characteristics of the treatment program, offender population, research

methodology, and recidivism effects. We used a coding protocol that was pilot tested by multiple coders. Items with poor agreement or items that mapped poorly onto the characteristics of the studies were modified or dropped. For example, we needed to modify the categories for the nature of the comparison group, adding wait-list controls as an option. We repeated this process until we arrived at a coding protocol that had an acceptable level of agreement between raters and that was consistent with the characteristics of the eligible studies.

We transformed recidivism outcome data presented in the studies into an effect size, which allowed us to compare results across studies. The effect size chosen was the standardized mean difference, a widely used effect size index that can be computed from a wide variety of summary statistics that are frequently reported in primary studies (Lipsey & Wilson, 2001). In particular, this effect size index can accommodate dichotomous indicators of recidivism, such as proportion or percentage of a sample reoffending, and continuous indicators of recidivism, such as the number of new arrests or convictions (see Hedges & Hasselblad, 1995). For purposes of this review, we excluded measures based solely on technical violations or summary data based on a subset of the program-comparison sample.

We computed a total of 74 effect sizes across the 20 studies. Most were based on dichotomous indicators of recidivism (62 effect sizes). A small number (10) were based on means and standard deviations (e.g., number of arrests), and two effect sizes were based on the odds-ratio from a Cox hazard regression model (see Lipsey & Wilson, 2001, for formulas). For the purpose of the analyses that follow, we computed a single mean effect size for each study. All analyses used the random effects, inverse variance weighted method of determining the mean effect size for a collection of studies (Lipsey & Wilson, 2001). This approach weights more heavily those studies with larger samples. The larger the sample, the greater the precision in the estimate of the effectiveness of an intervention, all other things being equal. Under a fixed effects model, this meta-analysis used the inverse of the squared standard error (the inverse variance), a statistical expression of the precision of an effect, as the optimal weights. A random effects model modifies these weights based on the variability across studies. As such, a random effects model assumes uncertainty due to subject-level sampling error and study-level sampling error.

When there is large variability across studies, it is unlikely that the studies are estimating a common population effect size. The random effects model incorporates this source of uncertainty into the statistical model. The assumption is that there are true sources of variation in the effect sizes across studies that are unexplained (and potentially unexplainable) by the coded study characteristics in addition to the uncertainty due to sampling error.

RESULTS

EVALUATIONS OF COGNITIVE-BEHAVIORAL PROGRAMS FOR OFFENDERS

The two dominant cognitive-behavioral programs for offenders are moral reconnection therapy (MRT) and R & R. Roughly two thirds of the available comparison group evaluations of cognitive behavioral programs examined these two program types (see Table 1). The remaining third was a mixed collection of cognitive-behavioral programs that placed an emphasis on modifying cognitive distortions. We identified seven evaluations of other cognitive-behavioral programs that represent a mixed bag of smaller programs, often implemented at a single site.

As shown in Table 1, a full 45% of the studies were government reports, dissertations, theses, or other unpublished manuscripts. Thus, the overall results of this synthesis are unlikely to be influenced by publication bias. The year in which these documents were published (or written, in the case of unpublished works) are recent, with well more than 65% having publication dates in the later part of the 1990s, increasing the generalizability of the findings from this collection of studies to the current correctional context and offender population. Furthermore, the programs were conducted in institutional correctional facilities, such as prisons and jails, and in the community while offenders were under correctional supervision.

The thrust of this review is on the effectiveness of this class of interventions in reducing criminal behaviors. In this context, it is important to examine the evidence of effectiveness in light of the internal validity of the research designs that generated the data. We rated each study

TABLE 1: Description of Studies

<i>Variable</i>	<i>Frequency</i>	<i>Percentage</i>
Document type		
Journal article	10	50
Book chapter	2	10
Government report	1	5
Thesis/dissertation	5	25
Other	2	10
Document year		
1985-1989	3	13
1990-1994	5	22
1995-1999	15	65
Program types		
Moral reconnection therapy	6	30
Reasoning and rehabilitation	7	35
Other cognitive-behavioral	7	35
Program setting		
Prison/jail	12	60
Community (e.g., probation)	7	35
Both prison/jail and community	1	5

on a scale of 1 to 4 with a score of 4 representing the highest-level design (a true experiment), 3 a high-quality quasi-experimental design (a non-equivalent comparison group design that either constructed groups designed to be highly similar prior to the treatment or incorporated pretest measurement of offender characteristics in the analysis), 2 a lower-quality quasi-experimental design (a non-equivalent comparison group design that used a comparison group of offenders eligible for the program), and 1 equaling a minimum-level design (a non-equivalent comparison group design with obvious sources of non-equivalence between the treatment and comparison group, such as the comparison group being comprised of individuals who declined program participation). These scores are similar to scores of 5 to 2 used in the Maryland Crime Prevention Study (Sherman, Farrington, Welsh, & MacKenzie, 2002). We collapsed the lower two categories due to the small number of studies rated as 1 on this scale.

Overall, we found many strong studies to include in the review, with 20% employing random assignment to conditions (see Table 2). These true experiments provide the strongest case for the effec-

TABLE 2: Description of Methodological Characteristics

<i>Variable</i>	<i>Frequency</i>	<i>Percentage</i>
Nature of comparison group		
Wait-list control group	5	25
Nonparticipation in program(s) or management as usual	13	65
Treatment dropouts or unsuccessful participation	1	5
Alternative treatment	1	5
Quality of research design		
Experimental ^a	4	20
High-quality quasi-experimental ^b	7	35
Low-quality quasi-experimental ^c	9	45

a. Used random assignment to conditions.

b. Did not use random assignment to conditions but made attempts to control for group differences, either through design or statistical methods.

c. Obvious threats to internal validity from selection bias or other observed group differences.

tiveness of cognitive-behavioral programs. One of these four studies (D. Robinson, 1995), however, was compromised in terms of design integrity because the offenders who were randomly assigned to the wait-list control, but for whom a treatment slot became available, were dropped from the study, raising the possibility of bias from differential attrition.¹

We judged seven studies, or 35%, as using a high-quality quasi-experimental design. Despite having used nonrandomly constructed treatment and comparison groups, these studies made efforts to statistically adjust for initial group differences or provided evidence on the similarity of the treatment and comparison groups prior to the intervention. The designs for these studies had reasonably controlled for selection bias (e.g., both groups volunteering to participate in some form of a self-help program), and no other threats to internal validity were obvious. The studies with designs that we judged as low-quality were run-of-the-mill quasi-experimental designs for which selection bias posed a real threat to the validity of the findings. The typical study in this category compared individuals who self-selected into the treatment program with those who declined to participate in the program.

We examined the evidence of these studies on the effectiveness of the various program types in light of this design weakness.

We discuss the effectiveness of each of the program types in reducing recidivism below. We then discuss the overall effects and how they compare in magnitude to effects reported in studies of educational, vocation, and employment programs for offenders. Table 3 lists each study included in this synthesis, along with the study's research design, sample size, outcomes, and effect sizes.

MORAL RECONATION THERAPY (MRT)

MRT was developed by Little and K. D. Robinson (1988) for the purpose of improving social, moral, and behavioral deficits in offenders. In addition to being firmly grounded in the theoretical framework of cognitive-behaviorism, MRT draws on theoretical ideas from Kohlberg's (1976) cognitive-developmental theory of moral development. Kohlberg's theory posits that moral development progresses through six stages and that only a small percentage of the adult population ever attains the highest level of moral reasoning. Individuals with higher levels of moral development are less likely to choose behaviors that are harmful to others and, as such, are less likely to engage in criminal activities. Higher levels of moral development involve abstract thinking and perspective taking. Research has generally supported the hypothesis that juvenile delinquents and adult criminals tend to be at early stages of moral development and reasoning (Arbuthnot & Gordon, 1988). MRT views offenders as having deficits that go beyond delayed moral development. Little and K. D. Robinson stated that "clients enter treatment with low levels of moral development, strong narcissism, low ego/identity strength, poor self-concept, low self-esteem, inability to delay gratification, relatively strong defense mechanisms, and relatively strong resistance to change and treatment" (p. 135).

Despite this rather broad theoretical basis for MRT, the therapeutic elements are largely cognitive-behavioral, drawing a clear connection between thought processes and behavior. Little, K. D. Robinson, Burnette, and Swan (1996) noted that MRT's treatment methods

(text continues on page 186)

TABLE 3: Listing of All Effect Sizes and Selected Study Features by Program Type

Citation	Design	n	Outcome	Months of Follow-Up	Effect Size ^a	95% CI	
						Lower	Upper
Moral reconnection therapy Burnett, 1996	Quasi-experimental matched group design	60	Rearrest	12	0.45	-0.37	1.27
			Reincarceration	12	1.13	-0.56	2.83
Godwin, Stone, & Hambrock, 1995	Volunteers versus nonvolunteers	5,217	Returned to jail	12	0.67	0.32	1.01
			Returned to jail	24	0.31	0.06	0.56
Krueger, 1997	Volunteers versus nonvolunteers	7,128	Rearrest	48	0.50	0.01	0.99
			Rearrest	60	1.35	0.33	2.38
Little, Robinson, & Burnette, 1991a	Wait-list control group	152	Rearrest	38 ^b	0.22	-0.15	0.59
			Rearrest	60 ^b	0.12	-0.29	0.52
			Rearrest	72 ^b	0.15	-0.26	0.57
			Reincarceration	38 ^b	0.33	-0.06	0.72
			Reincarceration	60 ^b	0.40	-0.04	0.76
			Reincarceration	72 ^b	0.32	-0.03	0.68
Little, Robinson, & Burnette, 1994	Random assignment	1,381	Reincarceration	84 ^b	0.35	-0.01	0.70
			Rearrest	42 ^b	0.34	0.18	0.50
			Reincarceration	42 ^b	0.36	0.23	0.50
			Reincarceration	60 ^b	0.33	0.20	0.47
Little & Robinson, 1989	Wait-list control group	180	Rearrest for DUI/DWI	17.41 ^b	0.25	-0.25	0.74
			Rearrest for DUI/DWI	29.8 ^b	-0.05	-0.50	0.39
			Rearrest for DUI/DWI	72 ^b	-0.05	-0.42	0.31

(continued)

TABLE 3 (continued)

Citation	Design	n	Outcome	Months of Follow-Up	Effect Size ^a	95% CI	
						Lower	Upper
			Rearrest/				
			Reconviction DUI/DWI	5.28 ^b	0.13	-0.43	0.69
			Rearrest for non- DUI/DWI	17.41 ^b	0.14	-0.22	0.49
			Rearrest (any offense)	17.41 ^b	0.16	-0.18	0.50
			Rearrest (any offense)	29.8 ^b	0.36	0.02	0.71
			Rearrest (any offense)	72 ^b	0.32	-0.04	0.69
			Reincarceration (any offense)	17.41 ^b	0.29	-0.15	0.73
			Reincarceration (any offense)	29.8 ^b	0.38	0.02	0.75
			Reincarceration (any offense)	42 ^b	0.20	-0.15	0.55
			Reincarceration (any offense)	72 ^b	0.27	-0.06	0.61
			Rearrest/ Reconviction (any offense)	5.28 ^b	0.23	-0.16	0.62

Reasoning and rehabilitation	Random assignment	98	Probation revocation	8 ^b	0.11	-0.38	0.60
Johnson & Hunter, 1995	Random assignment	98	Probation revocation	8 ^b	0.11	-0.38	0.60
Porporino, Fabiano, & Robinson, 1991	Wait-list control group	63	Reincarceration	19.7 ^b	0.16	-0.41	0.72
Porporino & Robinson, 1995	Wait-list control group	73	Reincarceration	18	0.31	-0.34	0.96
			Reincarceration	10-40	0.76	0.13	1.39
			Reconviction	10-40	0.45	-0.14	1.05
Raynor & Vanstone, 1996	Participants with other probationers	207	Reconviction	12	-0.04	-0.31	0.23
			Reconviction	24	0.00	-0.28	0.27
Robinson, D., 1995	Wait-list control group	2,125	Reincarceration	12	0.07	-0.05	0.20
			Reconviction				
			(any offense)	12	0.11	-0.03	0.26
			Reconviction				
			(violent offense)	12	0.20	-0.06	0.47
			Reconviction				
			(sex offense)	12	0.53	-0.08	1.14
			Reconviction				
			(drug offense)	12	0.16	-0.22	0.53
			Reconviction				
			(nonviolent property)	12	0.19	0.00	0.38
			Reconviction				
			(robbery offense)	12	0.06	-0.30	0.41
			No. of misdemeanor person arrests	6	0.12	-0.22	0.42
			No. of felony person arrests	6	-0.22	-0.55	0.12
			No. of misdemeanor property arrests	6	-0.09	-0.43	0.24
Robinson, S. C., 1995	Retrospective comparison group	137					

(continued)

TABLE 3 (continued)

Citation	Design	n	Outcome	Months of Follow-Up	Effect Size ^a	95% CI	
						Lower	Upper
Ross, Fabiano, & Ewles, 1988 Other cognitive-behavioral programs	Random assignment	50	No. of felony property arrests	6	-0.02	-0.35	0.32
			No. of misdemeanor public order arrests	6	-0.24	-0.57	0.10
			No. of felony public order arrests	6	0.02	-0.32	0.35
			No. of arrests (total)	6	0.01	-0.33	0.34
			Rearrest	6	0.20	-0.18	0.57
Baro, 1999	Other self-help program participants used as comparison	82	New conviction		1.29	0.51	2.06
			New prison sentence		1.66	0.01	3.30
Curulla, 1991	Treatment as usual comparison	49	Assaults while incarcerated	12	0.45	0.01	0.88
			Major misconduct while incarcerated	12	0.24	-0.19	0.68
Hamberger & Hastings, 1988	Program completers compared with program dropouts	71	No. of new charges	6	0.25	-0.35	0.84
			Any new charge	6	0.87	-0.33	2.07
Henning & Frueh, 1996	Retrospective comparison group	196	Recurrent spousal violence	12	0.30	-0.23	0.82
			Rearrest or technical violation	12	0.52	0.13	0.90

Kirkpatrick, 1996	Court-ordered participants compared with nonparticipants	643	Rearrest or technical violation	24	0.66	0.30	1.02
			Rearrest or technical violation	36	0.69	0.33	1.05
			Rearrest or technical violation	32 ^b	0.45	0.09	0.81)
			Recidivism	12	0.58	0.40	0.76
Menton, 1999	Comparison group consisted of jail inmates who left the jail prior to program participation	326	Rearrest	8	0.19	-0.09	0.47
			Rearrest	8	0.19	-0.09	0.47
			Rearrest	12	0.18	-0.11	0.47
			Rearrest	18	0.30	-0.04	0.63
			Rearrest	24	0.24	-0.17	0.64
			Rearrest	30	-0.14	-0.80	0.52
			Domestic violence reoffense	8	1.10	0.78	1.42
			Domestic violence reoffense	12	1.07	0.74	1.40
			Domestic violence reoffense	18	1.01	0.62	1.39
			Domestic violence reoffense	24	0.98	0.51	1.44
Moody, 1997	Quasi-experimental comparison group with generally similar characteristics	28	Domestic violence reoffense	30	0.52	-0.16	1.20
			Recommitted to training school	18	0.00	-0.82	0.82

a. Standardized mean difference effect size.

b. Mean or median length of follow-up in months.

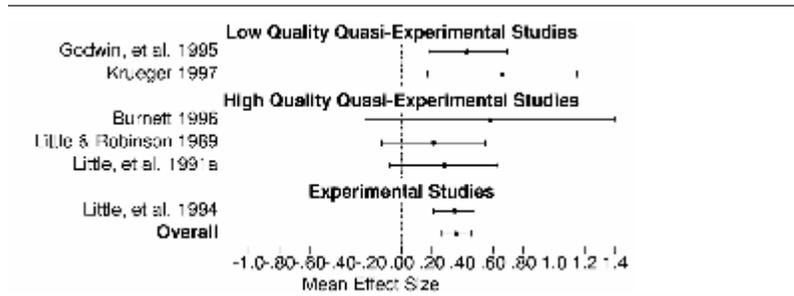


Figure 1: Mean Study Level Effect Size and 95% Confidence Interval by Methodological Quality for Moral Reconciliation Therapy

cifically, MRT is a structured program that makes use of a manual with clearly described exercises and lessons directed at groups of 10 to 15 offenders. Each session lasts 1 to 2 hours, and there are usually two sessions per week. Participants are given a workbook that contains the exercises and tasks that constitute the program (Little & K. D. Robinson, 1986). These exercises are highly varied and include a discussion of the source of unhappiness, prison disloyalty, identification of goals, an exploration of both the good and bad times in one's life, and the behaviors that help make the event bad.

We were able to identify six comparative evaluations of the effects of MRT on the future offending behavior of program participants (see Figure 1). The general pattern of results is positive across this collection of studies for all three levels of research quality. The single experimental evaluation of this cognitive-behavioral approach to offender treatment was conducted by Little, K. D. Robinson, and Burnette (1994). This study evaluated the effects of MRT for the general offender population in the Shelby County Correctional Facility in Memphis, Tennessee. The limited number of treatment slots allowed for the random assignment of offenders who expressed an interest in the program's treatment and control conditions. The follow-up recidivism data for the treatment group includes program completers and dropouts. The 5-year recidivism rate for the MRT condition was 41% compared with 56% for the comparison offenders (effect size = 0.33, $p < .001$). Furthermore, the MRT participants had lower levels of criminal involvement at all follow-up periods on all indicators of

recidivism, providing strong evidence of the effectiveness of this program.

All three high-quality quasi-experimental studies found positive effects of MRT, although the overall effect sizes are not statistically significant because small sample sizes resulted in large confidence intervals (i.e., low statistical power). The first of these, conducted by Burnett (1996), evaluated the effectiveness of MRT among parolees. This quasi-experimental design matched treatment and control individuals on age, gender, ethnicity, and time period under the jurisdiction of the corrections department. The 1-year rearrest and recidivism rates favored the treatment group. Given the rather small sample size of 60 offenders, the moderate to large average effect size of 0.58 did not reach conventional levels of statistical significance. This difference, however, is clinically significant—that is, it represents a meaningful reduction in the rate of reoffense (a reduction in the rearrest rate from 20% to 10% and a reduction in the reincarceration rate from 10% to 0%).

The second high-quality quasi-experimental design evaluated the effects of MRT on convicted drunk drivers in a southern state (Little & K. D. Robinson, 1989; Little, K. D. Robinson, & Burnette, 1990, 1991a, 1993a; Little, K. D. Robinson, Burnette, & Swan, 1995a). The study included 115 convicted drunk drivers in a county jail who agreed to participate in a treatment program compared with 65 convicted drunk drivers who volunteered but were not selected due to limited treatment slots. Study participants were followed, on average, for a total of 6 years. Early follow-ups showed a small difference favoring the moral reconnection participants with regard to rearrest for a DUI/DWI. However, this difference disappeared over time. The effect of moral reconnection on criminal behavior was generally more positive at all measurement points. The average effect across measurement points and different indices of recidivism was positive and modest (0.21), albeit statistically nonsignificant. Although this study did not use random assignment to conditions, a wait-list design generally has strong internal validity (Cook & Campbell, 1979). The threat of selection bias is reduced when all subjects volunteer for the program. Unfortunately, participation in MRT in this study was confounded with participation in other alcohol-related therapy—specifically, residence on the alcohol treatment unit during the offender's period of

incarceration. Thus, it is unclear whether the positive findings from this study are attributable to participation in MRT or to some other aspect(s) of the treatment regimen, such as Alcoholics Anonymous or other educational programming.

The third high-quality evaluation of MRT was conducted by Little, K. D. Robinson, and Burnette (1991b, 1993b) and assessed MRT effects with felony drug offenders (see also Little, K. D. Robinson, Burnette, & Swan, 1995a, 1995b). The control group consisted of felony drug offenders who applied for the treatment during the same time period as the treated offenders but did not participate due to an insufficient number of treatment slots—that is, a wait-list condition. Thus, both treated and nontreated offenders volunteered for the program and were drawn from the same larger population. Four measures of recidivism were used, and at the final follow-up point, study participants had 7 years, on average, at risk for reoffense. The effect attenuated only slightly from the first to the final follow-up period. The average effect was modest to moderate in size (0.28) and statistically nonsignificant. Two of the individual effects were reported as statistically significant by the authors, and all effects favored the moral reconditioning condition. Of the three high-quality quasi-experimental designs, this had the strongest interval validity and observed an average effect quite similar to one reported in the experimental study by Little et al. (1994).

A methodologically weak evaluation of the effects of MRT, conducted by Godwin, Stone, and Hambrock (1995), also showed a positive overall effect (average effect size of 0.43, $p < .01$). This study compared 98 male offenders who had voluntarily participated in the MRT program with all other offenders released during the same time period from the same short-term detention center in Florida. This study did not control for any offender differences that might be related to self-selection into the therapy program, and as such, it is impossible to determine whether the observed difference is due to self-selection or the moral reconditioning program. The difference is most likely a function of both.

Krueger (1997) reported the 4- and 5-year recidivism rates for participants in a county jail-based MRT program compared with a random sample of all other county jail inmates who did not participate in the program. The rearrest rates were substantially lower for the MRT

participants (e.g., 45% vs. 67% at 48 months, and 62% and 95% at 60 months). Unfortunately, this study did not control for selection bias and, as such, provides little basis for concluding that MRT is effective, despite the positive findings.

The mean recidivism effect size across the six evaluations of MRT is positive and statistically significant (mean effect size = 0.36; see Figure 1). Furthermore, this collection of studies is statistically homogeneous, indicating that the differences in observed effects across studies are no more variable than we would expect due to subject-level sampling error. Stated more simply, the studies tell a consistent story. All six evaluations found positive effects, although half were not statistically significant due to insufficient statistical power. Analyzing only the four higher quality studies produces essentially the same result, with a mean effect size of 0.33 ($p < .001$). Thus, there is reasonably strong evidence for the effectiveness of MRT at reducing long-term recidivism rates among offenders.

Three of the four methodologically stronger studies were conducted by the developers of MRT (see Little & K. D. Robinson, 1989; Little et al., 1991b, 1994), raising the question of whether the findings generalize to MRT programs run by other program personnel. The positive results from the studies not conducted by Little and colleagues are encouraging but currently insufficient to draw strong generalizations. The availability of a manual, as well as the highly structured nature of the program, increases the likelihood that the integrity of the program can be maintained when administered by a range of criminal justice personnel.

REASONING AND REHABILITATION (R & R)

R & R was developed by Ross and Fabiano (1985) and, like MRT, is based on the premise that offenders have cognitive and social competency deficits. Rather than focusing on moral reasoning, however, the program is directed at enhancing self-control, cognitive style, interpersonal problem solving, social perspective taking, critical reasoning, and values (e.g., prosocial attitudes). Ross, Fabiano, and Ewles (1988) stated that the

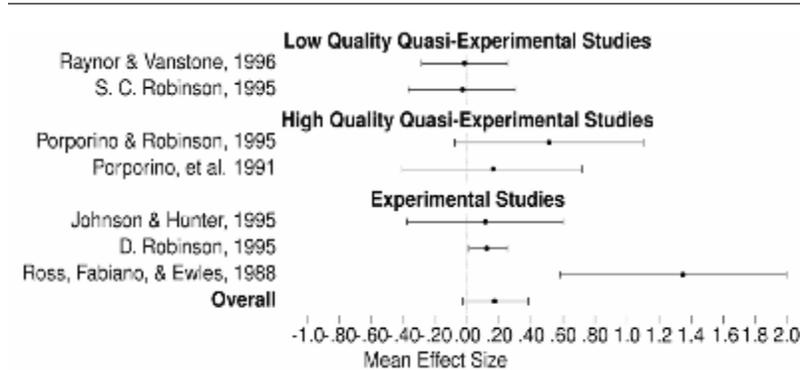


Figure 2: Mean Study Level Effect Size and 95% Confidence Interval by Methodological Quality for Reasoning and Rehabilitation

program focused on modifying the impulsive, egocentric, illogical and rigid thinking of the offenders and teaching them to stop and think before acting, to consider the consequences of their behaviour, to conceptualize alternative ways of responding to interpersonal problems and to consider the impact of their behaviour on other people, particularly their victims. (p. 31)

The goal is to develop “more effective problem-solving and coping skills, more reflective and deliberate thinking patterns, and both more pro-social and more consistent attitudes, values, and beliefs” (Porporino & D. Robinson, 1995, p. 161).

The program is divided into 35 sessions. The program generally runs 8 to 12 weeks depending on the number of sessions per week. The program occurs in a group context with 6 to 8 participants in a classroom-like setting. The sessions include a mix of “audio-visual presentations, games, puzzles, reasoning exercises, role-playing, modeling, and group discussion techniques and strategies” (Porporino & D. Robinson, 1995, p. 161).

We identified seven evaluations of R & R programs. Three of these were true experimental studies. The results are mixed across the seven studies, although all of the higher quality studies found that program recipients offended at lower rates than nonrecipients (see Figure 2).

The three true experiments all found positive results, although the difference in recidivism between conditions was not statistically significant in the Johnson and Hunter (1995) study. Johnson and Hunter randomly assigned drug offenders to the specialized drug offender program with the R & R program or the specialized drug offender program without the R & R component. At an average of 8 months after assignment to conditions, the R & R participants were recidivating at a slightly lower rate (26%) compared with the non-R & R participants (29%), translating into a small positive effect size (0.11). Recidivism was measured as probation revocations and outstanding warrants issued (absconsions).

A small effect favoring R & R was also found by D. Robinson (1995, 1996; Porporino & D. Robinson, 1995), with a mean effect size across outcome measures of 0.12 ($p < .05$). This was a large, 5-year study with 2,125 participants. During the first 3 years, participants were randomly assigned to either the program or a wait-list condition. However, the randomization process was abandoned during the final 2 years of the study. The control condition continued to consist of offenders who volunteered for the program but for whom there was insufficient space. Control group offenders for whom a space became available were allowed to participate in the program and were dropped from the study. This compromised the integrity of the randomization, for we do not know if the availability of slots for the controls followed a random process. Participants were offenders under federal jurisdiction in Canada, some of whom were institutionalized during participation in the program, whereas others participated while in the community. All effects favored the treatment condition, with effect sizes that ranged from small (0.06) to moderate (0.53). It is also worth noting that these effect sizes were based on analyses that included program dropouts (17% of the sample). As would be expected, the effects are substantially larger when based only on program completers.

Ross et al. (1988) also used an experimental design to evaluate the R & R program. This study was restricted to high-risk male probationers, and the program was delivered by trained probation officers. Offenders were randomly assigned to probation with or without R & R ($n = 25$ in each condition). The difference in the proportion con-

victed of new offenses or sentenced to prison favored the treatment condition by more than 2 to 1. The average effect size across these two outcome measures is very large (effect size = 1.35) and statistically significant. Even if we assume that the three treatment condition cases lost due to attrition recidivated and that the two control condition cases that also were lost due to attrition did not recidivate, the overall effect size is still large and statistically significant. The rather large effect, however, could be attributed to the instability of an estimate from a small sample.

Porporino and Robinson (1995; D. Robinson, 1995, 1996; D. Robinson, Grossman, & Porporino, 1991) reported on a small sample ($n = 73$) evaluation of R & R on high-risk offenders. The study used a wait-list control design without random assignment (i.e., it was a high-quality quasi-experimental study). All participants in this study volunteered for the program, and admittance into the program was independent of individual characteristics, such as motivation for treatment. Offenders for whom a slot in the program never became available served as the controls, and pretest data suggested that the treatment and control groups were similar on observed variables. This study found a positive and statistically significant difference favoring the R & R group on the proportion with a prison readmission (37% for the R & R group and 70% for the comparison group). The average effect size across the three indicators of recidivism was moderate (0.51) but statistically nonsignificant. Both this study and the previous study by Ross et al. (1988) suggest that the R & R programs can be effective with high-risk offenders.

Another study by Porporino and colleagues (Porporino, Fabiano, & D. Robinson, 1991; see also Porporino & D. Robinson, 1995; D. Robinson et al., 1991) also used a wait-list control group design without random assignment to conditions. This study served as the pilot study for the D. Robinson (1995) experimental evaluation of R & R discussed earlier. Participants were adult prison inmates in Canada. There was a small positive effect favoring the offenders who entered the program, whether or not they completed it, compared with the wait-list controls (effect size of 0.16, $p > .05$). The effect size was based on the reinstitutionalization rate for all offenders assigned to the program compared with the wait-list comparison group (reinstitutionalization rates of 45% and 52%, respectively). Furthermore, the

recidivism rate for those completing the program was lower than both the wait-list comparison group and the program noncompleters.

A variant of the R & R program, called Straight Thinking on Probation (STOP), was evaluated by Raynor and Vanstone (1996), who compared the 12- and 24-month reconviction rates for participants of STOP with several types of offenders (these being offenders on probation, given a suspended sentence, sentenced to community service, sentenced to prison, sentenced to other custodial institutions, and young offender sentenced to an institution). For purposes of computing effect size, the "other probation" condition was selected because it was judged to be the most comparable to the STOP with probation condition. This contrast showed a slightly negative effect for the STOP program. The only positive effects reported by the researchers were in analyses of STOP completers compared with other probation groups. This study suffers from obvious threats to internal validity, reducing the strength that can be placed on the overall finding of no program effect.

Using a retrospective comparison group, S. C. Robinson (1995) evaluated the effectiveness of R & R for juveniles sentenced to a Utah detention center. The retrospective controls were comparable to the program participants on demographics and prior criminal activity. The effects ranged from a small positive effect favoring the R & R condition (effect size of 0.20 for percentage recidivating) to a small negative effect favoring the controls (effect size of -0.24 for the number of public order offenses). The average effect across the eight indicators of recidivism was slightly negative. None of the observed effects were statistically significant. It is important to note that this study restricted the R & R sample to participants who attended 90% of the program sessions. The slightly negative effect is puzzling and might reflect some unobserved difference between the two groups.

Taken as a whole, the evaluation evidence supports the conclusion that R & R is effective at reducing future criminal behavioral among offenders, including high-risk offenders. The overall mean effect size for the experimental and high quality quasi-experimental studies is positive and statistically significant (mean effect size = 0.16, $p < .05$). The magnitude of this effect size is small, however. Furthermore, there is significant variability in the results across studies ($Q = 10.9$, $df = 4$, $p < .03$), suggesting differential effectiveness across studies.

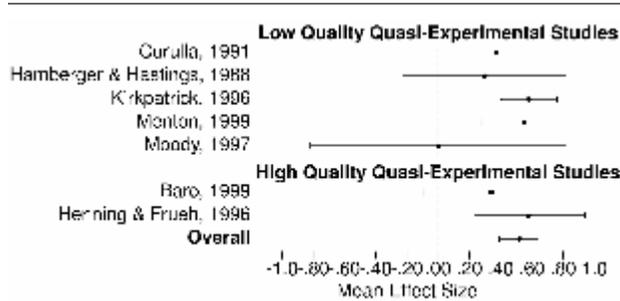


Figure 3: Mean Study Level Effect Size and 95% Confidence Interval by Methodological Quality for Other Cognitive-Behavioral Programs

R & R programs have been tested on a wider scale than MRT, with one evaluation examining effects across a large number of correctional institutions throughout Canada. Thus, the integrity of the program might have been compromised in the large-scale implementation. Additional research is needed to determine the sensitivity of the program to contextual changes and degradations to program integrity.

OTHER COGNITIVE-BEHAVIORAL PROGRAMS

This section includes a variety of structured cognitive-behavioral programs implemented in group settings. Many of these programs focus on cognitive restructuring, including the modification of cognitive distortions and faulty logic or perceptions. In contrast, the MRT and R & R programs have a distinctly deficit orientation. This is particularly true of the R & R program that attempts to strengthen cognitive deficits in several areas, including self-control, critical reasoning, social perspective taking, and interpersonal problem solving (Fabiano, D. Robinson, & Porporino, 1991).

All but one of the studies in this category reported lower rates of criminal offending behavior, generally of a moderate to large difference, between the cognitive-behavioral program participants and the comparison sample (see Figure 3). The single zero effect was for a small study (Moody, 1997) of a unique intervention. Furthermore, the research design for that study was flawed. Only two of the seven studies in this group had reasonably strong research designs; none were

true experiments with random assignment to conditions. We briefly discuss each of these seven studies.

The Strategies for Thinking Productively evaluated by Baro (1999) focuses on helping the offenders “identify key thinking patterns that have led to criminal behavior” and “realistic alternatives” (Baro, 1999, p. 470). Following an 8-week, highly structured program phase, the offenders enter a less structured phase that requires them to keep a journal of problematic situations and associated cognitions and to discuss these situations and cognitions with program staff. Participants in the program were compared with participants in other prison-based self-help programs, such as Narcotics Anonymous, Alcoholics Anonymous, religious and cultural programs, and education programs. Only offenders who participated in at least 8 weeks of an alternative self-help program were included in the comparison condition. The two groups were demonstrated to be similar on observed variables, including age and risk for property or assault offenses. Offenders in both conditions willingly participated in the programs that they selected. It is unknown whether this choice is related to future offending, but presumably, both groups were motivated to make positive changes in their lives, reducing the threat from selection. The difference in the 12-month follow-up rates for the number of assaults and major misconducts while incarcerated favored the strategies of the cognitive-behavioral participants (effect sizes of 0.45 and 0.24). The effect for assaults was statistically significant, despite the rather small sample size (41 offenders per group). The average effect across these two outcomes was small to moderate (effect size = 0.34, $p = .12$).

Henning and Frueh (1996) evaluated a cognitive-behavioral program that focused on the modification of cognitive distortions and the development of self-monitoring. The study participants ($n = 196$) were adult male prison inmates, and the research used a retrospective comparison group design—a generally weak research design from an internal validity perspective. However, the study retained treatment dropouts in the treated condition and was therefore a more conservative test of the effectiveness of the cognitive program. The researchers also used a Cox hazard regression model to statistically adjust for observed initial differences. Hence, we categorized this study as a high-quality quasi-experimental design. An effect size based on the

odds-ratio from the Cox hazard regression model was moderate in size and favored the program participants (effect size = 0.45, $p < .01$).

Using a sample of learning disabled offenders, Curulla (1991) evaluated the effectiveness of an aggression replacement training program that included social skills training, anger management, and moral reasoning. The comparison condition received no special treatment but was similar in their background characteristics, including being diagnosed as learning disabled. Participation in the program was mandated by a judge. The offenders in the control condition were found suitable for the program but were not mandated to attend. The overall effect size for the number and percentage with new charges was small to moderate and favored the aggression replacement training program (effect size = 0.37, $p > .05$). The weaknesses of this study are the very small sample size (16 persons in the treatment condition and 33 in the control condition) and the lack of control over the selection process.

Hamberger and Hastings (1988) conducted a methodologically weak evaluation of a cognitive-behavioral program for male batterers. The community-based violence abatement program consisted of a variety of components, including cognitive restructuring, communication skills enhancement, assertiveness training, and relaxation training. The quasi-experimental design compared program completers to dropouts and found that program completers had a lower rate of recurrent spousal violence (34% vs. 47%, respectively), translating into a small-to-moderate effect size, which was statistically non-significant (effect size = 0.30, $p > .05$). It is quite likely that program completers were more motivated to change their battering ways than program dropouts.

Also using a weak research design, Kirkpatrick (1996) evaluated the effectiveness in reducing criminal recidivism of a cognitive restructuring program with a strong moral reasoning component. The program focused on correcting 10 criminal thinking errors using Biblical references and Christian doctrines. The program also included social-skills and social-problems components. The research compared court-ordered program participants with nonparticipants and found a moderate difference in recidivism after 12 months between groups that favored the treatment condition (effect size = 0.58, $p < .01$). All participants were adult male offenders under community-based supervision. The research design did not control for selection

bias, and as such, little weight can be placed on these findings, although they are encouraging.

Menton (1999) conducted a low-quality quasi-experimental study examining a cognitive restructuring type program for male domestic abusers. The program was conducted while the offenders were in a county jail. The comparison offenders were domestic abusers who left the jail before having an opportunity to participate in the program. As such, they were more likely to have had a less serious domestic violence offense or criminal history. For purposes of comparison, the treatment condition included those who completed the program as well as those who did not. Recidivism effect sizes for any reoffense (including domestic violence) for 8- through 30-month follow-up periods were small to moderate and favored the treatment condition, with the exception of the final follow-up. None of these differences reached statistical significance. Effect sizes for domestic violence reoffenses were large at the 8-month follow-up (1.10) but moderate at the 30-month follow-up (0.52). The difference in recidivism rates between conditions was statistically significant for all but the last follow-up. The average effect size across measures and time points was moderate and statistically significant (0.55, $p < .01$), suggesting a positive effect for cognitive behavioral programs with domestic abusers, especially if the author's assumption was correct that the program participants were at higher risk for recidivism without the treatment than the nonparticipants. Although this assumption seems reasonable, it is untestable.

Finally, Moody (1997) evaluated a "pair" counseling program with male juveniles in a residential facility. Pair counseling involves two previously unconnected adolescents who meet with a counselor to develop social interaction skills. The program includes discussion of moral dilemmas using cognitive-behavioral methods. The control group consisted of youths in the same facility who were of similar age to the youths in the treatment condition. No other attempts to control for differences between groups was employed, and the study had a small sample ($n = 28$). Half of the participants in both conditions were recommitted to a training school at the 18-month follow-up (effect size = 0.00). The higher level of prior criminal involvement of the youth in the treatment condition might have biased the study against finding a positive effect.

Overall, the mean effect across this diverse collection of cognitive-behavioral programs was moderate in size (mean effect size = 0.51) and statistically significant ($p < .001$). In general, the quality of the studies in this category was low. The mean effect size for the two higher quality studies was also moderate (mean effect size = 0.48, $p < .001$). As was the case with most of the programs in this category, both of these studies (Baro, 1999; Henning & Frueh, 1996) focused on cognitive distortions rather than cognitive deficits. This evidence suggests that cognitive-distortions-based treatment approaches to corrections-based offender rehabilitation can be effective, but the data are far from convincing given the methodological weaknesses of the studies in this category.

DISCUSSION

The evidence summarized in this article supports the claim that cognitive-behavioral treatment techniques are effective at reducing criminal behaviors among convicted offenders. All of the higher quality studies found positive effects favoring the cognitive-behavioral treatment program. The random-effects mean effect sizes for the higher quality studies is 0.32 ($p < .001$), a moderate effect size. Removing the single outlier (Ross et al., 1988) reduces the mean effect size only slightly (0.27). Furthermore, without this one extreme value, the distribution is homogeneous ($Q = 11.4$, $df = 9$, $p = .25$). Only 2 of the 20 studies found negative overall effect sizes, both of which were near 0 and from studies of low quality.

Comparing the mean effect sizes across higher quality MRT, R & R, and other cognitive-behavioral programs suggests that R & R might be less effective than the other two (mean effect sizes of 0.33, 0.16, and 0.49, respectively; all are statistically significant at $p < .05$). This should be interpreted cautiously, for the findings for R & R were less consistent across studies, with one R & R study reporting the largest effect across all studies in this review. The larger R & R effects were observed by the smaller studies, raising the possibility that the smaller effects might be due to treatment integrity problems associated with large-scale program implementation and not the effectiveness of R & R core technology.

An important issue is the practical significance of these findings: Are these effects large enough to produce meaningful reductions in recidivism? One method of interpreting the mean effects presented in the current article is to translate them into recidivism rate difference for treated and untreated offenders. The mean effect size of 0.33 for the high-quality MRT studies translates into a 16-percentage-point difference in recidivism rates between the conditions (42% for the treated and 58% for the untreated). This is by no means a large effect, but it is of clear practical value. The recidivism rate difference for the mean effect size of 0.16 for the R & R high-quality studies is 8 percentage points (46% for the treatment and 54% for the untreated). Effect sizes of 0.20 and less are considered small (Cohen, 1988), and clearly an 8-percentage-point reduction in recidivism is small. Lipsey (1992) has argued, however, that such small effects can lead to meaningful reductions in community-level criminal behavior when such programs are implemented on a large scale, as has occurred for this program. That is, a small reduction in the offending behavior of a large number of offenders will still represent a large number of crimes prevented.

Other benchmarks for interpreting the cognitive-behavioral program findings are the mean effects for other correctional programs. A recent meta-analysis of corrections-based education, vocation, and work programs (Wilson, Gallagher, & MacKenzie, 2000) showed that recidivism rate difference ranged from 7% for multicomponent programs to 13% for postsecondary education programs. The bulk of the studies synthesized by Wilson et al. (2000) failed to adequately control for selection bias. The typical evaluations of education, vocation, and work programs simply compared program participants with program nonparticipants. These effects, therefore, are likely to be upwardly biased. The evidence for the effectiveness of cognitive-behavioral programs is substantially stronger, and the effects of cognitive-behavioral programs are equal to or slightly larger than those of education, vocation, and work programs.

The various programs discussed here have different names. Some have a theoretical basis that emphasized cognitive deficits, such as problem-solving skills, whereas others emphasized cognitive distortions, such as blaming others. Despite these differences, all of these programs have common structures and contents. In general, the pro-

grams encourage offenders to become more aware of their thought processes that either initiate or sustain their choices to engage in criminal acts. What cannot be determined from the preceding literature are the specific elements or combinations of elements that are critical in producing positive effects on offenders' behaviors. The evidence suggests that both deficit and distortion approaches can be effective as well as programs that emphasize moral teachings and reasoning. Further research is needed to gain insight into the "active ingredients" of these programs.

From a policy perspective, the active ingredients are less important than distinguishing between effective and ineffective rehabilitation programs. A question that remains unanswered by this research is whether these programs will remain effective when implemented on a large scale and when the training of program staff is provided by someone other than the program developers. A common finding in the evaluation literature is that the effectiveness of programs is reduced as the integrity of program design and implementation is compromised. The small effect sizes found for the R & R program when evaluated on a large scale throughout the Canadian federal prison system provides some evidence of this compromise occurring with cognitive-behavioral programs. The highly structured nature of these programs helps ensure program integrity but does not guarantee it. Further research is needed to understand how best to train the staffs of these programs.

NOTE

1. Due to this potential threat to internal validity, this study was rated as a 3, not a 4, on the method quality scale. For clarity of exposition, it is displayed with the other experimental studies in Figure 2.

REFERENCES

References marked with an asterisk were included in the quantitative synthesis.

Allen, L. C., MacKenzie, D. L., & Hickman, L. J. (2001). The effectiveness of cognitive behavioral treatment for adult offenders: A methodological, quality-based review. *International Journal of Offender Therapy and Comparative Criminology*, 45, 498-514.

- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., Cullen, F. T., et al. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- Arbuthnot, J., & Gordon, D. A. (1988). Crime and cognition: Community applications of socio-moral reasoning development. *Criminal Justice and Behavior*, 15, 379-393.
- *Baro, A. L. (1999). Effects of a cognitive restructuring program on inmate institutional behavior. *Criminal Justice and Behavior*, 26, 466-484.
- Berman, J. S., Miller, R. C., & Massman, P. J. (1985). Cognitive therapy versus systematic desensitization: Is one treatment superior? *Psychological Bulletin*, 97, 451-461.
- *Burnett, W. L. (1996). *Treating post-incarcerated offenders with moral reconnection therapy™: A one-year recidivism study*. Unpublished master's thesis, University of Phoenix, Phoenix, AZ.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cullen, F., & Gendreau, P. (1989). The effectiveness of correctional rehabilitation. In L. Goodstein & D. L. MacKenzie (Eds.), *The American prison: Issues in research policy* (pp. 23-44). New York: Plenum.
- *Curulla, V. L. (1991). Aggression replacement training in the community for adult learning-disabled offenders (Doctoral dissertation, University of Washington, 1991). *Dissertation Abstracts International*, 53(2-A), 627. (UMI No. AAD92-16113)
- Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 57, 414-419.
- Dobson, K. S., & Block, L. (1988). Historical and philosophical bases of the cognitive-behavioral therapies. In K. S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (pp. 3-38). New York: Guilford.
- Dobson, K. S., & Khatri, N. (2000). Cognitive therapy: Looking backward, looking forward. *Journal of Clinical Psychology*, 56, 907-923.
- Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin*, 110, 204-214.
- Fabiano, E. A., Porporino, F. J., & Robinson, D. (1991). Canada's cognitive skills program corrects offenders' faulty thinking. *Corrections Today*, 53(5), 102-108.
- Fabiano, E., Robinson, D., & Porporino, F. J. (1991). *A preliminary assessment of the cognitive skills training program: A component of living skills programming: Program description, research findings and implementation strategy*. Ottawa: Correctional Service of Canada.
- Gendreau, P., & Andrews, D. A. (1990). Tertiary prevention: What the meta-analysis of the offender treatment literature tells us about "what works." *Canadian Journal of Criminology*, 32, 173-184.
- Gendreau, P., & Ross, R. R. (1987). Revivification of rehabilitation: Evidence from the 1980s. *Justice Quarterly*, 4, 349-408.
- *Godwin, G., Stone, S., & Hambrook, K. (1995). Recidivism study: Lake County, Florida Detention Center. *Cognitive-Behavioral Treatment Review*, 4, 6.
- *Hamberger, L. K., & Hastings, J. E. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, 3(2), 121-130.
- Hedges, L. V. (1990). Directions for future methodology. In K. W. Wachter & M. L. Straf (Eds.), *The future of meta-analysis* (pp. 11-26). New York: Russell Sage Foundation.
- Hedges, L. V., & Hasselblad, V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin*, 117, 167-178.

- *Henning, K. R., & Frueh, B. C. (1996). Cognitive-behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Corrections cognitive self-change program. *Criminal Justice and Behavior*, 23, 523-541.
- Husband, S. D., & Platt, J. J. (1993). The cognitive skills component in substance abuse treatment in correctional settings: A brief review. *Journal of Drug Issues*, 23, 31-42.
- *Johnson, G., & Hunter, R. M. (1995). Evaluation of the Specialized Drug Offender Program. In R. R. Ross, & B. Ross (Eds.), *Thinking straight* (pp. 215-234). Ottawa, Canada: Cognitive Center.
- Kendall, P. C., & Hollon, S. D. (1979). Cognitive-behavioral interventions: Overview and current status. In P. C. Kendall & S. D. Hollon (Eds.), *Cognitive-behavioral interventions: Theory, research and practice* (pp. 1-9). New York: Academic Press.
- *Kirkpatrick, B. L. H. (1996). Cognitive restructuring: Effects on recidivism (Doctoral dissertation, Ball State University, 1996). *Dissertation Abstracts International*, 57(6-A), 2680.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior* (pp. 31-55). New York: Holt, Rinehart & Winston.
- *Krueger, S. (1997). Five-year recidivism study of MRT-treated offenders in a county jail. *Cognitive-Behavioral Treatment Review*, 6, 3.
- Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, et al. (Eds.), *Meta-analysis for explanation: A casebook* (pp. 83-127). New York: Russell Sage Foundation.
- Lipsey, M. W., & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Little, G. L., & Robinson, K. D. (1986). *How to escape your prison: A moral reconnection therapy workbook*. Memphis, TN: Eagle Wing Books.
- Little, G. L., & Robinson, K. D. (1988). Moral reconnection therapy: A systematic step-by-step treatment system for treatment-resistant clients. *Psychological Reports*, 62, 135-151.
- *Little, G. L., & Robinson, K. D. (1989). Treating drunk drivers with moral reconnection therapy: A one-year recidivism report. *Psychological Reports*, 64, 960-962.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1990). Treating drunk drivers with moral reconnection therapy: A two-year recidivism study. *Psychological Reports*, 66, 1379-1387.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1991a). Treating drunk drivers with moral reconnection therapy: A three-year report. *Psychological Reports*, 69, 953-954.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1991b). Treating drug offenders with moral reconnection therapy: A three-year recidivism report. *Psychological Reports*, 69, 1151-1154.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1993a). 42-month alcohol treatment data: Multiple DWI offenders treated with MRT show lower recidivism rates. *Cognitive-Behavioral Treatment Review*, 2, 5.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1993b). Cognitive behavioral treatment of felony drug offenders: A five-year recidivism report. *Psychological Reports*, 73, 1089-1090.
- *Little, G. L., Robinson, K. D., & Burnette, K. D. (1994). Treating offenders with cognitive-behavioral therapy: 5-year recidivism outcome data on MRT. *Cognitive-Behavioral Treatment Review*, 3, 1-3.

- *Little, G. L., Robinson, K. D., Burnette, K. D., & Swan, S. (1995a). Six-year MRT™ recidivism data on felons and DWI offenders: Treated offenders show significantly lower reincarceration. *Cognitive-Behavioral Treatment Review*, 4, 1-5.
- *Little, G. L., Robinson, K. D., Burnette, K. D., & Swan, S. (1995b). Seven-year recidivism of felony offenders treated with MRT™. *Cognitive-Behavioral Treatment Review*, 4, 6.
- *Little, G. L., Robinson, K. D., Burnette, K. D., & Swan, S. (1996). Review of outcome data with MRT™: Seven-year recidivism results. *Cognitive-Behavioral Treatment Review*, 5, 1-7.
- Losel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works: Reducing reoffending—Guidelines from research and practice* (pp. 79-111). New York: John Wiley.
- MacKenzie, D. L. (2002). Criminal justice and crime prevention. In L. W. Sherman, D. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, & S. Bushway (Eds.), *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: National Institute of Justice.
- MacKenzie, D. L., & Hickman, L. J. (1998). *What works in corrections? An examination of the effectiveness of the type of rehabilitation programs offered by Washington State Department of Corrections* (Report to the State of Washington Legislature Joint Audit and Review Committee). College Park, MD: University of Maryland, Department of Criminology and Criminal Justice.
- Mahoney, M. J., & Arnkoff, D. B. (1978). Cognitive and self-control therapies. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 689-722). New York: John Wiley.
- Martinson, R. (1974). What works? Questions and answers about prison reform. *The Public Interest*, 35, 22-54.
- Meichenbaum, D. H. (1995). Cognitive-behavioral therapy in historical perspective. In B. Bongar & L. Beutler (Eds.), *Comprehensive textbook of psychotherapy: Theory and practice* (pp. 140-158). New York: Oxford University Press.
- *Menton, P. C. (1999). The effect of a domestic violence program on incarcerated batterers (Doctoral dissertation, Boston University, 1999). *Dissertation Abstracts International*, 59(8-A), 3217.
- Moody, E. E., Jr. (1997). Lessons of pair counseling with incarcerated juvenile delinquents. *Journal of Addictions and Offender Counseling*, 18, 10-25.
- *Porporino, F. J., Fabiano, E. A., & Robinson, D. (1991). *Focusing on successful reintegration: Cognitive skills training for offenders, R-19*. Ottawa: Research and Statistics Branch, Correctional Service of Canada.
- Porporino, F. J., & Robinson, D. (1995). An evaluation of the Reasoning and Rehabilitation Program with Canadian federal offenders. In R. R. Ross & B. Ross (Eds.), *Thinking straight* (pp. 155-191). Ottawa, Canada: Cognitive Centre.
- *Raynor, P., & Vanstone, M. (1996). Reasoning and rehabilitation in Britain: The results of the Straight Thinking on Probation (STOP) programme. *International Journal of Offender Therapy and Comparative Criminology*, 40, 272-284.
- *Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders*. Canada: Correctional Research and Development, Correctional Service of Canada.
- *Robinson, D. (1996). Factors influencing the effectiveness of cognitive skills training. *Forum on Corrections Research*, 8(3), 6-9.
- *Robinson, D., Grossman, M., & Porporino, F. J. (1991). *Effectiveness of the cognitive skills training program: From pilot to national implementation, B-07*. Ottawa: Research and Statistics Branch, Correctional Service of Canada.

- *Robinson, S. C. (1995). Implementation of the cognitive model of offender rehabilitation and delinquency prevention (Doctoral dissertation, University of Utah, 1995). *Dissertation Abstracts International*, 55(8-A), 2582.
- Ross, R. R., & Fabiano, E. A. (1985). *Time to think: A cognitive model of delinquency prevention and offender rehabilitation*. Johnson City, TN: Institute of Social Sciences and Arts, Inc.
- *Ross, R. R., Fabiano, E. A., & Ewles, C. D. (1988). Reasoning and rehabilitation. *International Journal of Offender Therapy and Comparative Criminology*, 32, 29-36.
- Sherman, L. W., Farrington, D. P., Welsh, B. C., & MacKenzie, D. L. (Eds.). (2002). *Evidence-based crime prevention*. London: Harwood Academic Publishers.
- Whitehead, J. T., & Lab, S. P. (1989). A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime and Delinquency*, 26, 276-295.
- Wilson, D. B., Gallagher, C. A., & MacKenzie, D. L. (2000). A meta-analysis of corrections-based education, vocation, and work programs for adult offenders. *Journal of Research in Crime and Delinquency*, 37, 347-368.