# 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 reconation 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

(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 reconation, 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 reconation 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**

| *Variable* | *Frequency* | *Percentage* |
|---|---|---|
| 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 reconation 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**

| *Variable* | *Frequency* | *Percentage* |
|---|---|---|
| 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 reconation 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 |
| | | | Reincarceration | 84[b] | 0.35 | –0.01 | 0.70 |
| Little, Robinson, & Burnette, 1994 | Random assignment | 1,381 | 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 | 95% CI 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 | | | | | | | |
| 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 |
| Robinson, S. C., 1995 | Retrospective comparison group | 137 | 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 |

*(continued)*

**TABLE 3 (continued)**

| *Citation* | *Design* | n | *Outcome* | *Months of Follow-Up* | *Effect Size*[a] | *95% CI* | |
|---|---|---|---|---|---|---|---|
| | | | | | | *Lower* | *Upper* |
| | | | 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 |
| Ross, Fabiano, & Ewles, 1988 | Random assignment | 50 | New conviction | | 1.29 | 0.51 | 2.06 |
| | | | New prison sentence | | 1.66 | 0.01 | 3.30 |
| Other cognitive-behavioral programs | | | | | | | |
| Baro, 1999 | Other self-help program participants used as comparison | 82 | Assaults while incarcerated | 12 | 0.45 | 0.01 | 0.88 |
| | | | Major misconduct while incarcerated | 12 | 0.24 | –0.19 | 0.68 |
| Curulla, 1991 | Treatment as usual comparison | 49 | No. of new charges | 6 | 0.25 | –0.35 | 0.84 |
| | | | Any new charge | 6 | 0.87 | –0.33 | 2.07 |
| Hamberger & Hastings, 1988 | Program completers compared with program dropouts | 71 | Recurrent spousal violence | 12 | 0.30 | –0.23 | 0.82 |
| Henning & Frueh, 1996 | Retrospective comparison group | 196 | Rearrest or technical violation | 12 | 0.52 | 0.13 | 0.90 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | 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) |
| Kirkpatrick, 1996 | Court-ordered participants compared with nonparticipants | 643<br>643 | 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<br>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 |
| | | | Domestic violence reoffense | 30 | 0.52 | –0.16 | 1.20 |
| Moody, 1997 | Quasi-experimental comparison group with generally similar characteristics | 28 | 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.

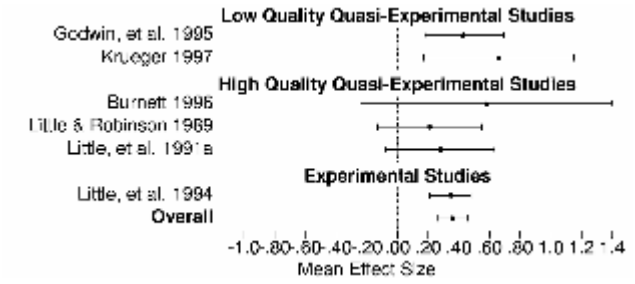


**Figure 1: Mean Study Level Effect Size and 95% Confidence Interval by Methodological Quality for Moral Reconation Therapy**

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 reconation participants with regard to rearrest for a DUI/DWI. However, this difference disappeared over time. The effect of moral reconation 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 reconation 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 reconation 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

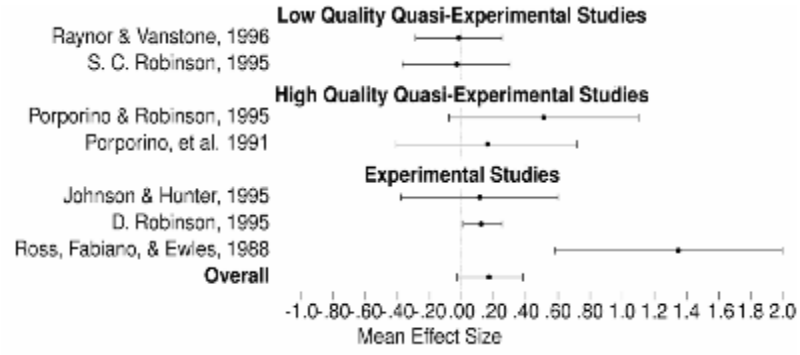


**Figure 2: Mean Study Level Effect Size and 95% Confidence Interval by Methodological Quality for Reasoning and Rehabilitation**

> 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.

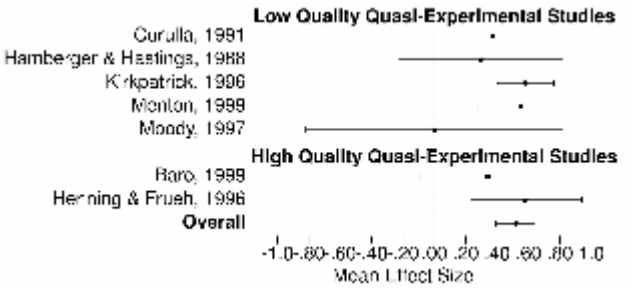


**Figure 3: Mean Study Level Effect Size and 95% Confidence Interval by Methodological Quality for Other Cognitive-Behavioral Programs**

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 nonsignificant (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 reconation 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., & Hambrock, 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 reconation therapy workbook*. Memphis, TN: Eagle Wing Books.

Little, G. L., & Robinson, K. D. (1988). Moral reconation 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 reconation 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 reconation 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 reconation 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 reconation 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.