Offender Coercion in Treatment: A Meta-Analysis of Effectiveness
Karen K. Parhar, J. Stephen Wormith, Dena M. Derkzen and Adele M. Beauregard
Criminal Justice and Behavior 2008; 35; 1109
DOI: 10.1177/0093854808320169

The online version of this article can be found at:
http://cjb.sagepub.com/cgi/content/abstract/35/9/1109

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
International Association for Correctional and Forensic Psychology

Additional services and information for Criminal Justice and Behavior can be found at:

Email Alerts: http://cjb.sagepub.com/cgi/alerts

Subscriptions: http://cjb.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://cjb.sagepub.com/cgi/content/refs/35/9/1109
Mandating offenders to attend correctional treatment is a controversial function of the justice system, in part because of the uncertainty about the effectiveness of such practice. A meta-analysis was conducted to compare the effectiveness of mandated, coerced, and voluntary correctional treatment in reducing recidivism. A search of correctional treatment studies resulted in 129 studies meeting the inclusion criteria. In general, mandated treatment was found to be ineffective in several analyses, particularly when the treatment was located in custodial settings, whereas voluntary treatment produced significant treatment effect sizes regardless of setting. Few significant differences in effect sizes were found between levels of coercion. The implications of mandating correctional treatment for offenders are discussed.

Keywords: meta-analysis; coercion; mandated treatment; voluntary treatment; correctional treatment

The evidence that correctional treatment programs can be effective in reducing recidivism rates is now well established (e.g., Hollin, 1999; Lipsey, Chapman, & Landenberger, 2001; McGuire, 2002). More specifically, principles of effective correctional treatment—such as risk, need, and responsivity (RNR)—that increase the likelihood of success and decrease recidivism rates most dramatically have been well documented (e.g., Andrews, Bonta, & Hoge, 1990; Andrews, Bonta, & Wormith, 2006; Andrews, Zinger, et al., 1990). Consequently, there appears to be strong support for correctional treatment programs within correctional agencies (McGuire, 2004) and among many in the judiciary (W. Hall, 1997).

Perhaps because of this revived confidence in the effectiveness of correctional treatment, the use of court-ordered or legally mandated treatment, which is accompanied by “threats of legal consequences if individuals refuse to comply with a referral to treatment” (Polcin & Greenfield, 2003, p. 650), and coercive treatment, whereby refusal to participate in treatment results in negative consequences (Day, Tucker, & Howells, 2004), has become increasingly
commonplace. For example, the number of legally mandated correctional treatment programs for substance abusers and spousal assailters has increased significantly within the past decade (Holtzworth-Munroe, 2001; Polcin, 1999; Rosenbaum & Geffner, 2002; Wells-Parker, 1994). Furthermore, court-ordered counseling for substance abusers has now become common practice in many regions and states of the United States (Shearer & Baletka, 1999).

Another possible explanation for the increased use of court-ordered programs is the belief that few offenders will enter treatment without some sort of external motivation (Farabee, Prendergast, & Anglin, 1998). Therefore, it may be argued that treatment should be legally mandated to safeguard the community. However, another perspective commonly held by clinicians is that treatment can be effective only if the participants choose to attend treatment because of their own internal motivations (Shearer & Baletka, 1999).

MANDATED AND NONMANDATED OFFENDER TREATMENT

Although only a few studies have compared the recidivism rates of participants attending treatment under a legal mandate to those who attend voluntarily (e.g., “Principles of Drug Addiction,” 2000), some evidence supports the concept of legally mandated treatment. After reviewing 11 empirical studies of compulsory substance abuse treatment programs, Farabee et al. (1998) concluded that legal referral to substance abuse treatment programs is an effective strategy for improving outcome and enhancing retention and compliance. This finding is supported by other drug treatment studies that have concluded that mandated treatment contributes to treatment completion (Rempel & Destefano, 2001; Siddall & Conway, 1988).

Moreover, comparative studies have found that those who are legally mandated to treatment are more likely to stay in treatment longer than those who are not legally mandated to treatment (e.g., Collins & Allison, 1983; Schnoll, Goldstein, Antes, & Rinella, 1980). For example, Rosenberg and Liftik (1976) compared retention rates between coerced and voluntary participants and found that offenders coerced to participate in an alcoholism treatment program had better attendance rates than did voluntary patients. In addition, Maxwell (2000) found that greater perceived legal threat increased retention in drug treatment, regardless of the client’s actual legal status.

Other studies have found less favorable results for mandated treatment. Howard and McCaughrin (1996) found that organizations in which 75% of the cases were court mandated had a higher rate of treatment failure than did organizations with few court-mandated clients. Examining outcome studies of court-ordered treatment programs for spousal assailters, Rosenfeld (1992) found that mandated treatment did not reduce recidivism. Rather, men who were arrested for spousal assault and were not treated were just as likely to recidivate as those who completed treatment. Moreover, court-mandated and nonmandated participants were equally likely to withdraw from treatment. Concerning two domestic violence treatment studies that found positive results for court-mandated treatment (i.e., Dutton, 1986; Waldo, 1988), Rosenfeld (1992) speculated that the participants who completed treatment were likely to have been the most motivated and least treatment resistant.

There is also evidence that nonmandated treatment programs are associated with treatment retention and compliance. Harford, Ungerer, and Kinsella (1976) found that offenders not under any legal pressure to attend drug abuse treatment remained in treatment for longer periods than did clients who were under legal pressure. Shearer and Ogan (2002) measured
treatment resistance for voluntary and forced participants who were residing in a substance abuse treatment program, in a prerelease therapeutic community, or in a therapeutic community in a substance abuse treatment facility as a condition of probation. They found that treatment resistance was significantly lower among offenders who perceived they had volunteered for treatment.

Comparative studies have also found no differences between the effectiveness of mandated treatment and nonmandated treatment programs. Research on substance abuse treatment by the National Institute on Drug Abuse has found that involuntary treatment of offenders is just as effective as voluntary treatment (“Principles of Drug Addiction,” 2000). More specifically, the report stated that “individuals who enter treatment under legal pressure have outcomes as favourable as those who enter treatment voluntarily” (p. 19). Similarly, Prendergast, Farabee, Cartier, and Henkin (2002) did not find any differences between voluntary and involuntary clients in a drug treatment program on most measures of psychosocial functioning. Moreover, a meta-analysis of the effects of coercion in a community sex offender treatment program found a negative, but nonsignificant, relationship between coercion and recidivism (Gray, 1998).

**RESEARCH ISSUES**

Existing comparisons of mandated and voluntary participation in correctional treatment may be misleading because these studies are plagued with conceptual and methodological problems. For example, truly voluntary participation does not exist in the criminal justice system because there is always some degree of external pressure (Wild, 1999). Wild (1999) concluded that the minimal difference found between mandated and nonmandated clients in substance abuse treatment may have occurred because he really compared mandated versus coerced clients and not mandated versus truly voluntary participants.

Furthermore, Farabee et al. (1998) identified a number of definitions of coercion, claiming that coercion can be used to refer to a range of situations, such as the following:

- A probation officer’s recommendation to enter treatment, a drug court judge’s offer of a choice between treatment or jail, a judge’s requirement that the offender enter treatment as a condition of probation, or a correctional policy of sending inmates involuntarily to a prison treatment program in order to fill the beds. (p. 3)

Farabee et al. stated that even being involved in the criminal justice system alone is sufficient to be considered coercion.

Another conceptual difficulty occurs because some programs do not accept clients into treatment unless they are at least somewhat willing to participate, even if they are legally mandated to attend treatment (e.g., Dutton, 1986; Taylor, Davis, & Maxwell, 2001). In other words, mandated clients who enter treatment may actually volunteer to enter treatment even though they are legally required to participate. More specifically, in psychological rehabilitation programs, offenders cannot be physically forced to attend a treatment session; even if they do attend, they cannot be forced to participate fully, although they can be pressured (Day et al., 2004).

One of the most confounding difficulties in assessing the effectiveness of mandated versus voluntary treatment is the offender’s motivation to attend treatment, regardless of mandate. One study found that more than three fourths of the patients who reported no control over their admission to an outpatient clinic planned to continue treatment when external pressure
was removed and thereby appeared to have volunteered for treatment (Farabee, Shen, & Sanchez, 2002).

Consequently, discussions about mandated treatment evoke questions about client motivation. Motivation is commonly described on a continuum from external motivation to intrinsic motivation. According to McMurran (2002), externally motivated behavior occurs in the presence of an external reward or some other kind of external contingency, whereas intrinsically motivated behavior occurs in the absence of any external reward. A person undertakes intrinsically motivated behavior out of interest because it is optimally challenging and is based on innate psychological needs. When behavior change is the result of intrinsic motivation, the changes last longer; behavior change that is extrinsically motivated lasts only as long as the extrinsic controls are in place (Ryan & Deci, 2000). Consequently, offenders who are intrinsically motivated to desist from crime may be more likely to be successful in the long term. Intrinsic motivation has been associated with positive behavioral changes such as better learning, performance, and well-being (e.g., Benware & Deci, 1984; Deci, Schwartz, Sheinman, & Ryan, 1981; Grolnick & Ryan, 1987).

Generally speaking, mandated treatment provides external motivation to attend treatment, and those who voluntarily attend treatment are intrinsically motivated. Thus, it is not surprising that a large study of the general public, substance abuse counselors, probationers, and judges reported that all respondents believed that compulsory substance abuse treatment is less effective than voluntary treatment (Wild et al., 2001). In light of the evidence supporting the benefits of voluntary treatment attendance, it is hypothesized that nonmandated or voluntary treatment will have greater impact on recidivism and will correspond with lower rates of treatment attrition than mandated treatment.

Because most studies do not assess offender motivation for treatment, the present meta-analysis cannot directly address this issue. However, many offender treatment studies describe the legal context of their intervention and the conditions that may apply to their clientele in some detail. This should allow one to estimate the extent to which external pressure may be found in “voluntary” treatment (e.g., pressure from family) and internal motivation may be present even in mandated treatment (e.g., the offender would have volunteered regardless of mandate). Specifically, a 5-level coercion–voluntary scale, ranging from 1 (mandated or involuntary) to 5 (nonmandated or voluntary), was used to assess the extent of coercion in a sample of treatment studies. In this manner, variations in the degree of coercion were established in an attempt to resolve Wild's (1999) concern regarding falsely classifying treatment participants into dichotomous categories of purely mandated and voluntary because categories fail to truly exist. Most interventions do not fall in either of these two extremes.

Another possible confound in the comparisons of mandated versus voluntary treatment concerns the effectiveness of the treatment strategy being adopted in any given study. In their meta-analysis, Andrews, Zinger, et al. (1990) recorded effectiveness of dramatically attenuated effect sizes between studies that adhered to the principles of RNR and those that did not. These findings have since been updated and supported (Andrews et al., 2006). It is quite possible that either mandated or voluntary treatment interventions are more or less effective, irrespective of the degree of voluntariness that is offered. Previous studies in this literature have not considered any other factors that might explain or camouflage differences in the effectiveness of voluntary or mandated interventions.

The goal of this article is to examine the relationship between the level of treatment coerciveness and treatment outcome as measured by treatment retention and recidivism. To
achieve this objective, the degree of coercion was correlated with percentage of treatment dropouts, and unweighted treatment effect sizes and magnitudes of the treatment effect for mandated, coerced, and voluntary treatment programs were compared.

METHOD

SAMPLE OF STUDIES

A search for all relevant articles pertaining to mandated and nonmandated treatment of offenders was conducted through a number of key social scientific computerized databases. These included PsycINFO, Criminal Justice Abstracts, Medline, and Sociological Abstracts. The search was based on the following keywords used individually or in various combinations with each other: mandatory, voluntary, compulsory, coercive, treatment, intervention, offender, criminal, inmate, recidivism, arrest, control, and comparison. Finally, the bibliographies of articles published in 2004, 2005, and 2006 that met the inclusion criteria were searched for additional articles that might meet the inclusion criteria.1

INCLUSION CRITERIA

The following information was required from each study to be included in the meta-analysis: Participants or the treatment in question were identified as either legally mandated or nonmandated, a comparison or control group was included in the study, and recidivism outcome measures were reported either as a Pearson product–moment correlation coefficient $r$ (effect size) or in a manner that could be transformed into a Pearson $r$. In 31 cases, authors were contacted by e-mail or telephone when the information was present but clarification on any of the above criteria was required (e.g., when it was unclear if the treatment group was mandated or nonmandated or when specific recidivism information was needed to convert to $r$). That is, authors were contacted when there was enough information to code the items but some information was unclear or specific details were omitted from the article. Consequently, studies were excluded from the analyses if there was no indication of whether the program or participants in the treatment group were legally mandated or not mandated, the recidivism data could not be analyzed (e.g., only $F$ statistics were reported), or no comparison or control group was included. In addition, studies comparing only mandated and nonmandated treatment participants, but with no control group, were excluded from the analyses because such a design does not afford an opportunity to compute a comparable effect size.

STUDY CHARACTERISTICS AND CODING

More than 500 studies were identified in the initial search and reviewed for possible inclusion. A total of 139 studies met the selection criteria. Of these, 46 (33.1%) studies involved exclusively mandated treatment programs, 83 (59.7%) involved exclusively nonmandated treatment programs, and 10 (7.2%) were mixed mandated and nonmandated treatment participants. A total of 69 items was used to code the articles. The item categories included background information (e.g., author, year, country, etc.), participants (e.g., gender, mean age, etc.), program information (e.g., location, research design, mandate, treatment type, treatment quality, sample size, treatment dropouts, etc.), and recidivism (e.g., length of follow-up, type of recidivism, number of rearrests, etc.).
The assessment of coercion and voluntariness was coded in three levels of specificity: a 2-level mandated–nonmandated variable, a 3-level mandated–coerced–voluntary variable, and a 5-point ordinal scale. The detailed, 5-point coercion–voluntary scale ranged from 1 (mandated or involuntary) to 5 (nonmandated or voluntary). Each level of the scale was developed by the authors and subjectively ordered to reflect the variations in levels of coercion reported in the treatment studies based on the legal mandate of any available information regarding the level of perceived or actual coercion and/or implementation of the mandate. The most coercive end of the scale, mandated or involuntary, and the less coercive midrange scores of the scale (2 to 4) were defined similarly to the definitions provided by Klag, O’Callaghan, and Creed (2005) of compulsory and coerced treatment in their review of legal coercion in the treatment of substance abusers.

The 5 levels and their definitions were as follows: 1 = mandated involuntary (it is clear that offenders must take the program [if they do not attend treatment they face incarceration or other negative consequences]; it is a condition of their release; they are court referred; treatment is included in sentence), 2 = mandated coerced (offenders are mandated to treatment, but there is evidence that consequences are minimal or nil if they do not participate [e.g., they are not consistently punished]), 3 = nonmandated coerced with legal consequences (there is no mandate for offenders to participate in treatment [they are volunteers], but there are some legal consequences if they do not attend or complete treatment [e.g., get sent back to court]), 4 = nonmandated coerced (there is no mandate for offenders to participate in treatment, but they may receive incentives if they attend treatment, such as early release [this is usually in a parole or institutional setting]), and 5 = nonmandated or voluntary (offenders freely volunteer to attend treatment without evidence of any external costs or benefits).2

Ten articles reported mixed mandated and nonmandated treatment participants. However, because the mixed treatment studies could not be analyzed with the remaining studies without jeopardizing the integrity of the study, they were excluded from the analyses, resulting in a sample size of 129 studies. This 5-level coding scheme was then used to create the two simpler measures of coercion–voluntariness, the 2-level measure where scores of 1 to 2 (mandated) and 3 to 5 (voluntary) were collapsed, and the 3-level measure where scores of 1 (mandated), 2 to 4 (coerced), and 5 (voluntary) were recoded. These multiple levels of precision were used to gain a better understanding of previous findings.

Because correctional treatment programs that adhere to the principles of RNR produce greater effect sizes (Andrews, Zinger, et al., 1990), a treatment quality variable was created to assess the extent to which the treatment program adhered to the principles of RNR. Treatment studies that adhered to zero or one of the principles of RNR were coded as inappropriate and treatment programs that adhered to two or all three of the principles of RNR were coded as appropriate.

QUALITY CONTROL AND RELIABILITY

Each study was coded by one of three independent coders, consisting of the first, third, and fourth authors of the study. To assess the reliability of the coding process, the three coders recoded a random selection of each other’s articles in a blind fashion. Of the articles meeting the inclusion criteria, 67 were randomly selected for recoding by a second coder. Any disagreements were discussed by all three coders and resolved among themselves in a
consensus fashion. The percentage agreement between the two coders of each article for several coded variables was calculated. To obtain an 80% agreement rate among coders on all coded variables, some minor modifications were made to the coding scheme, such as collapsing across similar categories of a variable, and agreement rates were reassessed. For example, the percentage agreement for research design was 76% across the five original levels used to code the variable. When research design was recoded and collapsed into two categories (strong vs. weak design), the percentage agreement was 82%. Following these modifications to the coding scheme, all variables used in the analyses reached or exceeded the interrater acceptability criteria of 80%, a level that has been considered acceptable in several other meta-analyses (e.g., Dowden, Antonowicz, & Andrews, 2003).

META-ANALYTIC STRATEGY

**Independence of comparisons.** The effect sizes collected for the meta-analysis report only independent comparisons in which a particular experimental group is statistically compared to just one comparison group. When several comparison groups were used in a single study, the highest quality comparison group was used. For example, random or matched-control groups were chosen over control groups composed of treatment dropouts. When more than one follow-up time was provided for the outcome data, comparisons based on the longest follow-up were used. Where studies reported more than one recidivism measure, recidivism was coded for arrest or the most inclusive recidivism measure reported. For example, if charges and convictions were reported, only charges were coded. Both general and specific (e.g., sexual, violent) recidivism rates were coded when reported.

**Effect size estimates.** Where possible, cell frequencies (treatment vs. control by crime desistance vs. recidivism) were calculated and transformed into the Pearson product-moment correlation coefficient ($r$). Where cell frequencies were unavailable, other statistical tests were transformed into $r$ (e.g., $t$ tests). Pearson’s $r$ and confidence intervals were selected to report the results to clearly describe the preciseness and magnitude of the results in a manner that is widely understood (Gendreau & Smith, 2007).

The Hunter–Schmidt (random effects) method for combining correlation coefficients was used to calculate the mean $r$, or effect size, because it tends to provide the most accurate estimates of the mean population effect size when the effect sizes of individual studies are heterogeneous (A. P. Field, 2001), as is commonly the case in meta-analytic studies of offender treatment programs. A “bare-bones” Hunter–Schmidt technique, which corrects only for sampling error (Hunter & Schmidt, 2004), was used in the present study as a method for combining correlation coefficients because there is rarely enough information reported in a study to use the full Hunter–Schmidt technique (A. P. Field, 2001).

The population correlation is estimated by using a weighted average in which each correlation is weighted by the number of participants in the study (Hunter & Schmidt, 2004). As cited by A. P. Field (2001), the equation used to calculate the mean $r$ ($Mr$), where $\Sigma =$ summation and $n =$ the sample size of each study, is

$$Mr = \frac{\Sigma n_i r_i}{\Sigma n_i}.$$  \hspace{1cm} (1)

The equation used to calculate the standard deviation of $Mr$ is
$$SD_r = ((\Sigma n_i (r_i - Mr)^2) / (\Sigma n_i))^{1/2}. \quad (2)$$

The equation used to calculate the standard error of the population correlation, where \(k\) = the number of effect sizes, is

$$SE_{Mr} = SD_r / k^{1/2}. \quad (3)$$

The equation used to calculate the \(z\) score of the population correlation is

$$z = Mr / SE_{Mr}. \quad (4)$$

The equation used to test the homogeneity of the effect sizes, where \(\chi^2 = \text{chi-square}\), is

$$\chi^2 = \Sigma ((n_i - 1)(r_i - Mr)^2 / (1 - Mr^2)^2). \quad (5)$$

Although the Fisher’s \(z’\) transformation is commonly used to correct for bias in the Pearson \(r\), Hunter and Schmidt (2004) do not recommend using a Fisher’s \(z’\) transformation because it produces an estimated mean correlation that is upwardly biased and less accurate than an untransformed mean correlation. Therefore, the statistical comparisons made in this study used the untransformed correlation coefficients.

Effect sizes were calculated for three different types of recidivism: general, specific, and any recidivism. Effect sizes of general recidivism were derived from all studies that reported general recidivism (\(k = 106\)), whereas effect sizes of specific recidivism were derived from studies that reported specific (e.g., sexual recidivism for sexual offender treatment programs) recidivism (\(k = 52\)). Effect sizes of any recidivism (\(k = 129\)) were based on all studies that reported general recidivism (\(k = 106\)); where general recidivism was not reported, specific recidivism was used (\(k = 23\)). Although all studies contributed only one coefficient to each calculation,3 a single study could have contributed to more than one of these calculations if it provided outcome data on more than one type of recidivism.

**RESULTS**

**STUDY CHARACTERISTICS**

Table 1 summarizes the 129 studies published between 1970 and 2005 (\(Mdn = 1998\)) that were used in the current analyses. The majority were conducted in the United States (80.6%) and Canada (14.7%). Most of the 118 studies (78.0%) reporting the age of the treatment participants involved participants exclusively in the adult justice system. More than half of the 120 studies (56.7%) that reported treatment setting took place exclusively in the community. More than half of the 113 studies (54.0%) reporting gender consisted of males only, whereas most of the remaining studies (42.5%) consisted of both males and females. More than half of the studies (55.0%) were classified as having a strong research design (random assignment or quasi-experimental). Nonmandated treatment programs were more common (64.3%) than mandated treatment programs (35.7%). Approximately half of the 126 studies (48.8%) that were coded for treatment quality were described as inappropriate and half were appropriate (51.2%).
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mandated Treatment</th>
<th>Nonmandated Treatment</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>Sample size (ns)</td>
<td>384.3</td>
<td>629</td>
<td>20-3,974</td>
</tr>
<tr>
<td>Follow-up (months)</td>
<td>26.5</td>
<td>25.8</td>
<td>6-126</td>
</tr>
<tr>
<td>Attrition (%) (ns)</td>
<td>39.4</td>
<td>19.6</td>
<td>10.7-77</td>
</tr>
<tr>
<td>Country (%)</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>87.0</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>10.9</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>2.2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>0.0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Setting (%)**</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community</td>
<td>71.7</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>4.3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Custody</td>
<td>23.9</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>24.3</td>
<td>8.3</td>
<td>14.4-35.9</td>
</tr>
<tr>
<td>Justice system (%)**</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult</td>
<td>57.5</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Juvenile</td>
<td>42.5</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Research design (%) (ns)</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong</td>
<td>60.9</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td>39.1</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Treatment quality</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriate</td>
<td>53.3</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Inappropriate</td>
<td>46.7</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Gender (%)</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>48.8</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>4.9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>46.3</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>
**TABLE 1: (continued)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mandated Treatment</th>
<th>Nonmandated Treatment</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>Level of coercion (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandated</td>
<td>35.7</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Involuntary</td>
<td>29.5</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Coerced</td>
<td>6.2</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Nonmandated</td>
<td>64.3</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>With legal cons.</td>
<td>8.5</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Coerced</td>
<td>18.6</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Voluntary</td>
<td>37.2</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Treatment targets (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance abuse</td>
<td>38.6</td>
<td>17</td>
<td>37.8</td>
</tr>
<tr>
<td>Sex offences</td>
<td>6.8</td>
<td>3</td>
<td>20.7</td>
</tr>
<tr>
<td>Juvenile delinquency</td>
<td>22.7</td>
<td>10</td>
<td>8.5</td>
</tr>
<tr>
<td>Violence</td>
<td>20.5</td>
<td>9</td>
<td>9.8</td>
</tr>
<tr>
<td>Unspecified</td>
<td>2.3</td>
<td>1</td>
<td>15.9</td>
</tr>
<tr>
<td>Drunk driving</td>
<td>2.3</td>
<td>1</td>
<td>3.7</td>
</tr>
<tr>
<td>High risk or serious</td>
<td>4.5</td>
<td>2</td>
<td>2.4</td>
</tr>
<tr>
<td>Theft or property</td>
<td>2.3</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>Treatment type (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive-behavioral</td>
<td>25.6</td>
<td>11</td>
<td>39.5</td>
</tr>
<tr>
<td>Therapeutic community</td>
<td>27.9</td>
<td>12</td>
<td>16.0</td>
</tr>
<tr>
<td>General counseling</td>
<td>9.3</td>
<td>4</td>
<td>18.5</td>
</tr>
<tr>
<td>Multisystemic therapy</td>
<td>14.0</td>
<td>6</td>
<td>8.6</td>
</tr>
<tr>
<td>Psychoeducational</td>
<td>16.3</td>
<td>7</td>
<td>4.9</td>
</tr>
<tr>
<td>Sanctions</td>
<td>4.7</td>
<td>2</td>
<td>8.6</td>
</tr>
<tr>
<td>Drug or medications</td>
<td>0.0</td>
<td>0</td>
<td>2.5</td>
</tr>
<tr>
<td>Vocational</td>
<td>2.3</td>
<td>1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

**p < .01, between-group comparisons.**
The treatment programs included in the study addressed a variety of offender issues. The most frequently reported objectives targeted substance abuse (38.1%), sex offending (15.9%), juvenile delinquency (13.5%), and violence (13.5%). The most common treatment modality was cognitive-behavioral treatment (34.7%), followed by therapeutic community (20.2%), general counseling (15.3%), multisystemic therapy (10.5%), and psychoeducational treatment (8.9%).

RECIDIVISM

A total of 106 (82.2%) articles included general recidivism information, and 52 (40.3%) articles included specific recidivism information that was directly related to the treatment in which offenders participated. Among the 106 studies reporting general recidivism, 104 indicated the source of the data; most were based on official records of recidivism (90.4%, \( k = 94 \)), few were based on unofficial records (3.9%, \( k = 5 \)) and both official and unofficial records (3.9%, \( k = 5 \)). With respect to type of recidivism, 105 studies provided data, and arrest was used in half (49.5%, \( k = 52 \)) of the studies, followed by incarceration (16.2%, \( k = 17 \)) and convictions (12.4%, \( k = 13 \)). Other types of recidivism reported were a combination of measures (5.7%, \( k = 6 \)), charges (4.8%, \( k = 5 \)), violations (3.8%, \( k = 4 \)), court contact (2.9%, \( k = 3 \)), other (2.9%, \( k = 3 \)), self-report (1.0%, \( k = 1 \)), and unspecified (1.0%, \( k = 1 \)).

Considering the 52 studies reporting specific recidivism, 50 reported the source of the data, and most were based on official recidivism (74.0%, \( k = 37 \)); few were based on unofficial records (12.0%, \( k = 6 \)) and both official and unofficial records (14.0%, \( k = 7 \)). Of the 51 studies reporting type of recidivism data, arrest was the most common type of specific recidivism reported (29.4%, \( k = 15 \)), followed by convictions (11.8%, \( k = 6 \)), violations (9.8%, \( k = 5 \)), and self-reported recidivism (9.8%, \( k = 5 \)). Other types of recidivism reported were a combination of measures (7.8%, \( k = 4 \)), charges (7.8%, \( k = 4 \)), police records (5.9%, \( k = 3 \)), drug use (5.9%, \( k = 3 \)), other (3.9%, \( k = 2 \)), unspecified (3.9%, \( k = 2 \)), incarceration (2.0%, \( k = 1 \)), and court contact (2.0%, \( k = 1 \)).

The overall mean weighted treatment effect size for the 129 studies using the Hunter–Schmidt (random effects) method was .06 (95% confidence interval = .05 to .08). The \( z \) score as measured by Equation 4 indicated that the effect size was significant at 7.10, \( p < .01 \). However, the chi-square test of the homogeneity of variance indicated that the sample of effect sizes had a heterogeneous variance, \( \chi^2(128, N = 129) = 908.96, p < .01 \).

Generally, weighted mean effect sizes are considered superior to unweighted mean effect sizes when estimating the population effect size (Hunter & Schmidt, 2004). However, when one of the studies has an extremely large sample size relative to the other studies, it may be advisable to use the unweighted average effect size. One study in the current review reported a sample size of 32,490. Because this was more than 8 times greater than the second largest sample size (\( N = 3,994 \)), this study was considered an outlier. In cases such as these, it is recommended that comparisons of interest be examined both with and without the outlier (Hunter & Schmidt, 2004). The total weighted mean \( r \) in the current study increased, but not significantly, to .09 (95% confidence interval = .07 to .11) when the outlier study was excluded. This increase occurred because the outlier study had an extremely large sample size but a small effect size. Therefore, weighted mean effect sizes, including and excluding the outlier, and unweighted effect sizes are reported below.
ANALYSIS OF STUDY VARIABLES BY LEVEL OF COERCION

Prior to analyzing the mean effect sizes of the mandated (k = 46) and nonmandated (k = 83) treatment programs, analyses were conducted to determine whether there were differences between the two groups on study characteristics. Mandated treatment programs were significantly more likely to be located in the community, \( \chi^2(1, N = 120) = 9.51, p < .01 \), and consist of juvenile offenders, \( \chi^2(1, N = 118) = 14.75, p < .01 \), than nonmandated treatment programs. Likewise, less voluntary treatment, as measured using the 5-point coercion–voluntary scale, was more likely to be located in the community, \( t(128) = -4.26, p < .001 \), and consist of juvenile offenders, \( t(116) = -3.879, p < .001 \), than more voluntary treatment. There were no significant differences between mandated and nonmandated treatment programs on quality of research design, \( \chi^2(1, N = 129) = 0.98, p > .05 \), treatment quality (appropriate or inappropriate), \( \chi^2(1, N = 125) = 0.13, p > .05 \), sample size, \( t(127) = -0.93, p > .05 \), or percentage of treatment dropouts, \( t(54) = 1.62, p > .05 \). Similarly, no differences were found between level of coercion and research design, \( t(127) = 1.028, p > .05 \), treatment quality, \( r(126) = .04, p > .05 \), or sample size, \( r(129) = .09, p > .05 \), using the 5-point coercion–voluntary scale. Unweighted treatment effect size was not correlated with the number of RNR principles to which the treatment adhered for any, \( r(126) = .10, p > .05 \), general, \( r(103) = .16, p > .05 \), or specific recidivism, \( r(51) = -.08, p > .05 \).

Because two potentially confounding differences in study characteristics were found between mandated and nonmandated treatment, further analyses were conducted to determine whether these variables were related to treatment effect size. Treatment programs conducted in the community were more effective than treatment offered in custody as measured by unweighted effect sizes on any, \( t(118) = 2.40, p = .02 \), and specific recidivism, \( t(46) = 2.03, p < .05 \), but not general recidivism, \( t(96) = 1.20, p > .05 \). Moreover, an ordinal measure of setting, ranging from 1 to 3, consisting of community, mixed settings, and custodial settings, respectively, resulted in significant correlations on the unweighted effect sizes of any, \( r(128) = -.21, p = .02 \), and specific recidivism, \( r(51) = -.28, p = .05 \), but not general recidivism, \( r(106) = -.12, p > .05 \). No differences were found between the unweighted effect sizes for treatment offered to adult and juvenile offenders on any, \( t(116) = -1.80, p > .05 \), general, \( t(30.83) = -0.88, p > .05 \), or specific, \( t(48) = 0.30, p > .05 \), recidivism. Therefore, it was necessary for comparisons of treatment effectiveness between mandated and nonmandated treatment in the following analyses to consider treatment setting.

RELATION BETWEEN LEVEL OF COERCION AND UNWEIGHTED EFFECT SIZE ESTIMATES

To examine the relationship between amount of coercion on the 5-point coercion–voluntary scale and magnitude of effect size, correlational analyses were conducted using the unweighted effect sizes of all studies (any recidivism), studies reporting general recidivism, and studies reporting specific recidivism, regardless of setting. Significant positive correlations were found between effect size and degree of voluntariness for any recidivism, \( r(129) = .21, p = .02 \), and general recidivism, \( r(106) = .24, p = .01 \), but not for specific recidivism, \( r(52) = .13, p > .05 \). These correlational analyses were repeated separately for custody and community settings. Significant positive correlations were found between unweighted effect sizes and degree of voluntariness in custody and community settings, respectively, for any recidivism, \( r(52) = .41, p = .003 \), \( r(68) = .24, p = .05 \), in custody only for general recidivism, \( r(47) = .43, p = .003 \), \( r(51) = .20, p > .05 \), and neither custody nor community for specific recidivism, \( r(20) = .22, p > .05 \), \( r(28) = .18, p > .05 \).
MANDATED, COERCED, AND VOLUNTARY TREATMENT EFFECT SIZE COMPARISONS

Preliminary analyses comparing mandated and nonmandated treatment resulted in a similar pattern of results as comparisons among mandated, coerced, and voluntary treatment. In addition, only eight mandated treatment studies were nonmandated coerced and separating coerced treatment from voluntary and mandated treatment would be more informative than analysis of mandated versus nonmandated treatment. Therefore, mandated versus nonmandated comparisons were not reported to avoid redundancy and to provide a more detailed analysis of the impact of levels of coercion and treatment effectiveness. To assess the impact of the most extreme levels of coercion, the 5-point coercion–voluntary scale was collapsed into three groups as follows: 1 (mandated involuntary), 2 to 4 (coerced), and 5 (nonmandated or voluntary). Table 2 reports the overall, custody and community weighted (including and excluding the outlier), and unweighted effect sizes of mandated, coerced, and voluntary treatment of any, general, and specific recidivism.

All of the overall effect sizes were significant for the coerced and voluntary groups in all weighted and unweighted analyses, $p < .05$. The overall effect sizes for mandated treatment were significant for any and specific recidivism, but not for general recidivism using weighted .05 (95% confidence interval = −.00 to .10) and unweighted .05 (95% confidence interval = −.01 to .12) effect sizes. No significant differences were found between the overall weighted effect sizes for the three levels of coercion.

Three one-way ANOVAs were conducted to compare the overall unweighted effect sizes of the three levels of coercion. The results found significant differences between the levels of coercion for any recidivism, $F(2, 126) = 3.75, p = .03$ and general recidivism, $F(2, 103) = 3.12, p = .05$, but not specific recidivism, $F(2, 49) = 1.62, p > .05$. Post hoc multiple comparison tests indicated that voluntary treatment ($M = 0.17, SD = 0.15$) was significantly greater than mandated treatment ($M = 0.08, SD = 0.19$) for any recidivism only. Although the coerced group did not significantly differ from either the mandated or the voluntary group, the effect size of the coerced group was consistently closer to the voluntary group than the mandated group.

When the three levels of coercion were analyzed by treatment setting, custody settings were only effective in voluntary treatment using weighted and unweighted effect sizes for any, general, and specific recidivism. Community settings were effective for all three levels of coercion using any, general, and specific recidivism for all weighted and unweighted effect sizes. The weighted effective size of mandated treatment in community settings was significantly greater than voluntary treatment in custodial settings when the outlier was included using any recidivism, .10 (95% confidence interval = .07 to .14) > .04 (95% confidence interval = .02 to .06), respectively, but did not significantly differ when the outlier was excluded.

The weighted effective size of coerced treatment in community settings was significantly greater than voluntary treatment in custodial settings when the outlier was included using any, .11 (95% confidence interval = .07 to .15) > .04 (95% confidence interval = .02 to .06), and general .10 (95% confidence interval = .07 to .14) > .04 (95% confidence interval = .02 to .06) recidivism, respectively, but did not significantly differ when the outlier was excluded.

Three two-way ANOVAs were conducted to compare the three levels of coercion using the unweighted effect sizes of all studies (any recidivism), studies reporting general recidivism, and studies reporting specific recidivism. Treatment coercion main effects were found in
TABLE 2: Mean Effect Size by Level of Coercion for Any, General, and Specific Outcomes

<table>
<thead>
<tr>
<th>Treatment Effect</th>
<th>Mandated</th>
<th>Coerced</th>
<th>Voluntary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k</td>
<td>N</td>
<td>R</td>
</tr>
<tr>
<td>Any recidivism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>38</td>
<td>16,056</td>
<td>.08</td>
</tr>
<tr>
<td>Outlier excluded</td>
<td>38</td>
<td>16,056</td>
<td>.08</td>
</tr>
<tr>
<td>Unweighted*</td>
<td>38</td>
<td>16,056</td>
<td>.08</td>
</tr>
<tr>
<td>Custody</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>10</td>
<td>2,446</td>
<td>-.03</td>
</tr>
<tr>
<td>Outlier excluded</td>
<td>10</td>
<td>2,446</td>
<td>-.03</td>
</tr>
<tr>
<td>Unweighted*</td>
<td>10</td>
<td>2,446</td>
<td>-.00</td>
</tr>
<tr>
<td>Community</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>27</td>
<td>13,418</td>
<td>.10</td>
</tr>
<tr>
<td>Unweighted</td>
<td>27</td>
<td>13,418</td>
<td>.11</td>
</tr>
<tr>
<td>General recidivism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>32</td>
<td>11,219</td>
<td>.05</td>
</tr>
<tr>
<td>Outlier excluded</td>
<td>32</td>
<td>11,219</td>
<td>.05</td>
</tr>
<tr>
<td>Unweighted*</td>
<td>32</td>
<td>11,219</td>
<td>.05</td>
</tr>
<tr>
<td>Custody</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>9</td>
<td>2,231</td>
<td>-.05</td>
</tr>
<tr>
<td>Outlier excluded</td>
<td>9</td>
<td>2,231</td>
<td>-.05</td>
</tr>
<tr>
<td>Unweighted*</td>
<td>9</td>
<td>2,231</td>
<td>-.02</td>
</tr>
<tr>
<td>Community</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>22</td>
<td>8,796</td>
<td>.08</td>
</tr>
<tr>
<td>Unweighted</td>
<td>22</td>
<td>8,796</td>
<td>.09</td>
</tr>
<tr>
<td>Specific recidivism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>10</td>
<td>5,898</td>
<td>.12</td>
</tr>
<tr>
<td>Unweighted</td>
<td>10</td>
<td>5,898</td>
<td>.12</td>
</tr>
<tr>
<td>Custody</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>3</td>
<td>721</td>
<td>-.01</td>
</tr>
<tr>
<td>Unweighted</td>
<td>3</td>
<td>721</td>
<td>.06</td>
</tr>
<tr>
<td>Community</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted</td>
<td>7</td>
<td>5,177</td>
<td>.14</td>
</tr>
<tr>
<td>Unweighted</td>
<td>7</td>
<td>5,177</td>
<td>.15</td>
</tr>
</tbody>
</table>

Note. k = number of effect sizes; N = sample size; CI = confidence interval about r. Means in the same row that do not share subscripts differ at p < .05 using post hoc multiple comparison tests.
*Significant one-way ANOVA, p ≤ .05.
custody and community settings, respectively, for any recidivism, $F(2, 49) = 3.67, p = .03$, $F(2, 65) = 2.58, p > .05$, custody only, for general recidivism, $F(2, 44) = 3.68, p = .03$, $F(2, 48) = 1.05, p > .05$, custody only, and neither custody nor community for specific recidivism, $F(2, 17) = .97, p > .05, F(2, 25) = 1.25, p > .05$. Post hoc multiple comparison tests revealed that the effect sizes of voluntary treatment was significantly greater than mandated treatment in custodial setting for any ($M = 0.14, SD = 0.10 > M = 0.00, SD = 0.24$) and general ($M = 0.13, SD = 0.11 > M = 0.02, SD = 0.25$) recidivism.

To determine whether differences among the three levels of coercion confounded the results, three two-by-three ANOVAS were conducted to determine whether effect sizes differed by treatment mandate and setting for each of the recidivism outcomes. When setting was analyzed using any recidivism, a main effect was found for treatment setting, $F(1, 114) = 11.95, p = .001$, and level of coercion, $F(2, 114) = 5.43, p = .006$ (see Figure 1). A treatment setting by level of coercion interaction was not found, $F(2, 114) = .10, p > .05$. The effect size for voluntary treatment ($M = 0.17, SD = 0.15$) was greater than that for mandated treatment ($M = 0.08, SD = 0.19$), and the effect size for community treatment ($M = 0.16, SD = 0.17$) was greater than that for custodial treatment ($M = 0.09, SD = 0.15$).

When setting was analyzed using general recidivism, a main effect was found for treatment setting, $F(1, 92) = 4.75, p = .03$, and level of coercion, $F(2, 92) = 3.64, p = .03$. A treatment setting by level of coercion interaction was not found, $F(2, 92) = .49, p > .05$. When setting was analyzed using specific recidivism, a main effect was found for treatment setting, $F(1, 42) = 4.47, p = .04$, but not for level of coercion, $F(2, 42) = 1.98, p > .05$, or the treatment setting by level of coercion interaction, $F(2, 42) = .06, p > .05$.

### Figure 1: Mean Unweighted Effect Sizes for Any Recidivism as a Function of Setting and Level of Coercion

Note. Community treatment was significantly different from custodial treatment, and voluntary treatment was significantly different from mandated treatment. The setting by level of coercion interaction was not significant.
Finally, because a number of studies in the offender literature have suggested that mandated treatment increases treatment retention (e.g., Farabee et al., 1998; Rempel & Destefano, 2001; Siddall & Conway, 1988), correlational analyses were conducted between the percentage of clients dropping out of treatment and the 5-point coercion–voluntary scale. In the 56 studies reporting data on treatment dropouts, results found a significant negative correlation between the coercion–voluntary scale and the percentage of treatment dropouts, \( r(56) = -0.31, p = 0.02 \), indicating that greater voluntariness was associated with fewer client dropouts. However, no relationship was found between level of coercion and treatment dropouts when custodial, \( r(22) = -0.09, p > 0.05 \), and community, \( r(28) = -0.15, p > 0.05 \), settings were separately analyzed. Furthermore, a partial correlation controlling for setting was not significant, \( r(47) = -0.13, p > 0.05 \). Treatment effect size was not related to the percentage of treatment dropouts for any, \( r(56) = -0.02, p > 0.05 \), general, \( r(48) = -0.05, p > 0.05 \), or specific recidivism, \( r(24) = -0.01, p > 0.05 \).

**DISCUSSION**

The research on mandated treatment programs for offenders has been complicated by many issues (see Seddon, 2007), one being definitional terminology. For example, the terms coerced, compulsory, mandated, involuntary, legal pressure, and criminal justice referral have all been used to discuss coerced treatment and sometimes even used interchangeably within the same article (Farabee et al., 1998). Not only has inconsistent terminology been used to define coerced treatment, but also the concept of mandated versus voluntary treatment covers a variety of formal and informal degrees of legal pressure (Prendergast, Farabee, et al., 2002).

The current investigation attempted to create a more precise measure of mandated and voluntary treatment that included various degrees of coercion rather than simply coding mandated and voluntary treatment in a dichotomous manner. It commenced with a review of the commentary provided in the offender treatment literature and was followed with the construction of a reliable coding scheme to rate the degree of treatment coercion on a 5-point coercion–voluntary scale. Correlational analyses were conducted on this most detailed scale, and comparative analyses were conducted on various degrees of coercion. Analyses were also conducted separately for treatment setting because it became apparent that the relationship between mandated–nonmandated treatment and treatment effect size was confounded by setting. Specifically, institutional programs were more voluntary, but community programs produced a larger mean treatment effect size.

Previous research on the effects of mandated and voluntary treatment has produced mixed and inconclusive results. For example, studies have found that mandated treatment is effective in reducing recidivism and increasing treatment retention (e.g., Farabee et al., 1998), that mandated treatment is less effective than nonmandated or voluntary treatment in retaining treatment participants (Harford et al., 1976), and that there is little or no difference between mandated and voluntary treatment (e.g., Wild, 1999). The current study attempted to provide more conclusive results by conducting a meta-analysis of a large number of studies using a variety of measures of coercion.

The results of this study partially support those studies that found no treatment effects for mandated programs (e.g., Rosenfeld, 1992), partially support reviews finding nonmandated or voluntary treatment superior to mandated treatment in treatment in outcome and
retention (e.g., Harford et al., 1976), and partially support studies finding no differences between mandated and nonmandated treatment programs (e.g., “Principles of Drug Addiction,” 2000). Our study also challenges a number of studies that concluded mandated treatment is effective and superior to nonmandated or voluntary treatment in outcome and retention (e.g., Farabee et al., 1998; Rempel & Destefano, 2001). But most important, the current study may provide an explanation for the inconclusive results of previous research.

**INTERPRETATION**

Both weighted (including and excluding the outlier) and unweighted effect sizes are presented in the current study, with several differences emerging among them. Reminiscent of the differences in results that are often found when comparing effectiveness studies to efficacy studies (e.g., Lipsey, 1999; McGuire, 2002), smaller studies tended to have larger effect sizes. Consequently, unweighted effect sizes tended to be larger than the weighted effect sizes, most notably in voluntary treatment excluding the outlier study. Larger effect sizes are needed to achieve significance in studies with small samples because of the decreased power associated with smaller samples (Rosenthal & DiMatteo, 2001). Moreover, studies that do not yield results that achieve statistical significance are less likely to be published (also known as the “file drawer problem”); studies with small sample sizes and large effect sizes are more likely to be published than studies with small sample sizes and small effect sizes, introducing possible selection bias into the sample. Moreover, the presence of one (voluntary) treatment study with an extremely large sample size, and an average to small effect size, further complicated our interpretation of the current findings. Consequently, it was not surprising to find the unweighted mean effect size was stronger than the weighted mean effect size for the total effect size and for voluntary treatment, on both any and general recidivism, and that the weighted mean effect sizes of the total and voluntary samples increased when the outlier was excluded. This finding also explains the significant differences found between mandated and voluntary treatment in the unweighted effect size analyses.

Most of the coerced interventions (35 of 43) were nonmandated treatment programs. This was one reason why the results reported analyses of mandated, coerced, and voluntary treatment rather than mandated versus nonmandated treatment. In fact, voluntary treatment was found to be effective regardless of treatment setting, whereas coerced treatment in custodial settings did not result in a treatment effect, similar to mandated treatment. Therefore, the findings of this meta-analysis indicate caution in providing treatment for what is most likely the most common treatment scenario in the criminal justice system, specifically conditions that are less than completely voluntary but are short of enforced treatment in terms of offenders’ participation. It appears that some element of coercion may adversely affect the outcome of voluntary treatment.

The largest effect size across the three measures of outcome and three levels of voluntariness (i.e., .24) was found using unweighted effect sizes when voluntary treatment was applied to specific recidivism. When the studies were analyzed according to treatment setting, the largest effect size (i.e., .29) was found using unweighted effect sizes when voluntary treatment was applied to specific recidivism in community settings. This compares favorably with the other kinds of therapeutic intervention with offenders and with other clientele (Lösel, 1995; McGuire, 2002; Pearson & Lipton, 1999). It is particularly encouraging to find such a strong treatment effect for what may prove to be the most vital kind of treatment services for offenders (e.g., sex offender treatment, substance abuse offender treatment).
An examination of the effect sizes for general and specific recidivism reveals a similar pattern with one important difference. First, the overall effect size is larger for specific recidivism than general recidivism. Second, for general recidivism, the effect sizes for coerced and voluntary treatment in the three-group comparison are virtually the same, although modest in size, whereas the effect size for mandated treatment is low and not significantly different from zero (i.e., indicating a nil effect of treatment). Third, for specific recidivism, there is a consistent and relatively steady increase in the unweighted effect size as one goes from mandated to coerced to voluntary treatment, although these increases are not significant, quite possibly because of a lack of statistical power or difference in the types of recidivism measures reported. Most impressive, however, the strongest effect sizes in this study were found for voluntary treatment (without coercion), and the weakest effect sizes were found for mandated treatment.

One important finding of the current meta-analysis concerns the lack of treatment effect that mandated treatment had on general recidivism and custodial treatment, whereas voluntary treatment displayed a reliable impact on all offenders’ recidivism regardless of treatment setting. Moreover, mandated treatment did have a significant effect on specific recidivism in spite of the small number of studies on which this mean effect size was based. As one might expect, studies that examined the treatment effect of mandated treatment on general recidivism included more studies with smaller effect sizes. On the other hand, nonmandated treatment consisted of more studies with a much larger sample size, providing more confidence for its effect on recidivism.

Several explanations may account for why voluntary treatment was effective regardless of treatment setting. Offenders in custodial settings mandated to attend treatment may perceive less personal choice to attend treatment than those in community settings. Therefore, custodial settings may increase the perception of forced treatment, whereas offenders mandated to treatment in community settings may perceive less force to attend treatment and perhaps more personal choice. According to the self-determination theory, intrinsic motivation is said to be fostered by personal autonomy, or behavior that is determined by oneself and under one’s own control rather than the control of external forces (Deci & Ryan, 1991).

OTHER CONSIDERATIONS

Some might contend that effect sizes tended to be lower in mandated treatment because mandated treatment participants are more likely to be at higher risk for recidivism than nonmandated treatment participants. However, the risk principle, although counterintuitive to some, is supported by considerable empirical evidence demonstrating that larger treatment effects are actually found with the higher risk offenders and in fact treatment can be iatrogenic with low-risk offenders (Andrews & Bonta, 2003). Because mandated and nonmandated treatment programs were not found to differ according to treatment quality, or the number of RNR principles the treatment adhered to, this explanation does not seem plausible. Furthermore, mandated studies were more likely to be conducted in community settings, and community settings were found to have greater effect sizes than custodial programs (Andrews, Zinger, et al., 1990). Therefore, if anything, mandated treatment would be expected to have an advantage in generating a positive treatment effect.

The present meta-analysis revealed that mandated treatment has no effect on recidivism when the program is administered in custodial settings. This implies that if offenders are
being required by courts to attend treatment in custody settings, the treatment is likely to have no effect and cost the criminal justice system and the courts both time and money. On a positive note, coerced treatment produced effect sizes similar to voluntary treatment both in custody and in the community. Further research should be conducted on methods of motivating offenders to attend treatment on their own and/or increasing choice and reward for attending treatment (McMurran, 2002). On the other hand, it should not come as a surprise that community treatment is effective regardless of coercion and that voluntary treatment is effective regardless of treatment setting, although in both cases the degree of impact varies.

The current meta-analysis of offender treatment differs from previous ones in various ways. Being more recent, it includes some research that was not available in previous meta-analyses. It is more selective than the previous meta-analyses in that studies were excluded if there was insufficient information to rate the intervention on the coercion–voluntary dimension. Yet the mean unweighted effect size for all studies on any recidivism (.14) is consistent with those found in other analyses of offender treatment (e.g., Andrews, Zinger, et al., 1990; McGuire, 2002). The fact that the current overall effect size was comparable to those of previous meta-analyses suggests that the selection process that was necessary to investigate the impact of participant voluntariness did not distort or unduly bias the collection of studies in terms of their overall impact on offender treatment outcome.

A review of the ethics of coercion and the legality of mandated treatment for offenders was beyond the scope of the current investigation. Clearly, many jurisdictions have introduced mandated treatment of offenders, whereby their participation is required by law, particularly in the area of domestic violence treatment (Rosenbaum & Geffner, 2002). Many researchers, clinicians, and civil libertarians have questioned the wisdom and ethics of such services. One of the ethical arguments pertains to the provision of potentially ineffectual interventions. The current study, therefore, addressed not only the efficacy of mandated treatment but also, indirectly, the ethics of these services insofar as it may be deemed unethical to insist that offenders participate in interventions that have not been demonstrated to be effective. Required participation in offender treatment that targets general offending may be such a case. On the other hand, it is clear that treatment designed to address specialized kinds of outcome can have a positive effect on the desired outcomes and may not be considered unethical, at least in terms of requiring an offender to participate in unproven treatment.

LIMITATIONS

As is the case with any meta-analytic treatment study, the use of comparisons between groups of studies that share specific characteristics—in this case, the level of coercion—constitutes a correlational, as opposed to an experimental, design. In other words, the degree of voluntariness of offenders in treatment was correlated with the magnitude of the unweighted treatment effect for most of the analyses. But the specific reason for this positive correlation remains in question. It is unclear whether the difference in effect size was because of the act of volunteering or other client characteristics (e.g., motivation) or treatment characteristics (e.g., the intensity of the programs being offered). Another possibility may be that treatment facilitators are more responsive to clients who have volunteered and/or are motivated to attend treatment, thereby improving outcomes.

The concepts of mandated and voluntary treatment are neither simple nor discrete. In the current study, the determination of the degree of coercion was dependent on the information
provided by the authors. Moreover, the motivation of each individual participant in the treatment programs was unknown. Some offenders who enter mandated treatment may do so voluntarily, regardless of their requirement to do so (Dutton, 1986; Taylor et al., 2001). Many offenders who enter voluntary treatment do so not only because of known systemic consequences imposed by the criminal justice system, including reduced security, parole, and other forms of early release, but also because of more subtle forms of coercion by family members and health care professionals (Polcin & Weisner, 1999). Furthermore, there may be inherent differences among types of programs targeting mostly mandated or voluntary treatment participants. Consequently, considerable effort was taken to consider the degree of coercion present in each treatment scenario. However, as a meta-analytic investigation, this was limited by the information provided by the authors and therefore generally done at the level of the individual study and not at the level of the individual offender.

The analyses conducted on any and general recidivism found significant differences among levels of coercion using unweighted effect sizes. Although there was an overall treatment effect on specific recidivism, there was no difference in the mean unweighted effect size for the different levels of coercion. In addition to there being truly no difference between the impact of different levels of coercion under these conditions of specific treatment, a statistical explanation is also possible. Because considerably fewer studies included in the current meta-analysis reported specific recidivism ($k = 52$), statistical power was considerably reduced in these comparisons, as evidenced by the large confidence intervals, particularly for mandated treatment ($k = 10$). Furthermore, the lack of significance found between different levels of coercive treatment when specific treatment was analyzed may be because of the reduced power associated with the small sample of studies, particularly when custody ($k = 3$) and community ($k = 7$) treatment were separately analyzed. Lack of power because of small sample sizes may have also caused greater levels of coercion associated with greater rates of treatment attrition overall, but not when custody and community treatment were separately analyzed. Further research should be conducted on the effects of specific versus general recidivism, and researchers should be encouraged to report both general and specific recidivism.

CONCLUSION

The practice of requiring offenders to participate in treatment is one of the most controversial aspects of service provision to offenders, in terms of both effectiveness and ethics (Day et al., 2004). The current meta-analysis was limited to an investigation of the impact of different levels of coercion to attend treatment on offender recidivism and attrition. Most notably, the study found that mandated treatment was ineffective, particularly when the treatment was located in custodial settings, whereas voluntary treatment produced significant treatment effect sizes regardless of setting.

NOTES

1. The inclusion of a diverse set of offender treatment programs was deemed appropriate to avoid possible confounds due to selection bias.

2. An earlier version of the coercion–voluntary scale included six levels, where a score of 3 was mandated willing (offenders are mandated to participate in treatment, but there is evidence that they also took the treatment willingly or would have if it was not mandated); this level was excluded as none of the 129 studies included in the meta-analysis fit this description.
3. One study (Marshall, Eccles, & Barbaree, 1991) contributed two effect sizes that were calculated from the outcomes of two independent treatment groups.

4. Examples of the types of recidivism categorized as other include contact with any law enforcement agency, no further commitment (e.g., supervision, jail, prison), and rearrainment.

5. The outlier study was a transcendental meditation program with a treatment group of 220 and a control group of 32,270 (Bleick & Abrams, 1987). It was a voluntary program with an effect size of .02.

6. Treatment programs delivered in both custody and community settings \((n = 8)\) were excluded from analyses conducted on community and custody settings.

7. Effectiveness studies, which are conducted in practical settings where operative variables are less controlled, typically yield effect sizes lower than those found in efficacy studies, which are assessed purely in scientific terms, such as randomized controlled trials.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.


*Bourgon, G., & Armstrong, B. (2005).* Transferring the principles of effective treatment into a “real world” prison setting. *Criminal Justice and Behavior, 32*, 3-25.


Karen K. Parhar is a doctoral student in applied social psychology at the University of Saskatchewan. She has worked on various research projects with community and government agencies. Her interests include crime desistance, correctional treatment, and program evaluation.

J. Stephen Wormith is a professor in the Department of Psychology, University of Saskatchewan. He has worked with federal and provincial correctional agencies across Canada as a clinician, researcher, and administrator. His research interests include risk assessment of offenders and correctional intervention, and he is a coauthor of the Level of Service/Case Management Inventory.

Dena M. Derksen is a doctoral student in the Department of Psychology, Carleton University, and a researcher at Correctional Service of Canada, Ottawa, Ontario. Her research interests include the assessment and treatment of violent offenders, impulsivity, and program evaluation.

Adele M. Beauregard completed the bachelor of arts honours degree program, majoring in psychology at the University of Saskatchewan. She is currently working with the North East School Division as a counseling consultant and is pursuing her master of school and counseling psychology degree in the Department of Educational Psychology and Special Education at the University of Saskatchewan, majoring in psychology.