

---

**Research Report**

---

Safe and Humane Corrections  
Through Effective Treatment

This report is also available in French. Ce rapport est également disponible en français. Veuillez vous adresser à la direction de la recherche, Service Correctionnel du Canada, 340 avenue Laurier ouest, Ottawa (Ontario) K1A 0P9. Should additional copies be required they can be obtained from the Research Branch, Correctional Service of Canada, 340 Laurier Ave., West, Ottawa, Ontario, K1A 0P9.

**Safe and Humane Corrections  
Through Effective Treatment**

Sheila French

Paul Gendreau

Department of Psychology  
and  
Centre for Criminal Justice Studies  
University of New Brunswick at Saint John

**May 2003**

### **Authors' Note**

Sheila French is a graduate student in the department of Psychology, University of New Brunswick, Saint John. The second author is University Research Professor and Director of the Centre for Criminal Justice Studies, University of New Brunswick, Saint John. The authors can be contacted by email: [gendreau@unbsj.ca](mailto:gendreau@unbsj.ca).

This project was funded by the Addictions Research Centre, Correctional Service of Canada, Research Contract no. 21120-1-6607.

## Abstract

A meta-analysis was conducted examining the effects of prison-based treatment programs on inmates' misconduct rates. A total of 70 studies generated 103 effect sizes in this regard. Behavioral treatment programs produced considerably larger effect size estimates ( $r = .26$ ) compared to non-behavioral programs ( $r = .10$ ), educational/vocational programs ( $r = .02$ ), and an unspecified group of treatments ( $r = .02$ ). The same result applied when effect sizes were weighted ( $z^+$ ). Stronger research designs were associated with a larger reduction in misconducts in the case of  $r$  but not  $z^+$  values. Yet, treatment programs that targeted more criminogenic needs and were rated higher on therapeutic integrity, generated greater the reduction in misconducts for both  $r$  and  $z^+$  estimates. Generalization effects were also found: the greater the misconduct treatment effect, the larger the reductions in recidivism ( $r = .44$ ) reported for those prison programs that followed their offenders into the community. As a result, the reductions in prison misconduct carried over to the community. In conclusion, despite the fact that the studies in the database lacked a good deal of essential information, the results offer the strongest support for the policy recommendation that offering more and better treatment programs in prison is the preferred choice for maintaining safe and humane prison environments.

## Introduction

At one time, prisons were truly closed institutions. Extensive lockups (e.g. 16-24 hours daily) were not uncommon. Except for work details of various sorts, treatment programming and related activities were nominal. As a consequence, since movement and opportunity for social contact was restricted, the potential for inmate displays of antisocial behavior was minimized. By the 1970s, however, it certainly appeared to be the case that the modern prison, with its rehabilitative ethos and all the activities that entailed, was becoming difficult to manage (Gendreau, Tellier, & Wormith, 1985). Shortly thereafter, in response to these developments, a prison management literature began to emerge (e.g., Dilulio, 1987) which is now voluminous (Gendreau & Keyes, 2001).

A major theme in this body of knowledge has concerned itself with preventing prison antisocial behavior, which is typically operationalized in terms of prison misconducts (see Gendreau, Goggin, & Law, 1997). The general consensus in the literature nowadays is that in recent times disruptive prison misconduct behaviors have reached a serious level and the situation may be worsening (Armstrong, 2002; Byrne & Brewster, 1993; Gendreau, 2003; Walrath, 2001). While the evidence is persuasive that reducing the prevalence and incidence of misconduct behavior in prisons has enormous savings in cost and human terms and that measures should be taken (Gendreau & Keyes, 2001; Lovell & Jemelka, 1996), there has been differing views as to how to best accomplish this task.

Some “get tough” advocates of prison reform (Corcoran, 1993; Nossiter, 1994)) have suggested a return to “no frills” prisons (e.g., Pelican Bay, CA; see Haney, 2003) which feature, among a variety of things, less TV, recreation, visits, internal traffic, more use of solitary, more gun coverage, and a return of the lash and chain gangs (cf. Gendreau, Goggin, Cullen, & Paparozzi, 2002). While there exist little data on this topic, it has been shown that severe punishments or restrictions on inmates have not been associated with meaningful reductions in inmates’ disruptive behavior (Bidna, 1975). Nevertheless, any evidence emanating from the Pelican Bay’s of the world in the future might demonstrate otherwise, given the trend to resort to extreme physical restraints on offenders. On the other hand, it is hard to conceive how extreme lockup conditions would ever encourage prosocial behavior in prisons or deter future criminal behavior. As to the latter point, recent research has unequivocally shown no support for this view, even during the “get tough” era 30-50 years ago when some prisons resembled the Pelican Bays<sup>1</sup> of today (Gendreau, Goggin, & Cullen, 1999).

The prison management literature has provided a wide range of other kinds of recommendations that may impact upon prison disruptive behavior (see Gendreau & Keyes, 2001). It should however be approached with caution, as

---

<sup>1</sup> Pelican Bay is a landmark development in American corrections (see Haney, 2003). Exemplars of this form of incarceration offer complete isolation (i.e., no social interaction whatsoever) of offenders within a 60-80 square foot cell. In some cases, correctional officers monitor offenders by camera and communicate only via intercom. Computerized locking and tracking systems govern inmate movement and further reduce human contact. Typically, inmates are allowed one hour or less per day of "recreation" time in a caged or walled-in yard area. The offender must travel outside of his cell or unit and restraints are applied prior to movement via openings in cell doors. Inmates are never in the presence of another person (even physicians and psychiatrists) without restraint. In some institutions, visits with family, physicians, or psychiatrists are accomplished using a videoconferencing. Severe restrictions are placed upon personal possessions and on access to prison amenities.

most of the literature is characterized by case studies and the wisdom of prison management experts and clinicians. What policy recommendations and “hard” evidence that comment on the validity of some of the recommendations are as follows:

1. *Reduce overcrowding*. There is a sizeable body of evidence that indicates that overcrowding, typically defined as spatial density, has not played a major role in prison misconduct behavior (Bonta & Gendreau, 1990; Gendreau, et al., 1997). At first glance, this appears to be a contentious result, but it is likely that the rate of inmate turnover, type of inmate management style, and program availability are factors that mediate the overcrowding relationship with misconducts. Gendreau et al. (1997) proposed that inmates’ perception of feeling overcrowded, not the spatial density of their living space, would be a better predictor of misconducts.
2. *“Good time”*. It has been suggested that prisons will become safer environments if inmates are given presumptive parole dates contingent upon good conduct (i.e., if good time is taken away; inmates cannot earn time off their sentence which is preferable strategy according to behavior modification principles, see Gendreau, 2003). Results of the few and admittedly dated studies in this area have not produced clear cut results on misconducts one way or the other (Emshoff & Davidson, 1987; Shafer, 1982; Stone-Meierhoeffer & Hoffman, 1982).
3. *Prison “structural/operational” issues*. Based on a minimal amount of evidence and quasi-experimental designs, there is tentative evidence that

three prison “structural/operational” variables may have some influence on prison misconducts. They are: the weather<sup>2</sup> (Haertzen, Buxton, Covi, & Richards, 1993), the use of podular designs whereby correctional officers manage cell blocks, i.e., mini prison units where there is more face to face contact and bottom-up decision making (Senese, 1997), and food services (Schoenthaler, 1983c). The latter is not a flippant observation; a good nutrition habit on the part of inmates has been shown to reduce prison infractions by about 10% (a few studies in this area involved strong designs, e.g., Gesch, Hammond, Hampson, & Eves, 2002).

4. *Effective classification.* Better classification methods is a frequently nominated proposal in the management literature. In fact, there is a reasonable amount of empirical support for this policy. Gendreau, et al. (1997) conducted a meta-analysis that reported on several hundred correlations between various offender and prison situational factors, and prison misconduct behavior. While the database had serious limitations, some useful relationships were reported. First, they found that situational predictors were associated with serious prison disruptions (e.g., frequent inmate turnover among high risk youths and sentence length)<sup>3</sup>. Secondly, they discovered that the dynamic and static predictors of prison misconducts and offender recidivism were almost identical. The convergence in these two risk prediction literatures means that at-risk

---

<sup>2</sup> Anyone who has the pleasure of working in the summer in a maximum security prison is a strong proponent of air-conditioning.

offenders for prison disruptions can be identified<sup>4</sup> and placed in programs that target their criminogenic needs, which will not only result in reduced prison offending but likely lower recidivism rates. This point nicely segues into the final and pre-eminent management policy recommendation, which is at the core of this paper's research investigation.

5. *Effective treatment.* Of the hundreds of management recommendations for reducing prison misconducts, the most favoured, by far, has been to increase treatment programming (Gendreau & Keyes, 2001). To our knowledge, Keyes (1996) provided the first quantitative synthesis of the "treatment of prison misconduct" outcome literature. He gathered 33 effect sizes categorized along the lines suggested by Andrews, et al. (1990). Fifteen of them were designated "appropriate" (e.g., radical behavioral, social learning, and cognitive behavioral treatments) and 18 were determined to be "inappropriate" (e.g., psychodynamic, non-directive, group milieu, and intermediate sanctions). The overall ( $N = 33$ ) effect size estimate was  $r = .06$ <sup>5</sup>, while the effect sizes for the 15 appropriate and 18 inappropriate categories were  $r = .17$  and  $r = -.02$ , respectively. Seventy-one percent of the time, appropriate treatments produced larger reductions in misconducts. Effect sizes were also greater for stronger designs and juvenile samples.

---

<sup>3</sup> The evidence for situational factors is in urgent need of replication. In addition, some of the data is reported in aggregate form, which inflates correlations (Freedman, Pisani, Purves, & Adhikari, 1991).

<sup>4</sup> CSC already employs a useful risk measure (CNIA) and there are others that could be used that perform well in predicting prison misconducts (LSI-R, HCR-20)

<sup>5</sup> These  $r$  values can be interpreted at face value (see Cullen and Gendreau, 2000; Rosenthal, 1991). This means that  $r = .06$  is equivalent to a 6% reduction in misconducts.

Subsequently, Morgan & Flora (2002) reported on 13 effect sizes for prison group psychotherapy programs. Our estimate of their effect size was  $r = .21$  for misconduct outcomes. Behavioral/cognitive approaches tended to produce better results, but a precise estimate was not available as the results reported in this comparison encompassed other outcomes. As with Keyes (1996), Morgan and Flora (2002) remarked on the vast amounts of missing information in the studies in their database.

The purpose of this investigation was twofold. The first and primary objective, also central to the interests of prison managers, was to confirm the existing findings from the quantitative summaries of the effects of treatment programming on prison misconducts and to provide a more precise estimate of the treatment effect sizes on a much expanded database. Treatment programs were categorized as appropriate, inappropriate and "other" (e.g., educational/vocational). The second goal was to assess whether the results of these studies generalize to reductions in recidivism in the community. The latter issue is an important one, as it has been asserted that misconduct behavior in prison is a proxy for antisocial behavior in the community (see Gendreau et al., 1997). Thus, those studies that are associated with a reduction in prison misconducts after programming should provide a similar effect on recidivism upon release. If the answer to this question is affirmative, then the role of treatment programs in prison will be shown to have viable long-term consequences and provide additional support to the effective correctional treatment literature, which has convincingly demonstrated that behaviorally-

oriented programs produce the greatest reductions in criminal behavior (Andrews & Bonta, 2003; Cullen & Gendreau, 2000). In addition, this study assessed the relationship of the various moderators on misconduct effect size magnitude. The moderators, identified in the literature as possible sources of variability, were design strength, the therapeutic integrity of the treatment program, the number of criminogenic needs targeted and experimenter involvement, among others (Andrews & Bonta, 2003).

## **Method and Procedure**

### ***Sample of Studies***

Using library abstracting services (PsycINFO, Medline, Academic Search Elite) and the ancestry method, a literature search was conducted for studies examining prison treatment programs/interventions and their relation to inmate misconduct which was defined as violent/serious misconduct, non-violent misconduct, unspecified misconduct, disciplinary infraction, or institutional adjustment measures. To be included, studies were required to employ a randomized or comparison control group and to contain sufficient data to calculate an effect size (i.e., Pearson  $r$ /phi coefficient) between the treatment and criterion. In each study, data from the largest sample, the longest follow-up period, and the most serious outcome was recorded. The levels of outcome in order of seriousness included violent misconduct, non-violent misconduct, misconduct unspecified, and institutional adjustment rating/scale indices. A study could contribute more than one effect size if the treatment and/or control groups were distinct as to their content/composition (see Andrews et al, 1990). Recidivism data was recorded when available. Appendix A lists the studies included in the meta-analysis with their treatment categorization, sample size, and effect size.

### ***Coding of Studies***

The coding guide consisted of 100 items. The general coding categories with examples of their sub-components were as follows:

1. Study/author characteristics (e.g., type of publication, author affiliation, publication year).
2. Institutional factors (e.g., security level, institution type, location).
3. Sample variables (e.g., adult/juvenile, race, gender).
4. Research design (e.g., type of design, attrition, sample size, length of follow-up).
5. Treatment descriptors (e.g., type of treatment, control groups, treatment dosage). Treatment type was coded into one of three categories: behavioral (e.g., radical behavioral, social learning, cognitive behavioral, or punishment), non-behavioral (e.g., non-directive therapy, psychodynamic, group milieu), educational/vocational, and a non-specified grouping.
6. Therapeutic integrity descriptors (e.g., director qualifications, staff training, responsivity assessment, risk principle, criminogenic needs targeted).
7. Effect size descriptors (e.g., type of outcome, calculated effect size).

The coding manual for this meta-analysis can be found in Appendix B.

### ***Effect Size Calculation***

Phi coefficients ( $\phi$ ) were calculated using contingency tables for those studies providing information with respect to the frequency or proportion of post-treatment misconduct (and recidivism, if applicable) for both experimental and control groups. Where statistics other than  $r$  were reported (i.e.,  $F$ ,  $t$ ,  $\chi^2$ ,  $p$ ), the appropriate formula for conversion to  $r$  was employed (Rosenthal, 1991). In cases where a report of non-significance or a  $p$  value greater than .05 was the

only significance information provided, a Pearson  $r$  of .00 was assigned. Finally, according to the procedures outlined in Hedges & Olkin (1985), weighted  $r$  values ( $Z^+$ ) that account for both the influence of sample size per effect and the number of effect sizes per category of interest (e.g., treatment type, level of therapeutic integrity, or number of criminogenic needs targeted) were calculated.

### ***Effect Size Magnitude***

The metrics used to estimate the magnitude of misconduct/institutional adjustment effect sizes were  $r$  and  $Z^+$  along with the associated 95% confidence intervals ( $CI$ ). Significance testing is not emphasized (see Gendreau, 2002; Schmidt, 1996); yet, for those readers wedded to “ $p < .05$ ”, the  $CI$ s reported in this study can be interpreted as significance tests if the  $CI$  includes zero.

### ***Effect Size Heterogeneity***

The influence of outliers for the comparison of treatment types was determined using the  $Q$  statistic (Rosenthal, 1991). For each effect size, a  $q$  value was calculated using the formula  $\left[ (n - 3) * (z_r - z^+)^2 \right]$ , where  $n$  is the total sample size per treatment type;  $z_r$  is the standardized  $r$  per effect size; and  $z^+$  is the weighted  $r$  for each treatment type. These  $q$  values were then summed for each type of treatment, yielding  $Q$ , which is an estimate of the heterogeneity of the effect sizes within each treatment category. To evaluate its significance, the  $Q$  for each treatment category was evaluated using the critical value of  $\chi^2$  with  $(k - 1)$  degrees of freedom. If significant heterogeneity was indicated, outlying effect sizes  $[(z_r)(n - 3)]$  positioned above or below the mean of the treatment category by two or more standard deviations were identified and removed.

This process was repeated until non-significance was achieved, or until the originally obtained Q was reduced by 50 % (Bonta, Law & Hanson, 1998).

### ***Common Language Effect Size Indicator***

McGraw and Wong's (1992) common language effect size indicator (*CL*) was used to determine the practical utility of each of the treatment types. The *CL* statistic converts an effect size into a probability that a score (i.e., an effect size) sampled from the distribution of one group will be larger than a score sampled from that of another. Note the *CL* statistic cannot be calculated for a  $z^+$  estimate of effect size, as this statistic does not produce a standard deviation.

### ***Fail Safe Estimation***

A fail safe estimate was employed to provide an index of how many previously uncovered or "file drawer" effect sizes would be required to alter the obtained results. An index of the number of effect sizes ( $r = .00$ ), needed for a treatment type of greater efficacy in the reduction of misconduct to approach that equal to one of lesser efficacy, was calculated using the formula

$[(k_B(r_B - r_A))]/(r_A - r_{B=0})$  (see Gendreau et al., 2002), where  $r_{B=0}$  indicates a null effect for the more efficacious treatment type.

As applied to this meta-analysis, we assume that the mean effect size Treatment Type A is .30 ( $k = 50$ ) and that of Treatment Type B is .35 ( $k = 40$ ). An estimate of the number of B treatments with  $r = 0$  to negate its supremacy using the above formula is seven. In other words, seven additional Treatment Type B effect sizes, each with a magnitude of  $r = .00$ , would have to be located to conclude that the two treatments were at parity.

## Results

### ***Effect Size/Study Characteristics***

Seventy studies produced 103 correlations between various types of treatment programs and prison misconduct.

The majority of studies provided little information as to their study characteristics. We report on study characteristics in terms of effect size percentages and for only those coding categories where information was available on 50% of the items. Within a coding category, unless otherwise noted, we report on sub-components where the frequency was 20% or greater.

Eighty-one percent of the effect sizes were reported in journal articles. The literature search began in 1952 and extended to 2002. The percentages of effect sizes produced by decades were: 1952-1960 (8%), 1961-1970 (18%), 1971-1980 (29%), 1981-1990 (25%), and 1991-2002 (18%). Eighty-two percent of the effect sizes originated from U. S. prisons, 73% were based on male samples, and 39% and 50 % of effect sizes were from adult and juvenile samples, respectively.

Seventy-seven percent of the effect sizes could be coded as belonging to a specific treatment modality. Thirty-eight percent of those were designated behavioral (e.g., 15% mixed, 12% cognitive behavioral, and 7% social learning) and 37% non-behavioral (e.g., 7% group milieu, 5% non-directive and diet). Thirty eight percent of effect sizes were from programs that targeted criminogenic needs. Thirty three percent of effect sizes were from programs rated higher (5 – 21) on a 37-item scale of program therapeutic integrity (an abridged version of

the *Correctional Program Assessment Inventory*, Gendreau & Andrews, 2001), while 35% came from studies rated lower (< 5).

Sample sizes ranged from 5 to 1478 for the treatment and comparison groups.

The percentage of effect sizes by length of follow-up was: less than 6 months 46%, 6 to 11 months 17%, and 12 months or longer 15%.

The frequency of missing information in terms of effect size percentages ranged from 51% to 100% for the following categories: a) prison characteristics (security, location, population, crowding, treatment orientation, prison climate, location of treatment programs); b) offender characteristics (race, risk level, misconduct history); c) staff characteristics (program director and staff qualifications, experience, education and training, therapeutic skills); d) treatment characteristics (specific responsivity, risk principle, and program manual); e) the criterion (i.e., type of misconducts and recidivism; and f) the precise metric used (i.e.,  $r$ ,  $t$ ,  $F$  values).

### ***Treatment Type: Effect on Misconduct***

A total of 103 effect sizes were cumulated involving 21,370 inmates. The mean  $r$  was .13 ( $CI$  .09 to .18), and after weighting by sample size and number of effect sizes, the weighted effect size ( $z^+$ ) was .14 ( $CI$  .13 to .15).

The mean effect sizes for misconduct by treatment type are summarized in Table 1 for all 103 effect sizes. For example, consider row one in Table 1, where there were 40 behavioral treatment effect sizes reported involving 5809 offenders. The mean  $r$  was .26 with a  $CI$  of .18 to .34, and after weighting by sample size and

the number of effect sizes, the weighted effect size ( $z^+$ ) was .39 with a *CI* of .36 to .42. By comparison, the mean  $r$  for the 31 non-behavioral treatment effect sizes ( $N = 7629$ ) was .10, with an associated  $z^+$  of .05. The 95% *CI*s about mean  $r$  and  $z^+$  for the behavioral group did not overlap with the *CI*s for the other treatment categories. According to the *CL* statistic, behavioral treatments produced greater reductions in misconduct 68% of the time versus non-behavioral treatments, and 77% and 92% of the time in the case of educational/vocational and unspecified treatments.

Fail safe analysis indicated that, in order for the behavioral treatment effect to be reduced to the effect size level for non-behavioral treatment, 64 additional behavioral treatment studies all with an  $r$  equal to .00 would need to be located. Testing for effect size heterogeneity using the Rosenthal (1991) procedure resulted in the eliminations of 3, 7, and 4 effect sizes for the behavioral, non-behavioral, and unspecified categories.

Table 2 depicts the effect size estimators for  $r$  and  $z^+$  with outliers eliminated. The results for each treatment category are markedly similar to those reported in Table 1. As a result, we did not remove outliers in future analyses (also see Hunter & Schmidt, 1990, for a critique of the practice of eliminating outliers based on significance testing)

Table 1

*Mean Effect Sizes for Misconduct by Treatment Type: All Effect Sizes<sup>a</sup>*

Treatment Type ( <i>k</i> )	<i>N</i>	<i>Mr(SD)</i>	<i>CI<sub>r</sub></i>	<i>Mz<sup>+</sup></i>	<i>CI<sub>z<sup>+</sup></sub></i>
1. Behavioral (40)	5809	.26 (.24)	.18 to .34	.39	.36 to .42
2. Non-Behavioral (31)	7629	.10 (.21)	.02 to .17	.05	.03 to .07
3. Educational/Vocational (8)	1117	.02 (.19)	-.14 to .18	.00	-.06 to .06
4. Unspecified (24)	6815	.02 (.14)	-.03 to .08	.06	.04 to .08

*Note.* *K* = effect sizes per treatment type; *N* = offenders per treatment type; *Mr(SD)* = mean Pearson *r* and standard deviation between treatment type and recidivism; *CI<sub>r</sub>* = 95% confidence interval about mean Pearson *r*; *z<sup>+</sup>* = weighted *r* value that accounts for both the influence of sample size per effect and the number of effect sizes per treatment type; *CI<sub>z<sup>+</sup></sub>* = 95% confidence interval about *z<sup>+</sup>*.

<sup>a</sup> Common language effect size indicators for mean *r* values; behavioral vs. non-behavioral treatment 68%; behavioral vs. educational/vocational strategy 77%; behavioral vs. unspecified treatment strategy 92%.

Table 2

*Mean Effect Sizes for Misconduct by Treatment Type: Outliers Eliminated*

Treatment Type ( <i>k</i> )	<i>N</i>	<i>Mr(SD)</i>	<i>CI<sub>r</sub></i>	<i>Mz<sup>+</sup></i>	<i>CI<sub>z<sup>+</sup></sub></i>
1. Behavioral (37)	2266	.26 (.24)	.18 to .34	.39	.36 to .42
2. Non-Behavioral (24)	3894	.11 (.21)	.02 to .20	.05	.02 to .08
3. Educational/Vocational (8)	1117	.02 (.19)	-.14 to .18	.00	-.06 to .06
4. Unspecified (20)	3000	.02 (.12)	-.04 to .08	.06	.02 to .10

*Note.* *K* = effect sizes per treatment type; *N* = offenders per treatment type; *Mr(SD)* = mean Pearson *r* and standard deviation between treatment type and recidivism; *CI<sub>r</sub>* = 95% confidence interval about mean Pearson *r*; *z<sup>+</sup>* = weighted *r* value that accounts for both the influence of sample size per effect and the number of effect sizes per treatment type; *CI<sub>z<sup>+</sup></sub>* = 95% confidence interval about *z<sup>+</sup>*.

### ***Design Strength***

Mean effect sizes after allowing for design strength as a moderator are outlined in Table 3. Inspection of Table 3 indicates that stronger designs were associated with larger effect sizes for behavioral and non-behavioral treatment for  $r$ , but not when sample sizes were weighted. There was considerable variation in educational/vocational  $r$  and  $z^+$  effect size estimates, but there were very few effect sizes involved in the analysis.

Table 3

*Mean Effect Sizes for Misconduct by Treatment Type: Strong<sup>a</sup> and Weak<sup>b</sup> Experimental Designs*

	<i>N</i>	<i>Mr(SD)</i>	<i>CI<sub>r</sub></i>	<i>Mz<sup>+</sup></i>	<i>CI<sub>z<sup>+</sup></sub></i>
Strong Design					
1. Behavioral (12)	429	.36 (.25)	.20 to .52	.40	.29 to .49
2. Non-Behavioral (13)	1747	.15 (.16)	.06 to .25	.07	.02 to .12
3. Educational/Vocational (2)	224	-.22 (.11)	-.1.18 to .73	-.19	-.32 to -.06
4. Unspecified (13)	2343	.02 (.05)	-.01 to .05	.03	-.01 to .07
Weak Design					
1. Behavioral (28)	5380	.21 (.23)	.12 to .30	.39	.35 to .41
2. Non-Behavioral (18)	5882	.06 (.23)	-.06 to .17	.04	.01 to .07
3. Educational/Vocational (6)	893	-.11 (.13)	-.03 to .24	.05	-.02 to .12
4. Unspecified (11)	4472	.03 (.20)	-.10 to .17	.08	.04 to .10

*Note.* *K* = effect sizes per treatment type; *N* = offenders per treatment type; *Mr(SD)* = mean Pearson *r* and standard deviation between treatment type and recidivism; *CI<sub>r</sub>* = 95% confidence interval about mean Pearson *r*; *z<sup>+</sup>* = weighted *r* value that accounts for both the influence of sample size per effect and the number of effect sizes per treatment type; *CI<sub>z<sup>+</sup></sub>* = 95% confidence interval about *z<sup>+</sup>*.

<sup>a</sup> Strong experimental designs were those designs employing randomization of participants to treatment and control or where a treatment group was matched to a comparison group on five or more criminal risk factors. <sup>b</sup> The criteria for classification as a weak experimental design were as follows: > 20 % attrition in experimental or control group; comparison group matched on fewer than five criminal risk factors; or studies calculating postdictive correlational analyses

### ***Therapeutic Integrity***

Analysis of the relationship of therapeutic integrity with misconducts was conducted using the following categorizations: treatment strategies receiving a score lower than five on the therapeutic integrity score variable were designated as having a lower level of therapeutic integrity, while strategies scoring five or higher were designated as rating higher. The mean effect for those treatments with higher therapeutic integrity ( $k = 34$ ,  $r = .24$ ,  $CI = .15$  to  $.33$ ) was superior to that resulting from treatments ( $k = 36$ ,  $r = .12$ ,  $CI = .04$  to  $.20$ ) with lower levels of therapeutic integrity. The higher level of therapeutic integrity produced larger effect sizes 63% of the time.

### ***Comparison of Criminogenic Needs Targeted***

There were a total of 81 effect sizes, involving 13,309 inmates, for which the number of criminogenic needs targeted could be determined. Table 4 contains the results.

For example, reading across the first row, it can be seen that there were 24 effect sizes where three to eight criminogenic needs were targeted ( $N = 4586$ ). The mean  $r$  was  $.29$ , while weighting by sample size substantially increased the mean effect size ( $z^+ = .47$ ). By comparison, those studies targeting one to two or no criminogenic needs produced lower effect sizes.  $CI$ s for these latter two categories overlapped. The  $CL$  index favored those studies targeting three to eight criminogenic needs over those targeting one to two 54% of the time, and those targeting no criminogenic needs 73% of the time.

Table 4

*Mean Effect Sizes for Number of Criminogenic Needs Targeted<sup>a</sup>*

Criminogenic Needs ( <i>k</i> )	<i>N</i>	<i>Mr</i> ( <i>SD</i> )	<i>CI<sub>r</sub></i>	<i>Mz<sup>+</sup></i>	<i>CI<sub>z<sup>+</sup></sub></i>
1. Three to eight (24)	4586	.29 (.23)	.20 to .39	.47	.44 to .50
2. One to two (40)	4289	.15 (.23)	.08 to .23	.05	.02 to .08
3. Zero criminogenic needs <sup>b</sup> (17)	4434	.06 (.17)	-.03 to .14	.04	.01 to .07

*Note.* *k* = effect sizes per category; *N* = offenders per category; *Mr*(*SD*) = mean Pearson *r* and standard deviation between category and misconduct; *CI<sub>r</sub>* = 95% confidence interval about mean Pearson *r*; *z<sup>+</sup>* = weighted *r* value that accounts for both the influence of sample size per effect and the number of effect sizes per category; *CI<sub>z<sup>+</sup></sub>* = 95% confidence interval about *z<sup>+</sup>*.

<sup>a</sup> Common language effect size indicators for mean *r* values; three to eight vs. one to two 54%; three to eight vs. no criminogenic needs targeted 79%; one to two vs. no criminogenic needs targeted 63%.

<sup>b</sup> Those studies targeting no criminogenic needs targeted from zero to six non-criminogenic needs.

### ***Recidivism Outcomes***

The relationship between the degree to which treatment programs ( $k = 23$ ) produced reductions in misconducts and reductions in recidivism was assessed. The results indicated a positive correlation in this regard ( $r = .44$ ,  $CI = .10$  to  $.78$ ).

The largest reductions in recidivism were found for those misconduct treatment programs that targeted the most criminogenic needs ( $k = 7$ ). The effect size was  $r = .16$  ( $CI = .02$  to  $.30$ ) and  $z^+ = .12$  ( $CI = .02$  to  $.19$ ).

Turning to therapeutic integrity, 14 effect sizes ( $N = 1159$ ) were recorded where the treatments involved could be categorized as having a higher therapeutic integrity score. In other words, these treatments were found to have recorded favorable results on five or more of the 37 therapeutic integrity descriptors. The mean  $r$  for recidivism was  $.06$  ( $CI = -.07$  to  $.19$ ) with an associated  $z^+$  equal to  $.06$ . There were eight effect sizes scoring lower on the index of therapeutic integrity (fewer than five favourable responses out of 37 descriptors coded for) resulting in a mean  $r = .03$  ( $CI = -.19$  to  $.24$ ), with a  $z^+$  of  $.06$ . The  $CL$  index was 54% in favor of higher therapeutic integrity and reductions in recidivism.

### ***Supplementary Analyses***

Five contingency management studies, primarily featuring token economy procedures, were located that employed within-subject designs (Bornstein, et al. 1980; Colman & Baker, 1969; Cullen and Seddon, 1981; Ellis, 1993; Marquis, Gendreau, Cousins, & Wormith, 1974).

They were not included in the meta-analysis because they did not have control group comparisons nor could an  $r$  value effect size be calculated. Collectively, these studies generated large reductions in misconducts (up to five fold decreases depending on the metric used to describe effect size).

Other analyses of note were: experimenter being involved in the treatment program was associated with a lower effect size ( $r = -.18$ ,  $CI = -.37$  to  $.00$ ), treatment location separate from the rest of the prison was linked with a higher effect size ( $r = .23$ ,  $CI = .05$  to  $.41$ ), programs with adults were associated with larger effect size than for juveniles ( $r = .17$ ,  $CI = -.03$  to  $.37$ ), longer length of follow-up was related to a lower effect size ( $r = -.14$ ,  $CI = -.09$  to  $-.19$ ), and the correlation with the number of weeks in the program, our estimate of dosage, was associated with a greater effect size ( $r = .25$ ,  $CI = .07$  to  $.43$ ).

## **Discussion**

Previous meta-analysis on the effects of prison-based treatment programs on misconducts reported on a total of 46 effect sizes (Keyes, 1996; Morgan & Flora, 2002). The present study increased the database to 103 effect sizes and a sample of 21,000 inmates, which is considerable for the purposes of deriving useful policies for managing prisons in a more safe and humane manner.

Before discussing the results, however, it must be noted that there were serious limitations regarding the quality of the information provided, which, unfortunately, is consistent with our findings from previous meta-analyses on

prison research (Gendreau et al., 1997; 1999).<sup>6</sup> The lack of information on key variables is of concern, as it was impossible to uncover important moderator influences on effect size estimations. As noted in the results, virtually no information was reported on the context within which the treatment programs occurred. Essential inmate characteristics such as risk level and misconduct history were almost invariably unrecorded, as were therapeutic integrity factors. As to the latter, only 2% of effect sizes were associated with a treatment program that scored 50% on the truncated version of the *Correctional Program Assessment Inventory* (Gendreau & Andrews, 2001).<sup>7</sup> Some of the critical principles of what is known to work in offender rehabilitation (e.g., risk principle, targeting criminogenic need, cf. Andrews, Dowden, & Gendreau, 1999; Gendreau, 1996) were ignored, although one must be mindful of the fact that 80% of the effect sizes were produced in studies before there was widespread dissemination of the “what works” literature.

In about a quarter of the effect sizes, it was impossible to identify the nature of the treatment; similarly, the precise outcome metric was not available 52% of the time. For example, the reporting of “non-significance” served to underestimate treatment effects. In some cases, there were seemingly modest to strong effects to report, but the authors dismissed their findings because sample sizes were too small to produce the magical  $p < .05$  arbitrator of success. The only recourse, as a meta-analytic

---

<sup>6</sup> While the prison literature is an easy target for criticism, the correctional treatment literature is not that much better in our view, particularly when it comes to therapeutic integrity (e.g., Andrews, Dowden, & Gendreau, 1999).

coder, in these cases is to assign an  $r$  value equal to zero. This may seem harsh, but there are more than a few cases in the treatment literature where programs produce increases in recidivism (Andrews et al., 1990), thus a zero estimate is prudent.

Granted these important caveats, the results that emerged from the present study were remarkably consistent with the limited data from the previous two meta-analyses and, possibly more significantly, the correctional treatment literature. Keyes (1996) and Morgan and Flora (2002) reported that behavioral treatments produced reductions in misconducts of approximately  $r = .20$ . This meta-analysis confirmed their results and indicated that effects of such treatments were more robust ( $r = .26$ ,  $z+ = .39$ ). In our experience, it is unusual for weighting to vary outcomes much, but in this case, one very large sample and effect size outcome inflated the  $z+$  value (Prendergast, Farabee, & Cartier, 2001). While not a strong design (recall that strong designs were not associated with lower effect sizes), this study's program scored relatively high on therapeutic integrity. Two other small sample but large effect size reports (Leeman, Gibbs, & Fuller, 1993) came from the Arnold Goldstein group, who have a long and distinguished association with effective service delivery programs for juveniles (Goldstein et al., 1987).

The pragmatic consequences of these results should not be underestimated. Correlational values can be interpreted at face value (Cullen & Gendreau, 2000). By applying Rosenthal's (1991) widely used BESD statistic to

---

<sup>7</sup> This very low estimate is in part due to limited page space in journals and the fact that concern over therapeutic integrity in treatment is a somewhat recent issue (Gendreau, 1996).

the more conservative estimate of  $r = .26$ , misconducts can be reduced by about 26%, or if choosing the CI, a range of 18%-34%, a finding that any prison administrator would eagerly embrace as a powerful effect<sup>8</sup>. Another practical effect size index, the CL statistic, strongly supports the use of behavioral programs. Sixty-eight percent and 77% of the time, they produced a better result than non-behavioral and academic/vocational programs<sup>9</sup>. Secondly, non-behavioral programs, still common in many prisons (Gendreau, Goggin, & Smith, 2001), will never achieve the level of behavioral programs. This study found that the next generation of behavioral programs would have to perform abysmally (64 consecutive effect sizes of  $r = 0$  reported in the literature) to be reduced in effectiveness to that of the non-behavioral category. Finally, behavioral programs certainly seem to be more useful than some of the alternative management strategies noted in the introduction.

When it comes to comparing the results to the correctional treatment literature, the assumption is that misconduct behaviors correlate highly with anti-social behavior. Fully recognizing the vagaries of misconduct data in some situations (see Light, 1990), the consensus is that misconduct records have reasonable validity (Gendreau et al., 1997; Van Voorhis, 1994). Consider the following results. The correlation between reducing misconducts and eventual recidivism was  $r = .44$ , a most gratifying result that demonstrates the beneficial long-term consequences of one type of sound management policy in prison. In

---

<sup>8</sup> For those readers with experience working in prisons, a small reduction (e.g., 5%-10%) in misconduct behavior might be the difference between a prison in turmoil and one that is coping adequately with the usual pressures.

the correctional treatment literature, the correlation between five types of criminogenic needs targeted and recidivism ranged from  $r = .06$  to  $r = .39$ , with a mean  $r = .29$  (Andrews & Bonta, 2003). The correlation in this study between criminogenic needs targeted and each of misconducts and recidivism was  $r = .31$ . In addition, of the 23 effect sizes involved in this database, those few programs that targeted at least three to eight criminogenic needs produced a reduction in recidivism of  $r = .16$ , almost identical to the effect size reported ( $r = .17$ ) for prison-based programs ( $k = 30$ ) and recidivism that adhered to the principles of effective treatment (Andrews & Bonta, 2003). Finally, the correlation of treatment dosage and recidivism in the Andrews and Bonta (2003) database was  $r = .20$ ; in this study the correlation between dosage and reduction in misconduct was  $r = .25$ .

Before becoming too enthusiastic about the congruence in results, it should be recorded that comparisons across different types of databases are at best approximations and there were instances where this study produced a result at variance with the recidivism literature (e.g., experimenter involvement with the program was associated with lower instead of higher effect sizes).

In closing, this meta-analysis, based on a large sample of incarcerates, indicated that, on average, prison-based programs have a modest effect ( $r = .13$ ,  $CI = .09$  to  $.18$ ) on misconducts and that one type of treatment - behavioral programs – produce large effects ( $r = .26$ ,  $CI = .18$  to  $.34$ ) that may carry over into the community. This result, indeed, represents a significant advance in

---

<sup>9</sup> This is not derogatory towards academic/vocational programs. They serve a very useful purpose in and of themselves and can be useful adjuncts to treatment programs.

knowledge concerning the prison management literature. However, given the quality of some of the literature in this field, more primary studies are needed. It is our expectation that future primary studies that pay close attention to issues of therapeutic integrity for prison programs might well produce much better outcomes (e.g., approximately 35%-50%) than those reported herein.

## References

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

- \*Abrams, A. I., & Siegel, L. M. (1978). The transcendental meditation program and rehabilitation at Folsom State Prison. *Criminal Justice and Behavior*, 5, 3-30.
- Andrews, D. A., & Bonta, J. (1998). *The psychology of criminal conduct* (2<sup>nd</sup> ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., Dowden, D., & Gendreau, P. (1999). *Clinically relevant and psychologically informed approaches to reduced re-offending: A meta-analytic study of human service, risk, need, responsivity, and other concerns in justice contexts*. Carleton University, Ottawa, Canada.
- \*Andrews, D. A., & Young, J. G. (1974). Short-term structured group counselling and prison adjustment. *Canadian Journal of Criminology and Corrections*, 16, 5-13.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- \*Armstrong, T. A. (2002). The effect of environment on the behavior of youthful offenders: A randomized experiment. *Journal of Criminal Justice*, 30, 19-28.

- \*Barratt, E. S., Stanford, M. S., Felthous, A. R., & Kent, T. A. (1997). The effects of phenytoin on impulsive and premeditated aggression: A controlled study. *Journal of Clinical Psychopharmacology*, *17*, 341-349.
- Bidna, H. (1975). Effects of increased security on prison violence. *Journal of Criminal Justice*, *3*, 33-46.
- Bonta, J., & Gendreau, P. (1990). Reexamining the cruel and unusual punishment of prison life. *Law and Human Behavior*, *14*, 347-372.
- Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, *123*, 123-142.
- Bornstein, P. H., Rychtarik, R. G., McFall, M. E., Bridgewater, C. A., Guthrie, L., & Anton, B. (1980). Behaviorally specific report cards and self-determined reinforcements: A multiple baseline analysis of inmate officers. *Behavior Modification*, *4*, 71-81.
- \*Brill, R. (1978). Implications of the conceptual level matching model for treatment of delinquents. *Journal of Research in Crime and Delinquency*, *15*, 229-246.
- Byrne, J., & Brewster, M. (1993). Choosing the future of American corrections: Punishment of reform? *Federal Probation*, *57*, 3-9.
- \*Cavior, H. E., & Schmidt, A. (1978). A test of the effectiveness of a differential treatment strategy at the Robert F. Kennedy Center. *Criminal Justice and Behavior*, *5*, 131-139.

- \*Clear, T. R., & Myhre, M. (1995). A study of religion in prison. *The IARCA Journal*, 6, 20-25.
- Colman, A.D., & Baker, S. L. (1969). Utilization of an operant conditioning model for the treatment of character and behavior disorders in a military setting. *American Journal of Psychiatry*, 125, 1395.
- Corcoran, L. (1993). Comfortable prisons push recidivism rates. *Peacekeeper*, 2, 7-8.
- \*Craft, M., Stephenson, G., & Granger, C. (1964). A controlled trial of authoritarian and self-governing regimes with adolescent psychopaths. *Journal of Orthopsychiatry*, 34, 543-554.
- \*Cullen, E. (1987). Group-based treatments for serious institutional offending. In B. J. McGurk, D. M. Thornton, & M. Williams (Eds.), *Applying psychology to imprisonment: Theory and practice* (pp. 315-327). London: Her Majesty's Stationery Office Books.
- Cullen, F., & Gendreau, P. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In J. Horney (Ed.), *Criminal Justice 2000* (pp.109-175). Washington, DC: U.S. Department of Justice, NIJ.
- Cullen, J. E., & Seddon, J. W. (1981). The application of a behavioral regime to disturbed young offenders. *Personality and Individual Differences*, 2, 282-292.
- \*Davis, R. (1988). Education and the impact of the family reunion program in a maximum security prison. *Journal of Offender Counseling, Services, and Rehabilitation*, 12, 153-159.

- Dilulio Jr., J. J. (1987). *Governing prisons: A comparative study of correctional management*. NY: Free Press.
- \*Eisenberg, L., Lachman, R., & Moling, P. A. (1963). A psychopharmacologic experiment in a training school for delinquent boys: Methods, problems, findings. *American Journal of Orthopsychiatry*, 33, 431-447.
- Ellis, J. (1993). Security officer's role in reducing inmate problem behaviors. *Journal of Offender Rehabilitation*, 20, 61-72.
- \*Empey, L., & Lubeck, S. (1971). *The Silverlake experiment*. Chicago: Aldine.
- Emshoff, J. G., & Davidson, W. S. (1987). The effect of "good time" credit on inmate behavior. *Criminal Justice and Behavior*, 14, 335-351.
- \*Feder, B. (1962). Limited goals in short-term group psychotherapy with institutionalized delinquent adolescent boys. *International Journal of Group Psychotherapy*, 12, 503-507.
- \*Fox, V. (1954). The effect of counseling on adjustment in prison. *Social Forces*, 32, 285-289.
- Freedman, D. R., Pisani, R., Purves, R., & Adhikari, A. (1991). *Statistics* (2<sup>nd</sup> ed.). NY: W. W. Norton.
- \*Friedland, D. M. (1960). *Group counseling as a factor in reducing runaway behavior from an open treatment institution for delinquent and pre-delinquent boys*. Unpublished doctoral dissertation, New York University, NY.

- Gendreau, P. (1996). The principles of effective intervention with offenders. In A. T. Harland (Ed.), *Choosing correctional options that work* (pp. 117-130). London: Sage.
- Gendreau, P. (2002). We must do a better job of cumulating knowledge. *Canadian Psychology, 43*, 205-210.
- Gendreau, P. (2003). Encouraging “good” behavior in prisons. (Research Contract No. 2112026620), Ottawa, Canada: Correctional Service of Canada.
- Gendreau, P., & Andrews, D. A. (2001). *Correctional program assessment inventory – 2000*. Ottawa, Canada: T3 Associates.
- Gendreau, P., Goggin, C., & Cullen, F. (1999). The effects of prison sentences on recidivism. Report to the Corrections Research and Development and Aboriginal Policy Branch (Cat. #J42-87/1999E), Solicitor General of Canada. Ottawa, Canada: Public Works & Government Services Canada.
- Gendreau, P., Goggin, C., Cullen, F., & Pappozzi, M. (2002). The common sense revolution and correctional policy. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programs and policies to reduce reoffending* (pp. 359-386). Chichester: John Wiley & Sons.
- Gendreau, P., Goggin, C., & Law, M. A. (1997). Predicting prison misconducts. *Criminal Justice and Behavior, 24*, 414-431.

- Gendreau, P., Goggin, C., & Smith, P. (2001). Implementation guidelines for correctional programs in the 'real world.' In G. A. Bernfeld, D. P. Farrington, & A. W. Leschied (Eds.), *Offender rehabilitation in practice* (pp. 247-268). Chichester, England: Wiley.
- Gendreau, P., & Keyes, D. (2001). Making prisons safer and more humane environments. *Canadian Journal of Criminology*, 43, 123-130.
- \*Gendreau, P., Ross, R. R., & Izzo, R. (1985). Institutional misconduct: The effects of the UVIC program at Matsqui penitentiary. *Canadian Journal of Criminology*, 27, 209-217.
- Gendreau, P., Tellier, M. C., & Wormith, J. S. (1985). Protective custody: The emerging crisis within our prisons. *Federal Probation*, 44, 55-63.
- \*Gesch, C. B., Hammond, S. M., Hampson, S. E., & Eves, A. (2002). Influence of supplementary vitamins, minerals and essential fatty acids on the antisocial behaviour of young prisoners: Randomized, placebo-controlled trial. *British Journal of Psychiatry*, 181, 22-28.
- \*Goldenberg, F., & Cowden, J. F. (1977). An evaluation of intensive group psychotherapy with male offenders in isolation units. *Corrective and Social Psychiatry and Journal of Behavior Technology, Methods and Therapy*, 23, 68-72.
- \*Goldstein, A. P., Glick, B. Reiner, S., Zimmerman, D. & Coultry, M. (1987). *Aggression replacement training: A comprehensive intervention for aggressive youth*. Champaign, IL: Research Press.

- \*Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: 2. Intervention. *Developmental Psychology*, 26, 269-277.
- \*Guttman, E. S. (1963). *MMPI measured changes in treated and untreated youth authority wards judged in need of psychiatric treatment* (Research report No. 25). CA: California Youth Authority.
- Haertzen, C., Buxton, K., Covi, L., & Richards, H. (1993). Seasonal changes in rule infractions among prisoners: A preliminary test of the temperature-aggression hypothesis. *Psychological Reports*, 72, 195-200.
- Haney, C. (2003). Mental health issues in long-term solitary and ‘supermax’ confinement. *Crime and Delinquency*, 49, 124-156.
- Hedges, L. V., & Olkin, I. (1985). *Statistical Methods for Analysis*. Toronto, Canada: Academic Press.
- \*Hollin, C. R. & Courtney, S. A. (1983). A skills training approach to the reduction of institutional offending. *Personality and Individual Differences*, 4, 257-264.
- \*Hollin, C. R., & Henderson, M. (1981). The effects of social skills training on incarcerated delinquent adolescents. *Journal of Behavioral Social Work and Abstracts*, 1, 145-155.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publications.

- \*Ingram, G. L., Gerard, R., Quay, H. C., & Levinson, R. B. (1970). An experimental program for the psychopathic delinquent: Looking in the "correctional wastebasket." *Journal of Research in Crime and Delinquency*, 7, 24-30.
- \*Jacobson, F., & Magee, E. (1965). *Englewood project: Re-education: A radical correction of incarcerated delinquents*. Englewood, CO.
- \*Johnson, B. R. (1987). Religiosity and institutional deviance: The impact of religious variables upon inmate adjustment. *Criminal Justice Review*, 12, 21-30.
- \*Kassebaum, G. G., Ward, D. A., & Wilner, D. M. *The man in the pastel prison: A study of correctional treatment outcome*. NY: John Wiley & Sons.
- \* Kessemeier, L. A. (1966, March-April). Does group counseling pay its way? *Correctional Review*, 26-30.
- Keyes, D. (1996). *Preventing prison misconduct behavior: A quantitative review of the literature*. Unpublished honor's thesis, University of New Brunswick at Saint John, Saint John, New Brunswick, Canada.
- \*Kirigin, K. A., Braukmann, C. J., Atwater, J., & Wolf, M. M. (1982). An evaluation of teaching-family (Achievement Place) group homes for juvenile offenders. *Journal of Applied Behavior Analysis*, 15, 1-16.
- \*Langenbach, M., North, M. Y., Aagaard, L., & Chown, W. (1990). Televised instruction in Oklahoma prisons: A study of recidivism and disciplinary actions. *Journal of Correctional Education*, 41, 87-94.

- \*Leak, G. K. (1980). Effects of highly structured versus nondirective group counseling approaches on personality and behavioral measures of adjustment in incarcerated felons. *Journal of Counseling Psychology, 27*, 520-523.
- \*Leeman, L. W., Gibbs, J. C., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. *Aggressive Behavior, 19*, 281-292.
- \*Levinson, R. B., & Kitchener, H. L. (1964). *Demonstration Counseling Project*. Washington, DC: National Training School for Boys.
- \*Liau, A. K., (1999). Evaluation of the peer helping component of a group treatment program for antisocial youth. (Doctoral dissertation, Graduate School of the Ohio State University, 1999). *Dissertation Abstracts International, 60(5-B)*, 2387.
- Light, S. C. (1990). The severity of assaults on prison officers: A contextual study. *Social Science Quarterly, 71*, 267-284.
- \*Lipton, D., Martinson, R., & Wilks, J. (1975). *The effectiveness of correctional treatment: A survey of treatment evaluation studies*. NY: Praeger.
- \*Longhurst, J. E., & Mazer, G. E. (1988). The effects of a low glycemic diet on antisocial behavior in juvenile offenders. *International Journal of Biosocial Research, 10*, 123-136.

- Lovell, D., & Jemelke, R. (1996). When inmates misbehave: The costs of discipline. *Prison Journal, 76*, 165-179.
- \*MacKenzie, D. L., & Shaw, J. (1990). Inmate adjustment and change during shock incarceration: The impact of correctional boot camps. *Justice Quarterly, 7*, 125-150.
- Marquis, H. A., Gendreau, P., Cousins, L., and Wormith, J. S. (1974). Application and social implications of a token economy in an adult institution. Paper presented at the Canadian Association for the Advancement of Research in Criminology and Criminal Justice, Montreal, Canada.
- \*McDougall, C. (1990). Anger control training with young offenders. In S. Duguid (Ed.), *The yearbook of correctional education* (pp. 319-324). Burnaby, Canada: Simon Fraser University.
- McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin, 111*, 361-365.
- \*Medve, W. J. (1961). The rehabilitative aspects of team sports in a reformatory. *Journal of Correctional Education, 13*, 4-11.
- Morgan, R. D., & Flora, D. B. (2002). Group psychotherapy with incarcerated offenders: A research synthesis. *Group Dynamics, 6*, 203-218.
- \*Morgan, R. D., Winterowd, C. L., & Fuqua, D. R. (1999). The efficacy of an integrated theoretical approach to group psychotherapy for male inmates. *Journal of Contemporary Psychotherapy, 29*, 203-222.

- \*Morrissey, C. (1997). A multimodal approach to controlling inpatient assaultiveness among incarcerated juveniles. *Journal of Offender Rehabilitation, 25*, 31-42.
- \*Moss, C. S., Hosford, R. E., Anderson, W. R., & Petracca, M. (1977). Personality variables of blacks participating in a prison riot. *Journal of Consulting and Clinical Psychology, 45*, 505-512.
- \*New York City Department of Correction. (1960, January). *A preliminary evaluation of the relationship between group psychotherapy and the adjustment of adolescent inmates in a short-term penal institution* (Research report). NY: Diagnostic staff of Rikers Island Penitentiary.
- \*Newburger, H. M. (1952). *The effect of group therapy upon certain aspects of the behavior and attitudes of institutionalized delinquents*. Unpublished doctoral dissertation, New York University, NY.
- Nossiter, A. (1994, September 17). Making hard time harder: States cut jail TV and sports. *New York Times*, pp. A1, A10.
- \*Ollendick, T. H., & Hersen, M. (1979). Social skills training for juvenile delinquents. *Behavior Research and Therapy, 17*, 547-554.
- \*Prendergast, M., Farabee, D., & Cartier, J. (2001). The impact of in-prison therapeutic community programs on prison management. *Journal of Offender Rehabilitation, 32*, 63-78.
- \*Pugh, D. N., (1993). The effects of problem-solving ability and locus of control on prisoner adjustment. *International Journal of Offender Therapy and Comparative Criminology, 37*, 163-176.

- \*Queralt, J. P., Caballero, J., Casals, A., Navarro, J. C., & Serra, S. (1997). The treatment of aggressive prisoners: A closed regime programme. In S. Redondo, V. Garrido, J. Perez, & J. Barberet (Eds.), *Advances in psychology and law: International contributions* (442-451). NY: Walter de Gruyter.
- \*Roberts, R. E., Cheek, E. H., & Mumm, R. S. (1994). Group intervention and reading performance in a medium-security prison facility. *Journal of Offender Rehabilitation, 20*, 97-116.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- \*Ross, R. R., & McKay, H. B. (1976). A study of institutional treatment programs. *International Journal of Offender Therapy and Comparative Criminology, 20*, 165-173.
- \*Roth, L. H., Rosenberg, N., and Levinson, R. B. (1971). Prison adjustment of alcoholic felons. *Quarterly Journal of Studies on Alcohol, 32*, 382-392.
- \*Rudoff, A. (1960). *The effect of treatment in incarcerated young adult delinquents as measured by disciplinary history*. Unpublished master's thesis, University of Southern California.
- \*Sarason, I. G., & Ganzer, V. J. (1973). Modeling and group discussion in the rehabilitation of juvenile delinquents. *Journal Of Consulting Psychology, 20*, 442-449.

- \*Schlichter, K. J., & Horan, J. J. (1981). Effects of stress inoculation on the anger and aggression management skills of institutionalized juvenile delinquents. *Cognitive Therapy and Research*, 5, 359-365.
- Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training researchers. *Psychological Methods*, 1, 115-129.
- \*Schoenthaler, S. J. (1983a). The Northern California diet-behavior program: An empirical examination of 3,000 incarcerated juveniles in Stanislaus County Juvenile Hall. *International Journal of Biosocial Research*, 5, 99-106.
- \*Schoenthaler, S. J. (1983b). Diet and crime: An empirical examination of the value of nutrition in the control and treatment of incarcerated juvenile offenders. *International Journal of Biosocial Research*, 4, 25-39.
- Schoenthaler, S. J. (1983c). diet and delinquency: A multi-site replication. *International Journal of Biosocial Research*, 5, 70-78.
- \*Seckel, J. P. (1965). *Experiments in group counseling at two youth authority institutions*. (Research report No. 46). CA: California Youth Authority.
- Senese, J. D. (1997). Evaluating jail reform: A comparative analysis of podular/direct and linear jail inmate interactions. *Journal of Criminal Justice*, 25, 61-73.
- Shafer, N. E. (1982). Good time and prisoner misconduct: A preliminary examination. In N. Parisi (Ed.), *Coping with imprisonment* (pp. 147-158). Beverly Hills, CA: Sage.

- \*Snyder, R., & Sechrest, L. (1959). An experimental study of directive group therapy with defective delinquents. *American Journal of Mental Deficiency, 64*, 117-123.
- \*Sowles, R. C., & Gill, J. H. (1970). Institutional and community adjustment of delinquents following counseling. *Journal of Counseling and Clinical Psychology, 34*, 398-402.
- \*Stallone, T. M. (1993). The effects of psychodrama on inmates within a structured residential behavior modification program. *Journal of Group Psychotherapy, Psychodrama & Sociometry, 46*, 24-32.
- Stonemeierhoeffer, B., & Hoffman, P. B. (1982). The effects of presumptive parole dates on institutional behavior: A preliminary analysis. *Journal of Criminal Justice, 10*, 283-297.
- \*Sultan, F. E., Long, G. T., Kiefer, S. A., Schrum, D. M., Selby, J. W., & Calhoun, L. G. (1984). The female offender's adjustment to prison life: A comparison of psychodidactic and traditional supportive approaches to treatment. *Journal of Offender Counseling Services and Rehabilitation, 9*, 49-56.
- Toch, H. (1995). The long-term inmate as a long-term problem. In T. J. Flanagan (Ed.), *Long-term imprisonment* (pp. 245-248). Thousand Oaks, CA: Sage.
- Van Voorhis, P. (1994). Measuring prison disciplinary problems: A multiple indicators approach to understanding prison adjustment. *Justice Quarterly, 11*, 679-709.

- \*Walrath, Christine (2001). Evaluation of an inmate-run Alternatives to Violence Project: The impact of inmate-to-inmate intervention. *Journal of Interpersonal Violence, 16*, 697-711.
- \*Walters, G. D. (1999). Short-term outcome of inmates participating in the Lifestyle Change program. *Criminal Justice and Behavior, 26*, 322-337.
- \*Watt, B. D., & Howells, K. (1999). Skills training for aggression control: Evaluation of an anger management program for violent offenders. *Legal and Criminological Psychology, 4*, 285-300.
- \*Wolk, R. L. (1966). The relationship of group psychotherapy to institutional adjustment. *International Journal of Group Psychotherapy, 16*, 141-144.
- \*Wormith, J. S. (1984). Attitude and behavior change of correctional clientele: A three year follow-up. *Criminology, 22*, 595-618.
- \*Zivan, M. (1966). *Youth in trouble: A vocational approach*. Final report of a research and demonstration project, May 31, 1961 – August 31, 1966. Dobbs Ferry, NY: Children's Village.

## Appendix A

### *Details of Studies Included in the Meta-Analysis*

Author (Year)	Treatment Strategy	<u>N</u>	<u>r</u>
Abrams & Siegal (1978)	Non-Behavioral	40	.00
	Non-Behavioral	49	.00
Andrews & Young (1974)	Behavioral	26	.33
	Behavioral	21	.24
Armstrong (2002)	Behavioral	256	-.06
Barratt et al. (1997)	Non-Behavioral	60	.15
Brill (1978)	Unspecified	22	.42
	Unspecified	21	.00
Cavior & Schmidt (1978)	Non-Behavioral	198	-.25
	Unspecified	222	-.27
	Behavioral	214	-.21
	Behavioral	259	-.08
Clear & Myhre (1995)	Unspecified	769	.24
Craft et al. (1964)	Non-Behavioral	44	.15
Cullen (1987)	Behavioral	26	.28
Davis (1988)	Educational/Vocational	325	-.02
	Non-Behavioral	332	.08
	Unspecified	324	.00
	Unspecified	325	.16
Eisenberg et al. (1962)*	Non-Behavioral	42	.30

Author (Year)	Treatment Strategy	<u>N</u>	<u>r</u>
Empey & Lubeck (1971)	Non-Behavioral	261	.03
Feder (1962)	Unspecified	40	.00
	Unspecified	40	.00
	Unspecified	40	.00
	Unspecified	40	.00
Fox (1954)	Unspecified	174	.15
Friedland (1960)*	Unspecified	36	.00
Gendreau et al. (1985)	Educational/Vocational	144	.13
Gesch et al. (2002)	Non-Behavioral	172	.00
Goldenberg & Cowden (1977)	Non-Behavioral	177	.00
Goldstein et al. (1987) <sup>+</sup>	Behavioral	60	.41
	Behavioral	51	.00
Guerra & Slaby (1990) <sup>+</sup>	Behavioral	80	.59
Guttman (1963)*	Unspecified	215	.00
	Unspecified	123	.00
Hollin & Courtney (1983)	Behavioral	16	.00
Hollin & Henderson (1981)	Behavioral	10	.41
Ingram et al. (1970)	Behavioral	61	.00
Jacobson & Magee (1965)*	Educational/Vocational	60	-.30
	Educational/Vocational	164	-.15
Johnson (1987)	Non-Behavioral	782	.02

Author (Year)	Treatment Strategy	<u>N</u>	<u>r</u>
Kassebaum, et al. (in press)*	Unspecified	968	.00
		249	
Kessemeier (1966)*	Unspecified	9	.07
Kirigin et al. (1982)	Behavioral	68	.33
	Behavioral	124	.23
Langenbach et al. (1990)	Educational/Vocational	200	.00
Leak (1980)	Non-Behavioral	47	.31
Leeman et al. (1993)	Behavioral	39	.68
	Behavioral	38	.88
Levinson & Kitchener (1964)*	Non-Behavioral	225	.00
Liau (1999)	Behavioral	39	.35
Longhurst & Mazer (1988)	Non-Behavioral	140	.37
MacKenzie & Shaw (1990)	Non-Behavioral	68	-.24
McDougall (1990)	Behavioral	36	.46
Medve (1961)*	Non-Behavioral	399	.10
Morgan et al. (1999)	Behavioral	36	.26
Morrissey (1997)	Behavioral	77	.26
Moss et al. (1977)	Educational/Vocational	98	.28
	Educational/Vocational	98	.00
	Non-Behavioral	98	.00
	Unspecified	98	-.28
	Non-Behavioral	98	.00

Author (Year)	Treatment Strategy	<u>N</u>	<u>r</u>
	Unspecified	98	-.20
New York City (1960)*	Unspecified	120	.00
Newburger (1952)*	Unspecified	47	.00
	Unspecified	47	.00
Ollendick & Hersen (1979)	Behavioral	18	.46
	Behavioral	18	.46
	Non-Behavioral	18	.00
		295	
Prendergast et al. (2001)			.53
	Behavioral	6	
Pugh (1993)	Behavioral	168	.05
Queralt et al. (1997)	Educational/Vocational	28	.24
Roberts et al. (1994)	Non-Behavioral	28	.52
	Non-Behavioral	23	.68
Ross & McKay (1976)	Behavioral	30	-.36
	Behavioral	30	.36
	Behavioral	30	.36
	Behavioral	30	.47
Roth et al. (1971)	Non-Behavioral	100	-.09
Rudoff (1960)*	Non-Behavioral	534	.00
Sarason & Ganzer (1973)	Behavioral	95	.10
	Behavioral	101	.24
	Behavioral	108	.15

Author (Year)	Treatment Strategy	<u>N</u>	<u>r</u>
Schlichter & Horan (1981)	Behavioral	19	-.13
	Behavioral	18	-.21
	Non-Behavioral	17	.07
Schoenthaler (1983a)		200	
	Non-Behavioral	5	.04
	Non-Behavioral	573	.00
Schoenthaler (1983b)	Non-Behavioral	276	.24
Seckel (1965)*	Unspecified	487	.09
Snyder & Sechrest (1959)	Non-Behavioral	32	.29
	Non-Behavioral	32	.40
Sowles & Gill (1970)	Unspecified	45	.00
	Unspecified	15	.00
Stallone (1993)	Behavioral	44	.43
	Behavioral	44	.33
Sultan et al. (1984)	Behavioral	61	.00
Walrath (2001)	Behavioral	56	.47
Walters (1999)	Behavioral	373	.23
Watt & Howells (1999)	Behavioral	38	.00
Wolk (1966)	Non-Behavioral	619	.10
Wormith (1984)	Behavioral	35	.33
Zivan (1966)*	Non-Behavioral	140	.00

*Note.* In those studies where more than one treatment of the same category was employed (e.g., behavioral, Stallone, 1993), the behavioral treatments were different, or the behavioral treatment was compared to different types of control groups, or the study was conducted in multiple sites.

\* Denotes studies for which the original publication was not accessible. The effect sizes for these studies were derived from Lipton, Martinson, and Wilkes (1975).

+ Denotes two studies, one where there were 2 *df* in the numerator and another where the authors reported separate pre-post *t* values for the treatment and comparison groups. The control group under either condition did not improve over time.

## Appendix B

### *Coding Manual*

#### *Study/Author Descriptors*

*Reference:* Write a complete citation in APA format

1. *Study ID number.* Study identification numbers correspond to the reference numbers assigned to each study by the Library Master Software program. If an article contains two or more independent studies, i.e., if the article reports two or more independent outcomes using different samples, then the suffix a, b, c, etc. is affixed to the study ID to distinguish each study within the report. Each outcome is then coded separately.
2. *Type of publication.* Record the type of publication for the study.
  1. Book
  2. Journal article
  3. Book chapter
  4. Thesis/doctoral dissertation
  5. Technical report
  6. Conference paper
  7. Other (specify):
3. *Published paper.* Identify whether the article has been published.
  1. Yes
  2. No
4. *Publication year.* What was the publication year (last two digits, 99 if unknown)?
5. *Author affiliation.* What was the affiliation of the study author? In the case of multiple author studies, record the affiliation for the lead author only.
  1. Academic
  2. Government
  3. Private agency
  4. Mixed (specify):
  9. Information not available

6. *Author discipline.* What was the academic discipline of the author? In the case of multiple author studies, record the discipline of the lead author only.

1. Criminology
2. Psychology
3. Sociology
4. Other (specify):
9. Information not available

*Institutional Descriptors*

7. *Security level.* What was the security level of the institution in which the study was conducted? Institutions that were residential (e.g., juvenile group homes, borstals, etc.) are coded as minimum security unless author(s) specify otherwise.

1. Minimum
2. Medium
3. Maximum
9. Information not available

8. *Institution type.* Record the type of institution in which the study was conducted.

1. Federal
2. Provincial/State
3. County
4. Residential/Half-way house
5. Other (specify):
9. Information not available

9. *Location.* Record the location of the institution.

1. Canada
2. United States
3. Europe
4. Other (specify):

10. *Inmate population.* Record the inmate population of the institution. If population information is not available, code "99999".

11. *Crowding indices*. Record an index of the level of crowding within the institution. Select “more” if a density index or author report indicates that the institution was crowded. Select “less” if a density index or author report indicates that the institution was not crowded.

1. More
2. Less
9. Information not available

12. *Program-oriented philosophy*. Specify whether the institution subscribed to a program-oriented philosophy.

1. Yes
2. No
9. Information not available

13. *Institutional climate history*. Code author report of institutional climate history, if available. A “poor” institutional climate refers to an institution that has experienced a history of disturbance and/or frequent staff turnover. A “good” institutional climate refers to an institution with no history of disturbance and infrequent staff turnover.

1. Poor
2. Good
9. Information not available

14. *Offender location*. Were program/intervention participants separated from the rest of the prison population for the duration of treatment?

1. Yes
2. No
9. Information not available

### *Sample Descriptors*

15. *Adult/Juvenile*. Was the sample comprised of adult and/or juvenile participants?

1. Adult (18 + years of age)
2. Juvenile (17 years of age and under)
3. Mixed
9. Information not available

16. *Predominant race.* Choose the code that best characterizes the makeup of the sample with respect to race.

1. >80% white
2. >80% black
3. >80% Hispanic
4. >80% other minority (specify):
5. Mixed (none greater than 80%)
6. Mixed (cannot estimate proportion)
9. Information not available

17. *Predominant gender.* Choose the code that best characterizes the makeup of the sample with respect to gender.

1. >80% male
2. >80% female
3. Mixed (Neither >80%)
4. Mixed (cannot estimate proportion)
9. Information not available

18. *Offender risk at onset of treatment.* Record the risk level of the sample prior to the treatment/intervention.

1. Low
2. Moderate
3. High
4. Mixed
9. Information not available

19. *Offender risk assessment method.* Specify the method used to determine risk level for sample.

1. Actuarial assessment
2. 2 or more previous convictions
3. Author(s) report risk
4. Other (specify):
9. Information not available

*Research Design Descriptors*

20. *Type of design.* Choose the code that best describes the method by which participants were assigned to treatment/intervention and control groups.

1. Random after matching, stratification, blocking, etc.
2. Random, simple (also includes systematic sampling)
3. Random with breakdown (e.g., 20% attrition in experimental or control)
4. Comparison group control (matched on 5 or more risk factors)
5. Comparison group control (matched on fewer than 5 risk factors)
6. Time series design
7. Correlational design
8. Other (specify):

21. *Length of follow-up.* Indicate the length of the post-treatment/intervention follow-up period.

1. >2 years
2. 1 to 2 years
3. 6 to 11 months
4. 3 to 5 months
5. <3 months
9. Information not available

22. *Total sample size.* Start of study.

23. *Treatment group size.* Start of study.

24. *Control group size.* Start of study.

25. *Participant institutional misconduct history.* Is information provided concerning the institutional misconduct history of participants?

1. Yes
2. No

## Treatment Descriptors<sup>10</sup>

26. *Is there enough information to determine the type of treatment strategy utilized?*

1. Yes
2. No

27. *Behavioral treatment strategy.* Indicate the strategy of the treatment/intervention program. If no behavioral strategy was utilized, code as 7 (No behavioral strategy).

1. *Radical behavioral.* If not explicitly stated in the study, this strategy can be inferred from the following: classical conditioning (e.g., desensitization); operant conditioning schedules of reinforcement; token economies; or contingency management.
2. *Social learning.* Code if modeling and behavioral rehearsal techniques designed to engender self-efficacy were employed in the treatment/intervention.
3. *Cognitive behavioral.* Code if any of the following were utilized: Cognitive Therapy; Cognitive Skills training, Problem Solving, Rational-Emotive Therapy; Self-Instructional Training; or Stress Inoculation Training.
4. *Punishment.* Code if the treatment/intervention employed strategies such as fines or loss of privilege.
5. *Mixed (specify):*
6. *Other structural therapy (specify):*
7. *No behavioral strategy*

28. *Non-behavioral treatment strategy.* Indicate the non-behavioral treatment strategy of the treatment/intervention program. If no non-behavioral strategy was used, choose code 12: (No non-behavioral strategy).

1. *Non-directive therapy.* Code if program appeared not to use any direct behavioral training procedures such as rehearsal, conditioning (classic or operant), modeling, role playing, etc.
2. *Psychodynamic.* This category is appropriate for those programs that were of the “talk” therapy, client-centered, Rogerian variety which concentrate on client “insight”.
3. *Group milieu.* Code if program was a group therapy, which did not employ any direct behavioral training procedures (e.g., 12 step program).

---

<sup>10</sup> The Behavioral and Non-Behavioral categories of the Treatment Descriptor Section were derived from: Gendreau, P. (1996). The principles of effective intervention with offenders. In A. T. Harland (Ed.), *Choosing Correctional Options that Work*, (pp. 117-130). London: Sage.

4. *Chemotherapy*. Code if program involved providing some kind of medical intervention (in the absence of behavioral treatment) as a means to reduce institutional misconduct (e.g., medication or plastic surgery)
5. *“Punishing smarter” program*. Code if program involved strategies such as extensive surveillance, frequent drug testing, shock incarceration, etc. to control institutional misconduct.
6. *Wilderness program*. Code if program was of the wilderness or “Outward Bound” variety.
7. *Boot camp*
8. *Recreation program*
9. *Diet program*. Code if program was designed to alter some aspect of participant diet to change behavior.
10. *Mixed (specify):*
11. *Other non-behavioral (specify):*
12. *No non-behavioral strategy*

29. *Educational/vocational treatment strategy*. Record the treatment/intervention as either educational or vocational in nature. If no education/vocation program was used, choose code 4 (No educational/vocational strategy).

1. *Education program*. Code if program was strictly educational.
2. *Vocational program*. Code if program was strictly vocational.
3. *Mixed*. Code if program was a combination of educational and vocational training.
4. *No educational/vocational strategy*

30. *Treatment dosage*. Record the length of time of the treatment/intervention.

1. >5 months
2. >2 months to 5 months
3. 1 to 2 months
4. < 1 month
9. Information not available

31. *Treatment intensity*. Record the total number of hours participants spent in the treatment/intervention program.

1. >100
2. 50 to 99
3. 25 to 49
4. 10 to 24
5. <10
6. Information not available

32. *Control group treatment.* Choose the code that most resembles the nature of the procedure for the control group. If the control group received a different treatment than the experimental group, choose the code that closest resembles the alternate (items 3 to 19).

1. Received no treatment
2. Wait list
3. Alternate treatment radical behavioral
4. Alternate treatment social learning
5. Alternate treatment cognitive behavioral
6. Alternate treatment punishment
7. Alternate treatment mixed behavioral (specify):
8. Alternate treatment non-directive therapy
9. Alternate treatment psychodynamic
10. Alternate treatment group milieu
11. Alternate treatment chemotherapy
12. Alternate treatment punishing smarter
13. Alternate treatment wilderness
14. Alternate treatment boot camp
15. Alternate treatment recreation
16. Alternate treatment diet
17. Alternate treatment mixed non-behavioral (specify):
18. Alternate treatment educational
19. Alternate treatment vocational
20. Alternate treatment mixed educational and vocational
21. Alternate treatment other (specify):
99. Information not available

*Therapeutic Integrity Descriptors.*<sup>11</sup>

33. *Director qualifications.* Record “qualified” if the director of the treatment/intervention was trained in a helping profession (e.g., education, nursing, psychology, social work, with some specialization in the correctional/forensic/legal area) and possessed post BA training, with at least an MA degree. If director had no such training, indicate “Not qualified”.

1. Qualified
2. Not qualified
9. Information not available

---

<sup>11</sup> The Therapeutic Integrity Section is derived from: Gendreau, P. & Andrews, D. A. (2001). *Correctional Programs Assessment Inventory – 2000 (CPAI-2000)*. Ottawa, Ontario: T3 Associates.

34. *Director experience.* Indicate “yes” if the director is reported as having had previous full time direct experience (e.g., case management) with an offender treatment program of any description for at least three years. Otherwise indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

35. *Director conducted program?* Indicate “yes” if the director conducted some aspects of the program that involved direct service delivery to clients and direct supervision of staff. Otherwise, indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

36. *Staff education.* If 75% of program staff had an undergraduate degree and 10% of staff had an advanced (e.g., MA) degree, code “yes”. Otherwise indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

37. *Staff training in area of study.* Code “yes” if 75% of program staff had training in criminal justice, education, nursing, psychology, social work, or specialized fields (e.g., addictions). Otherwise, indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

38. *Staff experience in treatment programs.* Indicate “yes” if 75% of staff had worked in treatment programs with offenders for at least 2 years. Otherwise, indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

39. *Staff personal qualities.* Code “yes” if staff had been selected for: A) relationship skills (e.g., enthusiasm, warmth, respectfulness, flexibility, non-blaming, genuineness, humour, self-confidence, empathy, engaging, reflective, maturity, intelligence); B) belief in offender rehabilitation, commitment to human service; and C) staff believed they had the skills to run the program effectively and that the program would be effective. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

40. *Staff selection based on skill?* Code “yes” if staff was selected according to a majority of the following characteristics: directive, solution-focused, structured, contingency-based, cognitive restructuring, pro-social modeling, effective reinforcement, disapproval, and problem solving. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

41. *Staff supervision.* Code “yes” if staff were assessed yearly on clinical skills related to service delivery and had frequent and immediate access to the program director. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

42. *Ongoing staff training.* Code “yes” if staff received training, at least once yearly, in theory and practice of interventions employed and in skill factors needed to deliver the service effectively. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

43. *Knowledge of risk/needs targeted.* Code “yes” if risk was assessed with a valid instrument. If the risk assessment was not valid, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

44. *Dynamic risk assessment.* Indicate “yes” if dynamic risk factors were assessed with a valid instrument.

1. Yes
2. No
9. Information not available

45. *Risk/need summary.* Code “yes” if participant personal characteristics were summarized as to the level of risk, either qualitatively (high, medium, low) or by using cut-off scores as in the case of the LSI. If not, indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

46. *Risk/need norms.* Indicate yes if the program had generated its own recidivism norms using their measure of choice. Otherwise, indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

47. *Attention to specific responsivity factors.* Indicate “yes” if the program had routinely taken into account the responsivity of offenders to different styles and modes of service, that is, how characteristics of offenders may have interacted with the style and mode of service delivery. Otherwise, indicate “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

48. *Responsivity assessment.* Code “yes” if actuarial measures were used to assess responsivity factors such as cognitive ability (General Aptitude Test Battery), psychopathy (PCL\_R factor 1), anxiety (Speilberger State-Trait Anxiety), and depression (Beck Depression Inventory). Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

49. *Responsivity matching.* Indicate “yes” if staff were assigned to a program best matching their skills and assigned to clients with whom they could work effectively. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

50. *Change measured.* Indicate “yes” if the measure used to assess risk/needs was re-administered periodically and change scores derived. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

51. *Appropriate intermediate targets.* Indicate “yes” if the program predominantly targeted criminogenic needs (dynamic risk factors). Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

52. *Criminogenic vs. noncriminogenic targets.* Indicate “yes” if the number of criminogenic targets exceeded the number of noncriminogenic targets by three or more. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

53. *Program manual*. Indicate “yes” if the program had a detailed program manual. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

54. *Client prosocial task involvement*. Indicate “yes” if clients spent at least 40% of their time (or 50 hours per week) in pro-social tasks (which could also include work or leisure activities or a related program). Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

55. *Intensity/risk*. Indicate “yes” if higher risk participants received the greatest intensity/duration of service. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

56. *Appropriate ratio (rewards: punishers)*. Indicate “yes” if rewards outnumbered punishers by at least 4:1. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

57. *Completion/termination criteria*. Indicate “yes” if criteria were clearly outlined as to when the program terminated for each participant. Progress in acquiring pro-social behaviors and beliefs while in program, or engaging in behavior that seriously jeopardized the safety of staff and/or other clients should have defined termination. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

58. *Self-monitoring*. Indicate “yes” if clients were trained to observe and anticipate problem situations. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

59. *Client prosocial response rehearsal*. Indicate “yes” if clients planned and rehearsed alternative pro-social responses. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

60. *Client behavior practice*. Indicate “yes” if clients practiced new pro-social behaviors in increasingly difficult situations. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

61. *Anti-criminal modeling*. Indicate “yes” if a coping model was being used (a model to display challenges similar to those experienced by the observer but include a self-corrective strategy). The model must have demonstrated the desired behavior in concrete and vivid ways. Participants were rewarded for demonstrating desired behaviors and the model was generally a source for reinforcement rather than always punishing or negative. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

62. *Effective reinforcement*. Indicate “yes” if staff immediately told participants that they liked the type of speech/behavior just exhibited. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

63. *Effective disapproval.* Indicate “yes” if staff immediately told participants why the behavior/speech just exhibited was inappropriate or if the antisocial expressions of the participant were not reinforced (extinction). Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

64. *Problem-solving techniques.* Indicate “yes” if staff aided participants in the identification of problems (focusing on the antecedents, behaviors, and consequences), as well as helped to clarify goals, generate alternative solutions, evaluate options, implement a plan, and evaluate a plan. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

65. *Structured skill learning.* Indicate “yes” if staff identified the skill to be learned, modeled the skill for the participants, provided opportunities for role play of the new skill, and provided feedback. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

66. *Effective use of authority.* Indicate “yes” if staff focused their message on the behavior exhibited and not on the person performing it, if they were direct and specific, used a normal tone of voice, specified choices with attendant consequences for compliance and non-compliance, gave encouraging messages, supported words with action, and remained firm but fair. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

67. *Staff relationship practices*. Indicate “yes” if staff were: open, warm, had respectful communication, were non-blaming, empathic, genuine, and flexible, used humour, were engaging, enthusiastic, and expressed optimism. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

68. *Structuring skills*. Indicate “yes” if staff were solution-focused, and the teaching of skills was structured and directive. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

69. *Motivational interviewing*. Indicate “yes” if staff avoided argumentation, developed discrepancy gently, and supported self-sufficiency. Otherwise, code “no”. If the information is not available, indicate as such.

1. Yes
2. No
9. Information not available

70. *Therapeutic integrity score*. To generate this score, add all of the “yes” responses from the *Therapeutic Integrity* section and record the total (items 33-69).

71. *Criminogenic Needs Targeted*. Record the number of criminogenic needs targeted.<sup>12</sup>

72. *Non-Criminogenic Needs Targeted*. Record the number of non-criminogenic needs targeted.

---

<sup>12</sup> The guidelines for determining criminogenic needs were adopted from Andrews, D. A. & Bonta, J. (1998). *The Psychology of Criminal Conduct* (pp. 354-357). Cincinnati, OH: Anderson.

### *Effect Size Descriptors*

73. *Study ID number.* Identification number of the study from which the effect size is coded. Study identification numbers correspond to reference numbers assigned by the Library Master Software program. If an article contains two or more independent studies, i.e., if the article reports two or more independent outcomes using different samples, then the suffix a, b, c, etc. is affixed to the study ID to distinguish each study within the report. Each study is then coded separately.

74. *Effect size number.* Assign each effect size within a study a unique number. Number multiple effect sizes within a study sequentially, e.g., 1, 2, 3, 4, etc.

### *Dependent Measure Descriptors*

75. *Type of outcome.* Choose the option that best describes the outcome measure employed for the effect size. Note: if a study reports both a measure of violent misconduct and another outcome measure (e.g., non-violent misconduct), code only the violent outcome.

1. Violent/serious misconduct
2. Non-violent misconduct or disciplinary infraction
3. Misconduct not specified
4. Institutional adjustment measure

### *Effect Size Data*

76. *Type of data effect size based on.*

1. Means and standard deviations
2. *t*-value or *F*-value
3. Chi-square (df=1)
4. Frequencies or proportions
5. Correlation
6. Other (specify):

77. *Page number where effect size was found.* Record the page number where the data for this effect size was found.

78. *Raw difference favors (i.e., treatment shows more success for...).*

1. Treatment group
2. Neither (exactly equal)
3. Control group
4. Cannot tell or only statistically significant difference reported

79. *Treatment group sample size (end of study).*

80. *Control group sample size (end of study).*

81. *Treatment group attrition.* Record the difference between the treatment sample size at the start of the study and the treatment sample size at the end of the study.

82. *Control group attrition.* Record the difference between the control sample size at the start of the study and the control sample size at the end of the study.

#### *Means and Standard Deviations*

83. *Treatment group mean.*

84. *Control group mean.*

85. *Treatment group standard deviation.*

86. *Control group standard deviation.*

#### *Proportions or Frequencies*

87. *n of treatment group with successful outcome.*

88. *n of treatment group with unsuccessful outcome*

89. *n of control group with successful outcome.*

90. *n of control group with unsuccessful outcome.*

91. *Proportion of treatment group with successful outcome.*

92. *Proportion of treatment group with unsuccessful outcome.*

93. *Proportion of control group with successful outcome.*

94. *Proportion of control group with unsuccessful outcome.*

*Significance Tests*

95. *t*-value.

96. *F*-value.

97. Chi-square value.

98. Correlation

99. Other Significance test (specify):

100. *Effect size*. Record the calculated effect size. Report to two decimals with an algebraic sign in front: plus if difference favors treatment and minus if difference favors control.