



# Investigating the Relationship between Program Integrity and Correctional Program Effectiveness

Christopher T. Lowenkamp, Ph.D., Assistant Director  
University of Cincinnati Corrections Institute  
Edward J. Latessa, Ph.D., Professor & Head  
Division of Criminal Justice  
University of Cincinnati

*This study adds to the growing body of corrections literature that examines the relationship between correctional program integrity and effectiveness. Measures of program integrity developed for the Correctional Program Assessment Inventory (such as staffing, program implementation, treatment activities, etc.) were used. Program integrity is shown to be associated with increased program effectiveness.*

## INTRODUCTION

Over the past several decades, research in corrections has focused on identifying particular types of programming that are effective in reducing recidivism. This research consists of a large body of literature that theoretically and empirically identifies behavioral and cognitive-behavioral strategies as the most promising in reducing criminal behavior. Recently research in corrections has begun to focus on the relationship between correctional program integrity (staff, implementation, treatment activities, etc) and effectiveness. One method of assessing program integrity that is gaining acceptance as the standard is the Correctional Program Assessment Inventory (CPAI) (Van Voorhis and Brown, 1996). While the CPAI is widespread in terms of its use, there is little research that empirically tests the CPAI's relationship with program effectiveness. The purpose of the current research is to determine if CPAI scores are related to program effectiveness with a sample of halfway houses (HWH) in Ohio.

## CPAI

The initial version of the CPAI contains 65 items that are used to rate six substantive areas of correctional programs (Gendreau & Andrews 1994). These areas are program implementation, client pre-service assessment, characteristics of the program, characteristics of practices and staff, evaluation, and miscellaneous. Each section contains anywhere from 6 to 22 items. Each item is scored based on the

presence or absence of what it is the item is measuring.

To date, hundreds of assessments with the CPAI have been conducted on correctional programs in the United States and Canada. Data indicate that overall, most programs are in need of improvement with many programs failing (scoring less than 50%), and only a few scoring in the very satisfactory range (70 to 100%). In a study by Latessa and Holsinger (1998), the authors review the scores of 51 programs assessed with the CPAI. The programs, on average, scored 56% on the CPAI. Most programs (60%) scored in the satisfactory but needs improvement or unsatisfactory range with only 12% scoring in the very satisfactory range.

Similarly, Matthews, Hubbard, and Latessa (2001) report on the CPAI assessments of 86 correctional programs. The data reported by Matthews et al. (2001) indicated again that a small percentage of the programs (10%) scored very satisfactory, 54% scored in the satisfactory or needs improvement range, and 35% scored in the unsatisfactory range.

Reviews of CPAI scores by other researchers have indicated a similar trend. For example, Gendreau and Goggin (2000) report on an early research effort in 1991 that assessed 101 correctional programs. These programs produced a mean CPAI score of 25%. Only 10% of the programs in this review received a satisfactory score. Another analysis of CPAI scores by Hoge, Leschied, and Andrews (1993) summarized the CPAI scores of 135 juvenile programs. This review again noted an average program integrity score that was failing (35%). Only 10% of the programs in Hoge's review received a CPAI score of satisfactory or better.

Given the vast number of CPAI assessments completed it is becoming clear that the field of corrections has, to some degree, realized the importance of program integrity assessment. The void that still exists in this area is empirically linking program integrity, as measured by a standardized

assessment like the CPAI, to program outcomes. As Gendreau, Goggin, and Smith (2001:260 fn 17) note "The ultimate validity, based on correlations between the CPAI scores and the recidivism rates associated with the programs being assessed, is a long way off." The only research published to date that uses the CPAI as a predictor of success rates of offenders that were served by the assessed programs are those by Gray (1997) and Holsinger (1999).

Gray (1997) used a 20-item abbreviated version of the CPAI to assess program quality from sixty-seven previously published evaluation studies on community-based interventions. The correlation ( $r$ ) between this abbreviated version of the CPAI and program outcome was 0.41.

Holsinger (1999) investigated the relationship between program integrity and program outcome for nine community-based juvenile correctional facilities in Ohio. The results of these analyses indicated that the CPAI score was a significant predictor of many post release outcome measures including re-arrest, arrest for a personal offense, and a new adjudication.

The analyses contained in this research add to the literature based on this topic in a number of ways. First it extends the application of the CPAI as a predictive measure to HWHs for adult parolees. Second, this research uses a more complete version of the CPAI when compared to that used in some previous research. Third, this research includes a sample of 38 programs where a treatment and comparison group were developed and the CPAI was applied. This is an advantage as the current research is not limited to the information reported in a primary study on each of the 38 programs.

## METHODS

In order to investigate the relationship between the CPAI and program effectiveness, data on program integrity and program effectiveness were collected for each of the 38 HWH programs in this sample. To develop data on program integrity individuals trained in the application of the CPAI conducted visits to each program. The CPAI<sup>1</sup> was scored for each program based on interviews with the program director, surveys of staff, and a review of program materials. The following sections provide some additional information on the measures used in these analyses, however, for a more detailed description see Lowenkamp (2004).

### Program Effectiveness

The measure of program effectiveness captures the difference in recidivism rates between an experimental group and a comparison group for each of the 38 HWHs included in this research. The

treatment group for each program contains offenders released from a state institution to one of 38 HWHs as part of their parole/post release control (PRC) or transitional control. The HWH treatment group included 3,237 offenders that were terminated from halfway houses during fiscal year 1999. The parolees/PRC releases were matched to parolees/PRC releases that were under parole supervision but were not placed in a HWH. The comparison cases (3,237) were drawn from a larger sampling frame ( $N=6,781$ ) and were matched with the treatment cases on county of conviction, sex, and risk level as determined by a modified version of the Salient Factor Score (see Lowenkamp, 2004 for a listing of these factors). Cases were further matched by crime type if the treatment case was coded as a sex offender.<sup>2</sup>

The log odds ratio was calculated based on the recidivism data for the treatment (only successful terminations) and comparison groups. The measure of recidivism used in this research was return to an Ohio Correctional Facility for any reason (technical violation or new arrest). The decision to use this measure over others was based on the unreliable nature of the arrest and conviction data available. While incarceration is a conservative measure of future criminality, the data on prison intakes is quite complete and easy to query. While the log odds ratio is not easily understood, it is suited well for the analyses of dichotomous data. Further, mathematically the log odds ratio can readily be converted to percentage point differences in recidivism (see Lipsey and Wilson, 2002).

### Program Integrity

The initial version of the CPAI contains 65 items across six substantive areas. Over the past several years the number of items on the instrument has increased. However, the substantive areas covered have remained the same. These areas are program implementation, client pre-service assessment, characteristics of the program, characteristics of practices and staff, evaluation, and miscellaneous. Each section contains anywhere from 6 to 26 items. Each item is scored based on the presence or absence of what it is the item is measuring.

The items on the CPAI are scored in such a fashion that one point is assigned for each characteristic that is observed. The total number of points in each section is divided by the total number of questions in each section providing a percentage score for each section. To obtain the overall score, all points from all sections are added together and then divided by the total number of items. The sections on the abbreviated version of the CPAI vary in the number of items from 5 to 19. Some items

deemed not applicable are not scored and are also removed for the base number of questions when calculating the percentages.

### Design And Analyses

The analyses in this research are fairly simple and straightforward. Correlations were calculated between the overall CPAI score and the measure of program effectiveness. In addition, categories of programs were developed based on their CPAI scores. The average of the program effectiveness of each category was then calculated. All analyses were weighted by the inverse variance weight of the log odds ratio which controls for differences in the number of offenders terminated at each program. This process gives more weight to the programs with larger samples and more stable rates of success.

## RESULTS AND DISCUSSION

Table 1 contains the treatment effect sizes for the 38 programs. With this measure 61% (23 out of 38) of the programs were associated with treatment effects that favored the treatment group.

**Table 1. Logged Odds Ratios and Weights Using Successful Terminations Only**

Program	Logged Odds	
	Ratio	Weight
Program 1	-1.17	2.42
Program 2	-1.08	9.96
Program 3	-0.89	5.45
Program 4	-0.67	1.49
Program 5	-0.65	13.71
Program 6	-0.62	0.83
Program 7	-0.35	5.64
Program 8	-0.32	3.15
Program 9	-0.21	4.68
Program 10	-0.15	6.67
Program 11	-0.10	0.42
Program 12	-0.07	7.25
Program 13	-0.05	9.56
Program 14	-0.05	0.45
Program 15	-0.03	0.47
Program 16	0.06	8.68
Program 17	0.07	7.43
Program 18	0.15	6.79
Program 19	0.23	2.09
Program 20	0.24	1.98
Program 21	0.29	3.41
Program 22	0.31	28.78
Program 23	0.33	8.98
Program 24	0.37	1.29

Program	Logged Odds	
	Ratio	Weight
Program 25	0.42	5.90
Program 26	0.44	13.53
Program 27	0.55	1.75
Program 28	0.61	16.94
Program 29	0.80	0.51
Program 30	0.82	3.51
Program 31	0.95	2.47
Program 32	1.03	12.62
Program 33	1.26	0.64
Program 34	1.30	4.42
Program 35	1.50	0.74
Program 36	2.16	0.40
Program 37	3.00	0.33
Program 38	4.48	0.40

The effect sizes calculated range from -1.17 to 4.48<sup>3</sup> with an average of 0.15 and a 95% confidence interval ranging from 0.01 to 0.29. This roughly equates to a four percentage point reduction in recidivism favoring the treatment group. This effect size would be considered fairly modest by most standards; however, one must consider that the four percentage point reduction assumes a 50% recidivism rate. Therefore, the four percentage point reduction equates to an 8% relative risk reduction. Nonetheless, given the fields understanding of effective correctional interventions, funding sources should be able to expect more than a four percentage point reduction from the programs they are funding. Ultimately, this research attempts to determine if abbreviated CPAI scores relate to treatment effectiveness.

Table 2 provides the descriptive statistics for the abbreviated CPAI measure. The average score for the 38 programs is 47 with a range of 33 to 69 and a standard deviation of 8. The patterning of scoring with this sample appears to be similar to the patterns of scoring noted in previous research on the CPAI. The next step in the analyses was to calculate weighted correlations between the CPAI and the log odds ratio.

The correlation between the CPAI and the log odds ratio is 0.32 and is significant at the  $p \leq 0.10$  level. While this is a moderate correlation further categorizing programs based on their CPAI does lead to a pragmatic grouping of programs that assists in the prediction of effective and ineffective correctional programs. The 38 programs were placed into three groups based on their percentage score for the CPAI. Those programs that fell under the threshold for an "unsatisfactory" score were placed in the first category. This category includes a total of 24 programs. The second category contains those

programs that would have attained score of satisfactory but needing improvement. A total of 13 programs fall into this second category. Finally, 1 program was categorized as satisfactory with a score between 60 and 69%. The average percentage point reductions in recidivism are displayed by category in Figure 1. Recall that these percentage point reductions assume a 50% recidivism rate in the comparison group. Therefore, the percentage point reductions should be doubled to arrive at relative reductions in risk.

**Table 2. Descriptive Statistics for Abbreviated CPAI Score**

Abbreviated CPAI Score	
Mean	47
Low	33
High	69
Standard Deviation	8

As Figure 1 indicates, the higher the CPAI score the larger the reduction in recidivism. While the programs in the lowest category for program integrity average 1.7 percentage point reductions in recidivism, the middle grouping of programs is associated with an average percentage point reduction of 8.1. The highest category of programs (N = 1) is associated with a 22 percentage point reduction in recidivism. When considering relative reductions in risk, these differences equate to 3.4, 16, and 44% reductions.

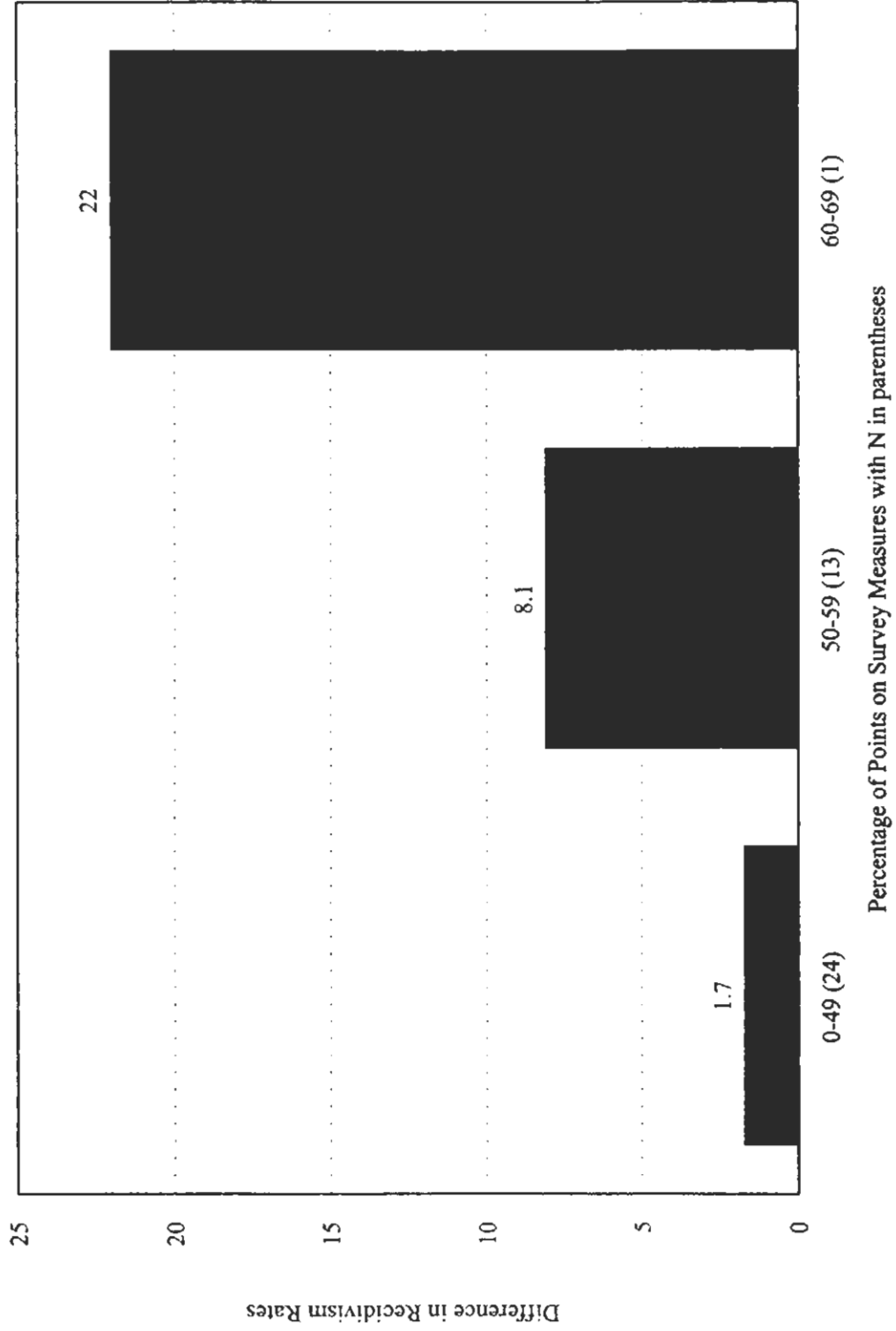
This research is limited in a number of ways. First, there was only one program that scored in the "satisfactory" range. Having only one program in this category causes the average reduction in recidivism to be unstable. It is unknown for sure that such reductions would remain so high for this category if additional programs reached that range of

the scale. In addition, the sample includes only HWH in Ohio, the treatment and comparison cases included parolees or offenders on other forms of post-release supervision only, the samples were limited to adult offenders, and were predominantly male. In spite of these limitations, this research indicates that in Ohio, HWH that serve parolees are much more effective when they adhere to the principles of effective interventions. Reductions in recidivism are 4 and 10 times larger when programs attain scores of 50% and above on the abbreviated version of the CPAI used in this research.

These findings are important for a number of reasons. First, this research indicates that there is in fact a relationship between correctional program integrity and effectiveness. Second, this research provides funding agencies with information that should assist in predicting how effective a proposed or existing intervention might be based on the criteria used to rate the CPAI items. Third, in finding support for the notion that program integrity is important, it provides program designers with a blueprint for designing and implementing new programs. Finally, the findings of this research indicate that the CPAI process can be used to identify deficiencies in programming and that those deficiencies are associated with program effectiveness. Addressing these deficiencies might also be associated with increased effectiveness.

While this research is important for correctional programming, it should be expanded to include different types of programs (non-residential and other residential programs), programs that provide services to juvenile offenders, other types of offenders (e.g. probation samples), and across geographic settings. Finally, future research should investigate how the differing components of program integrity impact program effectiveness.

Figure 1. Difference in Recidivism Rates Between Treatment and Comparison Group Based on Abbreviated CPAI Score



## ENDNOTES

<sup>1</sup> Due to the limited nature of the interviews with staff and program directors, the number of items to be scored on the CPAI will be reduced from 77 to 62.

<sup>2</sup> The experimental and comparison groups were contained in two separate files. A program was written that recorded the county of conviction and sex of the offender for the first case in the experimental group. The program then queried the comparison group file and found an offender that matched on county of conviction and sex. This matching case was then marked as a comparison case for the offender in the experimental group. If the experimental group offender was a sex offender, the program matched on county of conviction, sex of the offender, and current offense type. The program then went on to the second comparison case and continued iterating until no further matches could be made.

<sup>3</sup> Positive numbers in Table 1 indicate programs where the treatment group outperformed the comparison group. Negative numbers in Table 1 are associated with programs where the comparison group had a lower recidivism rate than the treatment group.

---

## REFERENCES

- Gendreau, P. & Andrews, D.A. (1994). *The correctional program assessment inventory*.
- Gendreau, P. & Goggin, C. (2000). Correctional treatment: Accomplishments and realities. In Patricia Van Voorhis, Michael Braswell, & David Lester (Eds.), *Correctional counseling & rehabilitation 4<sup>th</sup> edition*. (pp. 271-280). Cincinnati, Oh: Anderson Publishing.
- Gendreau, P., French, S., & Taylor, A. (2002). What works (what doesn't work)-revised 2002: The principles of effective correctional treatment. University of New Brunswick at Saint John: Unpublished Manuscript.
- 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 and Practice*. (pp. 247-268). New York: John Wiley and Sons, Ltd.
- Gray, G. A. (1997). *Does coercion play a significant role in community treatment programs that reduce offender recidivism?* Master Thesis. The University Of New Brunswick.
- Hoge, R., Leschied, A., & Andrews, D.A. (1993). *An investigation of young offender service in the province of Ontario: A report of the repeat offender project*. Toronto, Ontario: Ministry of community and Social Services.
- Holsinger, A. M. (1999). *Opening the 'black box': Assessing the relationship between program integrity and recidivism*. Doctoral Dissertation. The University of Cincinnati.
- Latessa, E. J., & Holsinger, A.M. (1998). The importance of evaluating correctional programs: assessing outcome and quality. *Corrections Management Quarterly*, 2(4):22-29.
- Lowenkamp, C.T. (2004). *Correctional program integrity and treatment effectiveness: A multi-site, program-level analysis*. Doctoral Dissertation. The University of Cincinnati.
- Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage Publications.
- Matthews, B., Jones-Hubbard, D. & Latessa, E.J. (2001). Making the next step: Using evaluability assessment to improve correctional programming. *The Prison Journal*, 81, 454-472.
- Van Voorhis, P. & Brown, K. (1996). *Risk classification in the 1990s*. National Institute of Corrections.